

Data Workbench

Contents

Data Workbench Help.....	5
Data Workbench Release Notes.....	5
Data Workbench 6.73 Release Notes.....	5
Data Workbench 6.72 Release Notes.....	6
Data Workbench 6.7 Release Notes.....	6
DWB Report Portal 2.1.....	8
Data Workbench 6.61 Release Notes.....	8
Data Workbench 6.6 Release Notes.....	9
Data Workbench 6.52 Update.....	11
Data Workbench 6.51 Update.....	12
Data Workbench 6.5 Release Notes.....	14
Data Workbench 6.4 Release Notes.....	18
Data Workbench 6.31 Update.....	41
Data Workbench 6.3 Release Notes.....	43
Data Workbench 6.21 Update.....	79
Data Workbench 6.2 Release Notes.....	81
Data Workbench 6.11 Update.....	108
Data Feeds Update for April MR 2014.....	109
Data Workbench 6.1 Release Notes.....	123
Data Workbench 6.0 Release Notes.....	144
Installing Data Workbench.....	172
Workstation requirements.....	172
Workstation setup.....	174
Upgrading Data Workbench.....	185
Data Workbench Client.....	198
Opening Data Workbench.....	198
Self-provisioning of users.....	200
Working offline and online.....	201
Profiles.....	202
Worktops.....	203
File versioning.....	205
Using the Worktop thumbnail menu.....	206
Configure the sidebar.....	209
Display workspaces.....	212
Keyboard shortcuts for the Data Workbench Client.....	224
Working with windows in a workspace.....	226
Exporting data using segment export.....	230
Analytics Reports.....	261
Using attribution reports.....	262
Visualizations.....	279
Analysis visualizations.....	304
Imagery layers.....	432
Administrative interfaces.....	446
Interface and analysis features.....	484
Configuration parameters.....	510
Query language syntax.....	515
Dataset Configuration.....	524
Understanding Dataset Construction.....	524
Understanding Dataset Configuration.....	526
Log Processing Configuration File.....	528
Transformation Configuration File.....	551
Dataset Include Files.....	557
Additional Configuration Files.....	570
Data Transformations.....	573
Extended Dimensions.....	606
Transform Functionality.....	617
Reprocessing and Retransformation.....	627

Dataset Configuration Tools.....	628
Conditions.....	634
Regular Expressions.....	642
Event Data Record Fields.....	646
Configuration Settings for Web Data.....	649
Time Zone Codes.....	658
Data Workbench Report Server.....	659
Understanding Report Sets.....	660
Generating Reports.....	660
Distributing Reports.....	660
Report Set Requirements.....	661
Working with Report Sets.....	661
Using the Report Portal.....	669
Advanced Search in the Report Portal.....	670
Installing Report Server.....	671
Administering Report Server.....	681
Installing the Report Portal.....	688
Report Server Parameters Reference.....	701
Report Sample Files.....	708
Data Workbench Page Tagging.....	711
What Kind of Data Can I Acquire?.....	711
How Do I Acquire this Data.....	712
Implementing Reference Page Tags.....	712
Acquiring Baseline Measurements.....	715
Acquiring Extended Measurements.....	721
Tracking Visitor Activity Within Flash Rich Media Content.....	728
P3P Considerations for Third-Party Page Tagging.....	730
Page Tag Variable Mapping Worksheet.....	731
Data Workbench Metrics, Dimensions, and Filters.....	732
Expanding Metrics, Dimensions, and Filters.....	734
Dimension Types.....	734
About Built-in Metrics, Dimensions, and Filters.....	735
Baseline Profiles.....	737
Data Workbench Geography Overview.....	743
Installing Data Workbench Geography.....	744
Working with Imagery Layers.....	746
Working with Data Services.....	760
Upgrading and Uninstalling Data Workbench Geography.....	765
Data Workbench Controlled Experiments.....	767
How Does Site Identify Visitors?.....	767
How Do Controlled Experiments Work?.....	767
What Should I Test?.....	768
What Are the Requirements?.....	769
Enabling Controlled Experimentation.....	769
Creating a Controlled Experiment.....	771
Validating the Experiment.....	777
Viewing the Results.....	778
Experiment Design Spreadsheet.....	783
Server Administration and Installation.....	787
Server System Requirements.....	787
Installing Data Workbench Servers.....	792
Administering Data Workbench Server.....	826
Transform.....	845
Insight Server Replication Service.....	847
Repeater Functionality.....	850
General Systems Administration.....	853
Upgrading and Uninstalling Adobe Software.....	854
Configuration Settings Reference.....	869
Time Zone Codes.....	874
Sensor Overview.....	875
How Does the Data Collection Process Work?.....	875
What are Basic Components.....	876
How Does Sensor Identify Visitors and Sessions?.....	877

Installing Sensor.....	.878
Administering Sensor.....	.936
Sensor Txlogd.conf File Parameters.....	.945
Event Data Record Fields.....	.951
Sample Configuration Files.....	.954
Sensor Transmitter Command-Line Options.....	.974
Sensor UNIX File Permissions.....	.974
Installing the Monitoring Profile.....	.976
Workspaces for Monitoring the Data Workbench Server978
Data Workbench Profile Dimensions and Metrics.....	.986
Data Workbench Dashboard User Guide.....	.995
Business Opportunities.....	.996
Features.....	.996
Glossary.....	.996
Getting Started.....	.997
Visualizations.....	.999
Dashboards.....	.1016
Making Selections within the Dashboard.....	.1027
Exporting Data.....	.1033
Dashboard Access Controls.....	.1036
User Account Management.....	.1036
Help Menu.....	.1037
Data Workbench Dashboard Administrator Guide.....	.1038
Minimum System Requirements.....	.1038
Download Required and Optional Components.....	.1038
Supported Browsers.....	.1038
Installation Prerequisites.....	.1039
Dashboard to Data Workbench Servers Communication.....	.1040
Dashboard Deployment.....	.1041
Managing from the Administration Console.....	.1043
Managing Profile Connections.....	.1045
Managing the Schema.....	.1047
Managing Folders.....	.1048
Managing User Accounts.....	.1051
Managing User Groups.....	.1053
System Logs.....	.1055
Implementing Adobe Data Workbench.....	.1056
Data Workbench Discovery and Requirements.....	.1057
Data Workbench Installation and Provisioning.....	.1059
Data Workbench Architecture.....	.1061
Data Workbench Administration.....	.1065
Data Workbench Configuration and Implementation.....	.1065
Data Workbench Implementation of Features.....	.1086
Data Workbench Glossary.....	1101
Third-Party Licenses.....	1110

Data Workbench Help

Data Workbench is part of *Adobe Analytics*. It has highly-flexible, powerful, multi-channel capabilities designed to collect, process, analyze, and visualize data from online and offline customer interactions across all types of web, CRM, POS, and other data types. Interactive reports in Data Workbench let you analyze and react to pertinent statistical information to monitor activity and predict customer trends.



[Download All Data Workbench documentation](#) in PDF.

New and Popular Topics

Data Workbench 6.6 is now released. It includes updates to Propensity Score and a new Segment Export Wizard.

Data Workbench 6.52 is released. It includes an AVRO upgrade for processing data source log files.

Data Workbench 6.51 is released. It includes a new AVRO Feed and gradient overlays to the Bar Graph display.

Data Workbench 6.5 is released. Install the latest version of Data Workbench with new features, including

- [Associations Table](#) and [Associations Chord](#) visualizations.
- [Regression Tree option](#) and other updates for Decision Trees.

Data Workbench 6.4 includes:

- [Exporting to Analytics Core Services](#)
- [Workstation Setup Wizard](#)
- [Presentation Layer for Annotating Visualizations](#)
- [Metric Dim Wizard](#)

Data Workbench also includes these predictive analytics features, data visualizations, and reports:

- [Best Fit Attribution](#)
- [Propensity Scoring](#)
- [Decision Trees](#)
- [Correlation Matrix](#)
- [Clustering](#)
- [Chord Visualization](#)
- [Attribution Profile](#)
- [Master Marketing Profile Export and Integration](#)

Last updated: 8/2/2016

Data Workbench Release Notes

Identify new features, upgrade instructions, and bug fixes for released versions of Data Workbench.

Data Workbench 6.73 Release Notes

New features and fixes in Data Workbench 6.73.

Fixes

- Fixed an issue in Workstation where users were unable to sign in on some hardware with high resolution and high DPI.
- Fixed an issue in the server where Email was missing in Archive file names when using IMS login.
- Updated OpenSSL to version 1.1.0h which includes several vulnerability fixes and new SSL Ciphers.
- Updated the below listed open source libraries to the latest stable versions
 - libssh2 1.8.0
 - Apache Xerces 3.2.1
 - Apache Xalan 1.11
 - libpng 1.6.34
 - libarchive 3.3.2
 - zlib 1.2.11
 - pcre 8.42
- Added error logging when Lookup file row count exceeds more than supported 357913908 rows.

Known issue

- Data Workbench Workstation version 6.73 does not connect to Data Workbench Servers versions 6.61 and older. The reason is, older server versions use a weak form of ciphers not supported in version 6.73. To enable support for older versions
 1. Override default SSL Ciphers list on the server with a strong cipher list supported by OpenSSL version 1.0.1h. To override, add key ‘SSL Ciphers’ in the ‘Communications.cfg’ files available in ‘Components’ and ‘Components for Processing Servers’ directories. For example: `SSL Ciphers = string: !aNULL:AESGCM`



Note: Ensure that the key is placed at the same level as the SSL Port. For details refer to [Communications Configuration Settings](#)

2. Place the latest trust_ca_cert.pem file on the server 6.61 and older servers. This setting is applicable to all Workstation 6.7x versions.

See [archived release notes](#) for Data Workbench 5.3 to 5.52.

Data Workbench 6.72 Release Notes

New features, fixes, and known issues in Data Workbench 6.72.

Fixes

- Fixed an issue in Workstation where users were unable to use Attribution when logged in with username and password.
- Fixed an issue in server where Email was missing in AUTHENTICATE audit logs.
- Fixed an issue in server where Client information was missing in Query Log file names when using Certificate and IP address.

See [archived release notes](#) for Data Workbench 5.3 to 5.52.

Data Workbench 6.7 Release Notes

New features, fixes, and known issues in Data Workbench 6.7.

New Features in Release 6.7

New authentication model for Data Workbench Workstation (IMS integration)

Data Workbench Workstation now supports user authentication through username and password. With this new method, administrators can create and manage their own user accounts, eliminating the need to contact Customer Care.

For more information, see [Self-Provisioning of Users](#).

Flat file lookup

Data Workbench Workstation now supports user authentication through username and password. With this new method, administrators can create and manage their own user accounts, eliminating the need to contact Customer Care.

Flat file lookup previously loaded the entire file into in-memory buffers, bloating memory usage and creating performance issues for other subsystems. The files can now be memory mapped and cached in Windows, optimizing memory usage by setting *Memory Mapped Lookup Files* to true in `MemorySettings.cfg`.

For more information, see [Self-Provisioning of Users](#).

Memory usage

Large Page usage can now be disabled by setting *Use Large Pages* to false in `MemorySettings.cfg`. See [Monitoring Memory Usage](#) for more information.

Security ciphers

Added support for ECDHE and DHE.

Email support in User List.cfg

Added support for Email attribute in `User List.cfg`. For more information, see [User Administration of Group Members](#).

Help Menu

Help menu now shows a shortcut to Open Certificates directory.

TargetBulkUpload export

URLs will be provided at the end of the export trace file and the `targetbulkuploadexportname.log.completed` file to track the record of stuck batches.

A new file, `TargetBulkUpload.cfg`, has been provided to configure the Max Timeout interval (in minutes). The file is found in `Server\Admin\Export\`.

Fixes

- Fixed an issue where the Campaign Clickthrough dimension was showing inflated values.
- Fixed an issue with generating excel files from the report server.
- RC4 cipher is now disabled by default.
- Fixed an issue causing the Data Workbench workstation to crash when adding a dimension element to a value legend table.
- Fixed an issue with Data Workbench to AAM exports that was causing timeouts.
- Fixed an issue causing the Data Workbench workstation to crash when a user without sufficient access level saved the workspace to Server.
- Fixed an issue with the date format in `report.cfg` being incorrect or not localized.
- Fixed an issue with the mobile and product rows in the AVRO feed displaying confusing information.
- Fixed an issue that prevented the ordering of `*.1cd` and `*.1ad` files in `order.txt`.
- Submit To Server option has been disabled for Expectation Maximization algorithm while running Clustering.

-
- Fixed an issue with the TargetBulkUpload executable stalling and failing to run completely.

Known Issues

- On log out, the user cache.db file is cleaned.
- IMS user email addresses containing '+' or '%' characters are not supported.
- User is unable to logout during an error in connection status. As a workaround, logout in offline mode.
- IMS login window does not render properly on some hardware with high resolution and high DPI. As a workaround, right-click on Insight.exe and navigate to **Properties > Compatability**, then check the box **Override high DPI scaling behavior**.

Upgrade Requirements

1. Update trust_ca_cert.pem on the Insight Servers which is part of build package.
2. Update Server's and Report Server's certificate by downloading new certificates from <https://aap.adobe.com>.
3. To automatically update Workstation and Report Server, manually update trust_ca_cert.pem for both by downloading it from the License Server.
4. Sensor's automatic update feature requires version 5.0 in order to communicate with Insight Server version 6.70. Also, Sensor's trust_ca_cert.pem must be updated manually by downloading it from the License Server.

System Updates

New files include:

1. Server\Admin\Export\TargetBulkUpload.cfg
2. Server\Components for Processing Servers\MemorySettings.cfg
3. Server\Components\MemorySettings.cfg

Updated files include:

1. trust_ca_cert.pem for all components.
2. Access Control.cfg to support IMS configuration.
3. Base\Context\meta.cfg for supporting Start Date and End Date formats in Report.cfg
4. Additions to Insight.cfg to support proxy for IMS authentication:

```
IMS Proxy Info = IMSProxyInfo:  
    Proxy Password = EncryptedString:  
    Proxy User Name = string:
```

See [archived release notes](#) for Data Workbench 5.3 to 5.52.

DWB Report Portal 2.1

Security update for Data Workbench Report Portal.

Important Security update

The Report Portal now features stronger hashing algorithms with salting support. If you are upgrading to Report Portal 2.1, add a new Text field, PasswordSalt with field size of 20 characters in users.mdb database. This field is required to store the password salt.

Data Workbench 6.61 Release Notes

Data Workbench 6.61 release notes include new features, upgrade requirements, fixed bugs, and known issues.

New Features

Data Workbench 6.61 includes these new features:

Updated MMP configuration for splitting record files

For the Data Workbench client, you can now split files containing records. In the `MMPExport.cfg` file, you can specify the number of records per split. The incoming records file is then split accordingly, and then record chunks are exported to AWS S3.

See [MMP Integration](#).

On selecting a particular localization profile, the month and day in the calendar are now localized.

System Updates

These features have been updated in this release:

- Page Views metric in Adobe SC and Adobe SC Avro profiles are no longer Page Event based. Page Views are recorded if the Page Name or Page URL is not null.

Fixed Bugs

- MetricDim format is enhanced to use the scale and size parameters. The scale and size parameters are set correctly based on bucket start and bucket size values respectively.
- Fixed Globe issue that was introduced in release 6.6. Errors occurred because of this issue when users tried to zoom-in on certain locations on the Globe.

[Additional Data Workbench Documentation online](#)

Data Workbench 6.6 Release Notes

Data Workbench 6.6 release notes include new features, upgrade requirements, fixed bugs, and known issues.

New Features

Data Workbench 6.6 includes these new features:

Smart Feature Selection

The Smart Feature provides automated feature selection when using the Propensity Scoring visualization. You no longer need to select the independent variables from a long list of metrics and dimensions.

Select **Options > Feature** to enable the feature selection menu. If the **Feature Reduction** option is not selected, you can make the system explore additional model inputs from the entire set of metrics while retaining user-added metrics. Checking **Feature Reduction** will make the system select the most relevant metrics from the user metrics.



Note: While using Smart Feature Selection for an empty model, **Feature Selection Settings** are redundant.

Selecting **More Features** or **Less Features** from the **Level Setting** menu makes the selection algorithm add feature sets to the model. In some cases both setting types may return the same model.

Segment Export Wizard

The [Segment Export Wizard](#) provides a step-by-step process to configure and export segments.

New export executable for Target exports

To export Adobe Target data, a new `TargetBulkUpload.exe` has been developed to replace the `ExportIntegration.exe` for Target exports. If a user wants to move legacy exports to `TargetBulkUpload.exe`, then exports have to be changed (change `mbox3rdPartyId` to `thirdPartyId`).

L4 Logging

An option for **L4 Logging** has been added under **Server > Admin > Export** menu ,as well as a new configuration file (`loggingDetails.cfg`). This allows you to turn GE and L4 logging on and off. By default, L4 Logging will be set to `true` and GE Logging will be set to `false` in the configuration file, but you can change these settings as needed. This is only for exports using the `exportintegration.exe`.

New Presentation Layer options

You can now set transparency and border thickness for Circle and Rectangle shapes within the Presentation Layer. Right-click on the object and select **No Fill** to make it transparent. Four border thickness options are present under **Outline**—Thin, Medium, Thick, Thickest.

New options for Correlation and Association Chord diagrams

New grouping options have been added in the Correlation Chord and Association Chord diagrams. The right-click menu now includes **Add Group**, **Remove Group**, and **Add to Group**.

Grouping will display once you populate your chord. All elements part of the same grouping will be based on the group's color. You can set the group's color like you would set an element's colors. If a grouping is deleted, then the elements are moved to Ungrouped.

Set transparency and border thickness for annotations

You can now set the transparency and border size of circles and rectangles to [Annotating visualizations](#) on page 296 in a workspace. New options for the circle and rectangle include a **No Fill** option for transparency and border thickness settings of **Thin**, **Medium**, or **Thickest**.

XML Decoder enhancements

You can now use the `#value` tag in XML paths to pull the value of an XML element. For example, you can now read the attribute value of the `<Hit><Page name="Home Page" index="20">home.html </Page></Hit>` tag using a `Hit.Page.#value` expression. See [XML Decoder Groups](#) on page 563 for additional information.

New features in Propensity Score

Within the toolbar button you can now choose to either **Save a filter** or automatically launch Propensity Scoring with a new feature selection capability. The Propensity Scoring opens the builder with the workspace selections applied for the dependent variable, enables the selected features, and lets you hit **Go** to start.

It is now possible to save a Propensity Score Model as a JavaScript file. After the model is built click on **Save > Save JavaScript**.

Additional Visit level predictive builders

Added new predictive builders with Visit as the root countable dimension for Propensity Scoring, Clustering, and Decision Trees.

Additional Latency visualizations

Added a set of Latency visualizations. The former latency visualizations looked at an equal period around the event (-7 to +7 days) to create an assignment. This works well for only one event per clip, but doesn't give adequate results when there are multiple events, such as multiple conversions.

We have reorganized and added additional latency visualizations to the menu that are configured to evaluate latency before an event, after an event, and before and after a single event.

Updates to `report.cfg`

You can now view dates in calendar format in the `report.cfg` configuration tree.

Chord Visualization updates

Implemented the ability to scale in the [Chord Visualization](#) on page 60 and [Pie Chart](#) on page 390 visualizations. You can scale by holding down <Ctrl> key and using the mouse wheel, or by right-clicking on the window border, selecting a value to scale, and entering a value.

Access to Detailed Status

Sensor 5.0 supports:

- 2048 bits certificates signed using SHA256.
- Support for Apache Server 2.4 on Linux, Sun Solaris, or FreeBSD
- Support for Apache Server 2.4 on Windows Server 2000 or later

Access to **Detailed Status** is now restricted to Power Users and Administrators by default.

System Updates

These features have been renamed, deleted, or the installation files or folders were restructured in this release:

- Default extension for a **reportserver.cfg** file is now **.xlsx**.
- Additions to **Insight.cfg** (the configuration tree).
 - printf format: You can now set the `printf` format from the **Insight.cfg** tree.
 - V3D: You can now set a V3D setting using three associated edit boxes.
 - Date and Time: You can now set the date and time from a dropdown menu.
- The Page Overlay feature has been removed from the options menu for a URL table element. If you have an older workspace with a page overlay visualization and click on the **Refresh Page** menu option, a dialog will appear stating: *This feature has been deprecated due to security concerns.*

Fixed Bugs

The following are prominent fixes made in Data Workbench 6.6 (since the release of Data Workbench 6.52).

- Fix to Best Fit Attribution (algorithmic attribution): Coefficients were all zeros when SGD was used to converge. This has been fixed.

Known Issues

- The **targetbulkupload.exe** requires command-line arguments in English. The output file name should be in English for Adobe Target Export, Profiles and Audiences Export, and Customer Record Service Export.

Data Workbench 6.52 Update

Data Workbench 6.52 release notes.

Upgrade Issue

AVRO upgraded for processing data source log files.

Fixed Bugs

- Fixed issue where the same decoder name was used for AVRO sources (decoders must be unique).



Note: Error message will now display if you do not have a unique decoder for every log path.

Known Issues

- If you have an older version of the workstation and choose to reprocess a profile, you will get a false warning—a "Decoders must be unique" error will display although you actually have unique decoders. You can exit and relaunch the workstation and the error will no longer appear, or upgrade to this version of the workstation (client) as a permanent fix.

Data Workbench 6.51 Update

Data Workbench 6.51 release notes include new features, upgrade requirements, fixed bugs, and known issues.

To view previous features and fixes for past releases, see the [release notes archive](#).

[New Features](#) on page 12

[Upgrade Requirements](#) on page 13

[System Updates](#) on page 14

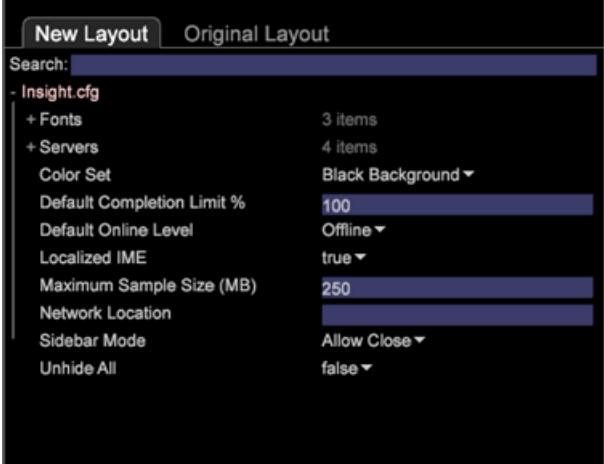
[Fixed Bugs](#) on page 14

[Known Issues](#) on page 14

New Features

Data Workbench 6.51 includes these new features:

DWB 6.51 Features	Description
Avro Data Feed rollout	Report suite hit data will be delivered in a new Apache Avro data source format providing updated features and new variable types for Adobe Analytics Premium (including additional evars, custom events, and solution variables). See Avro Data Feed .
Bar Graph improves display with color gradients overlay	To improve contrast of elements in the Graph visualization, a gradient of colors can be applied to highlight individual bars using the Cylinder effect or left and right background gradients. <ul style="list-style-type: none">• No Gradient Overlay. Select to show bars without gradient overlay applied.• Background from Left. Select to show gradation of colors across all bars from left to right.• Background from Right. Select to show gradation of colors across all bars from right to left.• Cylinder. Select to show gradation of colors from the center of each bar to the edge of each bar.
Updates to the Insight.cfg edit tree	The New Layout view is now the default to edit the Insight.cfg configuration file .

DWB 6.51 Features	Description
	 <p>Also,</p> <ul style="list-style-type: none"> • An Undo (Ctrl+Z) feature is provided for all but the required top-level values in the configuration tree. • When removing elements you are now provided a dialog to confirm actions to avoid erroneous changes.
Double-click to close workstation	Quickly close the workstation (client application) by double-clicking the upper left corner of the title bar.
XML Decoder Enhancements	You now have the ability to use <code>#value</code> tag in XML Paths to pull the value of an XML element.

Upgrade Requirements

Workstation (client) upgrade requirements:

To upgrade the client workstation for version 6.51 from version 6.50, you must run the Workstation Setup Wizard to install the workstation executable and supporting files (using the **Upgrade or repair** mode)—or copy or edit files manually on your server. For this upgrade only, the automated upgrade feature in the workstation from the *Software and Docs* profile is turned off (even though the `Update software = true` is set in the `Insight.cfg` file on your server).

Server upgrade requirements:

Update to the new **Adobe SC meta.cfg** file:

- The `server\Profiles\Adobe SC\Context\meta.cfg` file was renamed to `server\Profiles\Adobe SC\Context\Adobe SC meta.cfg`.
- Updated `Base\Context\meta.cfg` file.

Replace your server build with the updated files.

System Updates

These features have been renamed, deleted, or the installation files or folders were restructured in this release:

- **SSL Connectivity Change.** Certificate generation to upgrade key length was modified to 2048 bits and uses SHA256. This eliminates RC4 ciphers by using modern encryption algorithms.
- Files changed to Read-only.
 - base\context\serverdetails\detailed status.vw
 - base\menu\admin\detailed status for master.vw
 - base\menu\admin\detailed status for query.vw
 - base\workspaces\admin\dataset and profile\detailed status.vw
- **Bookmarks alphabetized.** Right-click in the Bookmark panel and select **Alphabetize**. The bookmarks will align based on alphabetic, case-insensitive characters for each language.

Fixed Bugs

The following are prominent fixes made in Data Workbench 6.51 (since the release of Data Workbench 6.5).

- Previously, **Algorithmic Attribution** coefficients displayed as all zeros when using SGD to converge. This is now fixed.

Known Issues

- Exporting an Association Table from an *Association Chord* visualization that contains *at least one metric* will result in duplicated elements in the rows/columns of the Association Table. To avoid duplicated elements, create a new Association Table and add the desired elements rather than exporting the elements from an Association Chord.
- If you build an Avro decoder, the default field button won't work. Please work with your account manager to get the appropriate configuration for your implementation.
- Upgrading the workstation (client) from DWB 6.50 must be done manually using the Client Setup Wizard in Upgrade mode). And automatic upgrade will not occur (see Upgrade Requirements for the client).
- Using the Chinese and Japanese install packages for Adobe SC AVRO as a stand-alone might give an error stating `Undefined Metric: "$Default Metric$"`. No mapping is currently defined for `$Default Metric$`. However, if you are using the Adobe SC AVRO package along with other packages, such as the Attribution-Premium package, then this error will not occur.

Data Workbench 6.5 Release Notes

Data Workbench 6.5 release notes include new features, upgrade requirements, fixed bugs, and known issues.

To view previous features and fixes for past releases, see the [release notes archive](#).

[New Features](#) on page 14

[Upgrade Requirements](#) on page 16

[System Updates](#) on page 16

[Fixed Bugs](#) on page 17

[Known Issues](#) on page 17

New Features

Data Workbench 6.5 includes these new features:

DWB 6.5 Features	Description
New Association Table and Association Chord visualizations	<p>The Association Table and Association Chord visualizations let you associate metrics with metrics, dimensions, and dimension elements using the Cramer's V algorithm. The Associations Table and Chord compares values using the Cramer's V calculation rather than using Pearson's correlation coefficient as employed in the Correlation Matrix and Correlation Chord visualizations (these can only compare metrics, while the Association Table and Chord can compare metrics, dimensions, and elements).</p> <p>To open, select Visualization > Predictive Analytics > Association Table or Association Chord from the workspace.</p>
New Regression Tree option for Decision Trees	<p>Evaluate a Decision Tree using the Regression Tree option by right-clicking and selecting Options > Regression Tree within a Decision Tree visualization.</p> <p>Updated Decision Tree builder: The new algorithm was introduced for building a Decision Tree. It handles more general data and provides a more informative visualization to improve the precision of the prediction.</p> <p>Improved data sampling module: An updated adaptive sampling scheme helps Decision Tree and Propensity Score achieve higher precision results.</p>
Report Server support for Predictive Analysis and Best Fit Attribution visualizations	Best Fit Attribution, Propensity Score, Correlation Matrix, Correlation Chord, Association Table, and Association Chord visualizations are now multi-pass visualizations, allowing them to run in Report Server. However, only one multi-pass visualization can run in a workspace. (See Known Issues below.)
Context-aware visualizations	When you drag and drop within a workspace, any visualization open in the workspace that can accept what is being dragged (metrics, dimension elements, etc.) visually identifies itself, allowing you to see the visualizations capable of accepting the dragged item. The visualization's border is highlighted in blue showing that the metric being dragged to the visualization is capable of accepting the dragged metric.
New Save Filter button	When you select an item within a workspace—a detail table, visualization, filter, or other item—a Save Filter button opens on the toolbar allowing you to create a filter for the item selected, allowing you to enter a name for the filter before saving it to the profile.
New Layout for editing configuration files	The New Layout interface lets you quickly open and edit .cfg, .dim, .metric and other files in the Configuration tree (the config file editor) by right-clicking and selecting from drop-down menus in a workspace.
New Ctrl > O and Ctrl > E quick keys to open and edit files	<p>New Keyboard Shortcuts</p> <ul style="list-style-type: none"> Type Ctrl > O in a workspace to open a visualization (.vw) file. It is the same as selecting Add > Open from the toolbar. Type Ctrl > E in a workspace to edit a configuration file by default, or browse to open other file types.
New Add > Edit menu option	Right-click and select Edit in a workspace (or from the toolbar) to open .cfg files by default, or Select All to browse to other .dim, .metric and .vw files.
Open recent or pinned files	<p>The Open files feature now lists the most current files and files you want "pinned" to the menu for quick access. Right-click and select Open from menu in a workspace, or select Add > Open and Add > Edit from the toolbar.</p> <p> Attention: Click the pin icon next to the item to save the workspace in the menu.</p>

DWB 6.5 Features	Description
Improved Filter descriptions	New easy-to-understand descriptions of the applied filters will now display in these visualizations and editors: <ul style="list-style-type: none"> • Filter Editor (when you hover over the title) • Filter panel on the Finder • Propensity Score visualization • Cluster Builder visualization • Decision Tree Builder visualization
International Formats for Metrics	Format configuration files were added to support international delimiters in metrics.
S/FTP Delivery	An SFTP/FTP export was added to the Export Integration framework. You can now configure settings from the client (workstation) and export using CSV, TSV, Segment Export, or Segment Export with Header using FTP and SFTP protocols, allowing you to send files to servers. Select any of the exports by right-clicking on a Detail Table . Once selected, an export configuration table opens allowing you to add or update export settings.
Using <i>Windows Certificates Store</i> in the workstation (client)	You can now store the SSL certificate and private key for the Data Workbench Client (Workstation) in the Windows Certificate store for SSL communication with servers. Using the <i>Windows Certificate Store</i> may be preferable for those who manage certificates in a single store. (See Known Issues .)
String Encryption as a service	A new encryption feature encrypts strings and stores them with keys in the <i>Window's Credentials Store</i> .
Windows in workspace resize proportionally	Windows in a workspace now resize proportionally when you resize from a corner by holding down the <Shift> key.
Correlation Matrix recalculates automatically	Correlation Matrix no longer requires a Recalculate option in the right-click menu.
Updates to User Administration of Group Members	Changes to <code>Access Control.cfg</code> and <code>User List.cfg</code> added to Base package for use with User Admin of Group Members .
Updated Query Memory Limit	Updated Query Memory Limit setting in <code>DPU.cfg</code> from <i>unsigned int</i> to <i>_int64</i> is now available for those who wish to use more than 4 GB for this value.

Upgrade Requirements

Follow these requirements and recommendations when upgrading to Data Workbench 6.5.

- Changes in the **Components for Processing Servers\Communications.cfg** file require you to update this file for the DWB 6.5 release. The *SourceListServer*, *SegmentExportServer*, and *NormalizeServer* entries were removed. (DPU's should not be running *sourcelist*, *segment export*, or *normalize servers*.)

System Updates

These features have been renamed, deleted, or the installation files or folders were restructured in this release:

- Correlation Chord, Correlation Matrix, Association Chord, Association Matrix, Propensity Score, and Best Fit Attribution visualizations are now **multi-pass visualizations**, allowing them to run in Report Server.



Important: You can only have one multi-pass visualization in a workspace.

When there are more than one multi-pass visualizations in a workspace, Report Server will fail to generate reports by default and throw the following error:

```
Too many Multipass visualizations in workspace ..... (has #, 1 allowed).
```

Avoid this error by updating your `ReportServer.cfg` file or add this line to your existing file in the *Reporting* section.

```
Max Multipass Per Slice = int: 1
```

The `Max Multipass Per Slice` value must remain set at 1 (See *Known Issues*).

- Changes made to files in the `Server\Menu\Admin` folder.
 - Removed Lookups.
 - Added Archive Files and User Files.
 - Modified Files, Export Files, Log Files.

Updated files include:

```
Base\Menu\Admin\Order.txt (changes to menu)
```

```
Base\Menu\Admin\Files.vw (removed Lookups and Events string, added Archive and Users strings)
```

```
Base\Menu\Admin\Archive Files.vw
```

```
Base\Menu\Admin\User Files.vw
```

- Changes made to files in the `Base\Context\ServerDetails` folder:

```
Base\Context\ServerDetails\Archive Files.vw
```

```
Base\Context\ServerDetails\Export Files.vw
```

```
Base\Context\ServerDetails\Log Files.vw
```

```
Base\Context\ServerDetails\User Files.vw
```

```
Base\Context\serverdetails\files.vw
```

```
Base\workspaces\admin\dataset and profile\server files.vw
```

- Changes made to the *self-administration of member access*.

A new file in the `Base` and `Server` packages includes a `User List.cfg` for Admins to update for member access.

Fixed Bugs

The following are prominent fixes made in Data Workbench 6.5 (since the release of Data Workbench 6.4).

- Fixed block decoding bug that was introduced in release 6.4. Issue was encountered during processing of a very large visitor.
- Fixed alignment bug introduced with compiler upgrade. Issue affected Replication in release 6.4

Known Issues

- Use of the *Windows Certificate Store* is limited to the English only (no Chinese or Japanese language support).

-
- *Regression Analysis* on graphs does not recognize new workspace selections. A possible workaround is to redo the regression after making a new selection.
 - The *Max Multipass Per Slice* value must remain set at 1. You cannot have more than one of the following visualizations in a single workspace intended for Report Server: Propensity Score, Best Fit Attribution, Association Matrix, Association Chord, Correlation Matrix, or Correlation Chord.

[Additional Data Workbench Documentation online](#)

Data Workbench 6.4 Release Notes

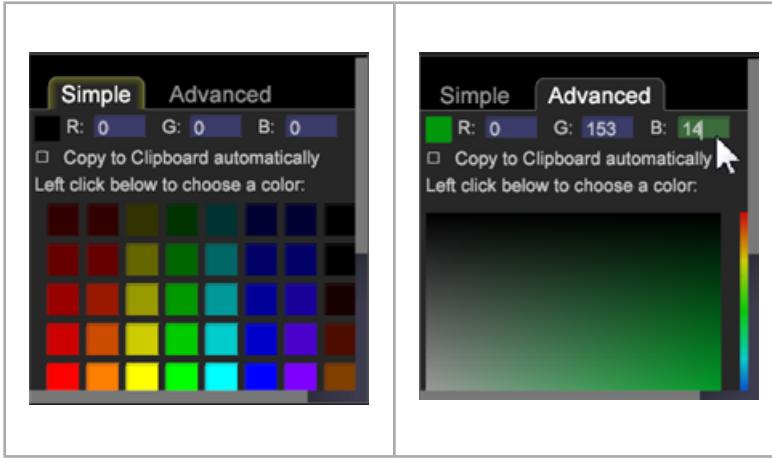
Data Workbench 6.4 release notes include new features, upgrade requirements, fixed bugs, and known issues.

To view previous features and fixes for past releases, see the [release notes archive](#).

New Features

Data Workbench 6.4 includes these new features:

DWB 6.4 Features	Description
Exporting to Analytics Core Services on page 23	The Customer Record Service (CRS) export feature lets you export Data Workbench data from the Details Table to the Adobe Analytics Core Services to integrate with other Analytics' capabilities, including <i>Reports & Analytics</i> .
Workstation Setup Wizard on page 25	Set up the Data Workbench (Client) using a step-by-step installation wizard. Download, set options, install the workbench, and sync up with servers.
List item. on page 27	After installation of the workstation, the Configure Connections to the Server workspace will open with additional information about entering server connection information in the Insight.cfg file. You can also view the connection status to your servers and select profiles.
Presentation Layer on page 28	Annotate and clarify visualizations using a presentation overlay. Add text call-outs, arrows, images, and color coding to highlight and clarify your data, and then share with others.
Metric Dim Wizard on page 30	Employ a step-by-step wizard to facilitate building metric dimensions.
User Administration of Group Member Access on page 33	Administrators can give workstation users the partial ability to manage access control for custom groups.
Locking Profiles in the Workstation on page 37	You can prevent profiles from being overwritten by users working in the Profile, Dimensions, Reports, Workspaces, Metrics, or Filters managers. In the Profile Manager , save the Internal.cfg file to a custom profile to prevent multiple profile files in your system from being overwritten on the server.
New User Interface Features on page 40	Data Workbench 6.4 adds new workspace icons, tool tips, splash screens, and an <F1> shortcut to the help. It also lets you open your log files by selecting the Help > Open Trace Directory from the toolbar.
Updated Clustering algorithm	Expectation Maximization added to Clustering feature. This is an <i>Adobe Analytics Premium</i> feature.
Updated Logging information	Data workbench now uses an expanded logging framework "L4" which provides the ability to configure logging based on the need. The default implementation that comes with the 6.4 package provides vital information on the software processing. Logging can be expanded with additional information to troubleshoot server events and help analyze underlying issues, including additional information for associated server, client and report server. For additional support in implementing additional L4 logging, please contact your account manager.

DWB 6.4 Features	Description
New cfg file for ExportIntegration.exe logging options	<p>A new httpLoggingEI.cfg configuration file (located at server\Admin\Export\httpLoggingEI.cfg) lets you stop INFO logging to the HTTP.log file during Export Integration exports. (The CRS, TNT, and MMP exports already capture verbose logging in individual export log files.)</p> <p>A <i>true</i> setting starts INFO logging (for testing and detailed reporting) to the HTTP.log file, and a <i>false</i> setting stops verbose logging. For a <i>false</i> setting, only a WARNING/ERROR level messages will be sent to the HTTP.log file.</p>
Zoom feature for Graph visualizations	<p>Use the zoom feature to better view metric labels when values reach a higher disparity. Previously the label would disappear with the change in the contrast of values—for example, when you set a higher metric regression value against previous values. You can now zoom in to the visualization by clicking <Ctrl> and moving the mouse wheel while hovering over the graph.</p>
New Color Picker tool	<p>A new color picker lets you select colors from a simple palette of colors, or select the Advanced tab to pick from a gradation scale or enter your own RGB values.</p> 
Finders now export more meta information	<p>More meta information is provided when exporting dimensions and metrics from the Finder.</p> <ul style="list-style-type: none"> Metrics Finder output now includes <i>Name</i> and <i>Formula</i>. Dimensions Finder output includes <i>Name</i>, <i>Type</i>, and <i>Parent</i>.
Insight.exe and InsightSetup.exe are now digitally signed.	<p>These executables are now digitally signed to ensure that the software downloads have not been altered or corrupted.</p>
Date format options	<p>You can change the date format based on your locale in the Standard Time Dimensions.cfg file. Change the default MM/DD/YYYY format to the DD/MM/YYYY format or choose other options.</p>
Files visualization broken out	<p>The Files visualization (Admin > Files) for Base profiles will not include larger directories (removed Logs, Exports, and Lookups) when reporting. This will increase the speed in displaying the report.</p> <p>The larger directories now have their own individual reports (Admin / Export Files, Lookup Files and Log Files).</p>

DWB 6.4 Features	Description
Device Atlas with In-Memory Cache on page 41	The DeviceAtlas.bundle file now uses in-memory cache to greatly improve the performance of look-ups.
Updated Chord visualization	Improved visibility when hovering over a section when viewing the Chord visualization.
Drag dimensions from Finder to a Detail Table	From the workstation, you can now drag dimensions from Finder panel directly to the Detail Table in a workspace.

Upgrade Requirements

Follow these requirements and recommendations when upgrading to Data Workbench 6.4.



Important: It is recommended that you use the newly installed default configuration files and customize them, rather than moving files from a previous installation—with these exceptions:

- Add **Excluded Processes** for *MS System Center Endpoint Protection in Windows 2012 Servers* for the following executables:
 - **InsightServer64.exe**
 - **ReportServer.exe**
 - **ExportIntegration.exe**
 This will allow "white list" rights for these interfacing executables.
- Update the **Trust_ca_cert.pem** certificate on the servers.
- Reorganization of Attribution Profiles.
 - The *Attribution* folder was renamed to **Attribution - Premium** (found in the default installation at *Profiles\Attribution - Premium*).
 - The *Premium* profile was removed and the workspace moved to the new **Attribution - Premium** folder.
- Update **Attribution-Premium settings**. If you have customized profiles with parameter settings that override the default *Adobe SC* profile, then you need to update the custom fields in these configuration files:
 - **Decoding Instructions.cfg**
 - **SC Fields.cfg**
- Because of this reorganization, you will want to remove the old *Attribution* and *Premium* folders from your server installation.

Change these settings

```
Profile = profileInfo:
  Active = bool: true
  Directories = vector: 6 items
    0 = string: Base\\
    1 = string: Geography\\
    2 = string: Predictive Analytics\\
    3 = string: Adobe SC\\
    4 = string: Attribution\\
    5 = string: Premium\\
```

to these settings:

```
Profile = profileInfo:
  Active = bool: true
  Directories = vector: 5 items
    0 = string: Base\\
    1 = string: Geography\\
```

```
2 = string: Predictive Analytics\\
3 = string: Adobe SC\\
4 = string: Attribution - Premium\\
```

- **Update custom Meta.cfg files** (if necessary).

The **Meta.cfg** files in **Base\Context** and **AdobeSC\Context** folders have been updated in this release.

If you override the **meta.cfg** file during installation, then your profile copy needs to be updated with these parameters and the **metadata vector** appropriately entered:

```
94 = meta:
    path = string: SegmentExport:CRS Configuration/CRS Attributes
    acceptable children = vector: 1 items
    0 = Template:
        name = string: CRS Attributes
        value = CRSAttributeConfiguration:
            Attribute Name = string:
            Attribute Type(int,string) = string:
            Field Name = string:

95 = meta:
    path = string: SegmentExportQuery:CRS Configuration/Report Suite
    acceptable children = vector: 1 items
    0 = Template
        name = string: Add Report Suite
        value = string:
```

- **Set Report Server permissions** to generate Microsoft Excel reports On Windows 2012 servers.

1. Set permission of the root folder (**E:\ReportServer**) to *Everyone = full control*.
2. Create the following folders with appropriate permissions:

```
C:\Windows\SysWOW64\config\systemprofile\AppData\Local\Microsoft\Windows\INetCache
C:\Windows\System32\config\systemprofile\AppData\Local\Microsoft\Windows\INetCache
C:\Windows\System32\config\systemprofile\Desktop
C:\Windows\SysWOW64\config\systemprofile\Desktop
```



Note: If you are running Report Server on Windows Server 2012, you need to have Windows Server 2012 R2 installed.

3. Assign "SYSTEM" as the owner for these folders.

- **Add fonts to the Report Server.**

In the **ReportServer.cfg** file, add these fonts (for all languages):

```
Fonts = vector: 3 items
  0 = string: Arial
  1 = string: SimSun
  2 = string: MS Mincho
```

- **Update your version of Microsoft Excel** (if necessary).

With the release of Data Workbench 6.4, support for Excel 2007 has been discontinued. Also because Data Workbench only runs on Microsoft Windows for 64-bit architecture, it is recommended that you also install a 64-bit version of Microsoft Excel.

- **64-bit architecture** required for Workstation (Client) installation.

- **Run the Workstation Setup Wizard.**

Install the new version of the workstation (client) by downloading and launching ***InsightSetup.exe*** and stepping through the setup instructions. The setup wizard will install your files to a new location by default:

Program files are now saved by default to:

```
C:\Program Files\Adobe\Adobe Analytics\Data Workbench
```

Data Files (profiles, certificates, trace logs, and user files) are now saved by default to:

```
C:\Users\<username>\AppData\Local\Adobe\Adobe Analytics\Data Workbench\
```

- **Add fonts to the Workstation.**

In the **Insight.cfg** file, add these fonts (for all languages):

```
FONTS = vector: 3 items
0 = string: Arial
1 = string: SimSun
2 = string: MS Mincho
```

System Updates

These features have been renamed, deleted, or the installation files or folders were restructured in this release:

- The **Base.zip** folder is no longer included in the version update package.
- The **DeviceAtlas.bundle** file now uses an in-memory cache to improve the performance of lookups.
- In the **Log Processing.cfg** file, the *Chunk Size* parameter under *Log Sources* was removed.
- In the **Disk Files.cfg** file, the *Detect Disk Corruption* parameter was removed in these locations:

```
\server\components\disk files.cfg
```

```
\server\components for processing servers\disk files.cfg
```

- New service descriptions for *Adobe Analytics Premium Services* and for *Adobe Analytics Premium Report Services* in the executable properties.
- The *Master Marketing Profile Export* feature in the Details Table was renamed to **Profiles & Audiences Export**.
- The *Test and Target Export* feature in the Details Table was renamed to **Adobe Target Export**.

Fixed Bugs

The following fixes were made in Data Workbench 6.4 (since the release of Data Workbench 6.31).

- Propensity score wasn't resetting when rerunning different inputs in the same workspace. This now resets properly.
- No countable dimensions available when first opening the Correlation Matrix has been fixed.
- Export of Target segments were failing because the mboxPC field was missing. This is now fixed.
- ID request formatted correctly. Using the mbox3rdpartyId identification instead of default PCIDs caused Adobe Target to reject requests generated via the Target/Data Workbench integration (using the **ExportIntegration.exe**). This ID request is now being formatted correctly and throughput is successful.
- Report Server memory leak when exporting to Excel has been fixed.

Known Issues

The following are known restrictions in Data Workbench 6.4.

- **ExportIntegration.exe** requires command-line arguments in English. The output file name should be named in English for *Adobe Target Export*, *Profiles and Audiences Export*, and *Customer Record Service Export*.
- In the *Profiles and Audience Export*, entering unauthorized characters ([CR] or [TAB]) as column names generates incorrect logs resulting in data not exporting correctly.

-
- In Chinese and Japanese versions, Unicode character encoding issue might be encountered in the Path Browser.

Exporting to Analytics Core Services

The Customer Record Service (CRS) export feature lets you export Data Workbench data to the Adobe Analytics Core Services to integrate with other Analytics' capabilities, including *Reports & Analytics*.



Note: In order for the CRS export feature to work, a visitor's Analytics Visitor ID (AVID) must be based on the Marketing Cloud Visitor ID (MCID). Although MCID may be populated in Data Workbench for a visitor, if the client is in the grace period or the visitor's cookie has not been replaced with MCID, the CRS export will not work for that visitor. For more information, see [Identifying Visitors](#) and [ID Service Grace Period](#).

From a **Detail Table** (right-click **Tools > Detail Table** in a workspace), you can set attribute values and the variables required to integrate with Analytics' Reports & Analytics (using Adobe Pipeline Services).

1. Right-click the table header and click **New Customer Record Service**.

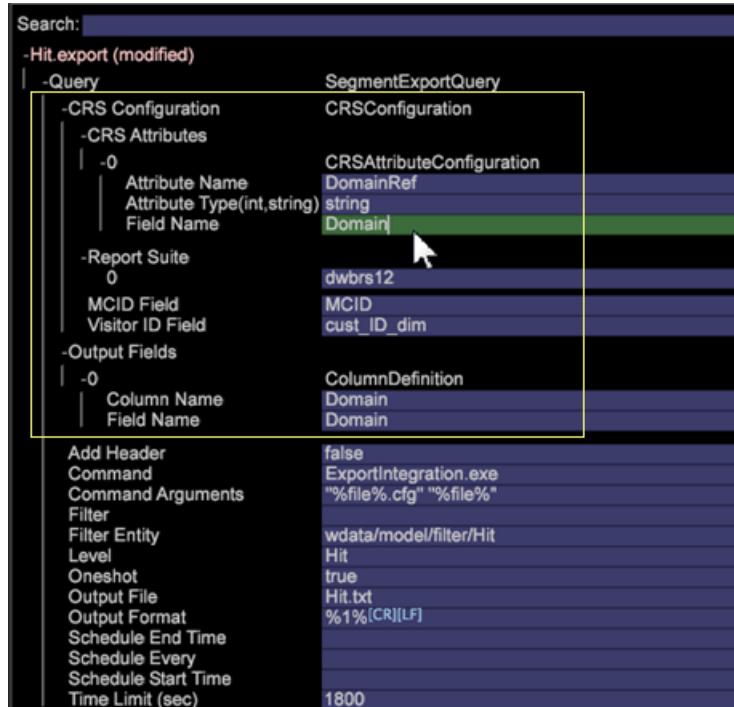


2. Name the export file and save.

The export file edit window will open.

3. Open Query > CRS Configuration.
4. Right-click CRS Attributes > Add New.
5. Enter *CRS Attributes* parameters.

Open the new entry and enter or verify values in the *CRS Attributes* section of the export file:



Attribute Name	Name of the <i>Customer Attributes</i> variable displayed in <i>Reports & Analytics</i> .
Attribute Type	<p>This parameter accepts values of <i>(int, string)</i>.</p> <p> Note: If an attribute is not subscribed to in Analytics:</p> <ul style="list-style-type: none"> The attribute will be created with any valid attribute type supported by Analytics (for this release it is limited to only <i>string</i> and <i>int</i>). If an invalid attribute type is entered, then you will receive an error stating a failure to subscribe to Analytics. <p>If an attribute is already subscribed to in Analytics:</p> <ul style="list-style-type: none"> Make sure to enter the right attribute type for the already subscribed to attribute. If you enter the wrong type for the attribute, then its behavior will be dependent on Analytics' handling of attribute types.
Field Name	<p>Name of the dimension or metric from which the attribute values are selected.</p> <p> Note: The <i>Field Name</i> under <i>CRS Attributes</i> should be the same as the <i>Output Fields > Field Name</i> (which is populated automatically based on the attribute selected). If the <i>Field Name</i> is invalid then the export will not run.</p>

6. Right-click Report Suite > Add New.

7. Enter the *Report Suite ID*.

Open the new entry and enter or verify values in the *Report Suite* section of the export file:

Report Suite	ID of the report suite in <i>Reports & Analytics</i> identifying the <i>Customer Attribute</i> variables being exported.
	<p> Note: Although <i>Reports & Analytics</i> lets you add to multiple report suites, Data Workbench 6.4 will only export a single report suite identified at <i>index 0</i>.</p> <p>The report suite value entered in this field is the report suite ID (and not the name of the report suite).</p>

8. Enter MCID Field parameter.

MCID field	Name of the dimension in your profile that represents the <i>Adobe Marketing Cloud ID</i> . This is a mandatory field and any invalid dimension value entered will not export.
-------------------	--

9. (optional) Enter Visitor ID Field parameter.

Visitor ID Field	If the user wishes to send any other custom ID for a visitor in his/her data, this is where they enter the name of the dimension which represents the custom visitor id. This is an optional field and can be left empty.
-------------------------	---

Workstation Setup Wizard

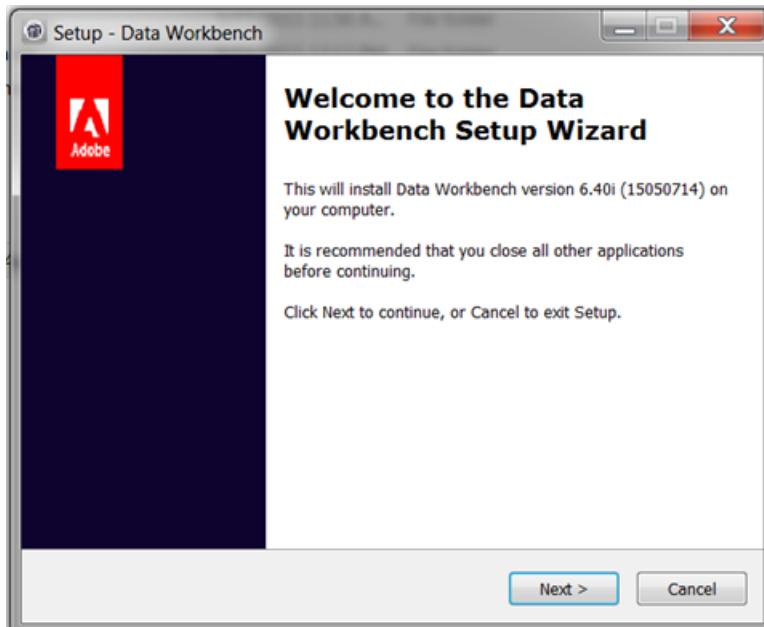
Data Workbench provides a set up wizard to install the workstation (client) application.

Installing the Workstation using the Setup Wizard

Launch the installation wizard executable and walk through each step to install the workstation client program. After installation of the workstation, you can connect to servers and profiles.

1. Double-click the workstation installer executable.
2. Click **Yes** to allow the program to install on Windows.
3. Select a **Language** for the setup wizard.

The wizard will open:



4. Click **Next** on the **Welcome to Data Workbench Setup Wizard** dialog.
5. Select to install a **New Installation** or to **Upgrade or repair** an existing installation.

New Installation overwrites any previously installed files.

Upgrade updates your Workstation to the latest version or lets you repair an existing installation. Data Workbench will compare installed **Insight.exe** files and run the Workstation Setup Wizard if a newer version of the client is available.

6. Select install location:

Typical installs to a default folder and location.

- Program files are saved by default to:

C:\Program Files\Adobe\Adobe Analytics\Data Workbench

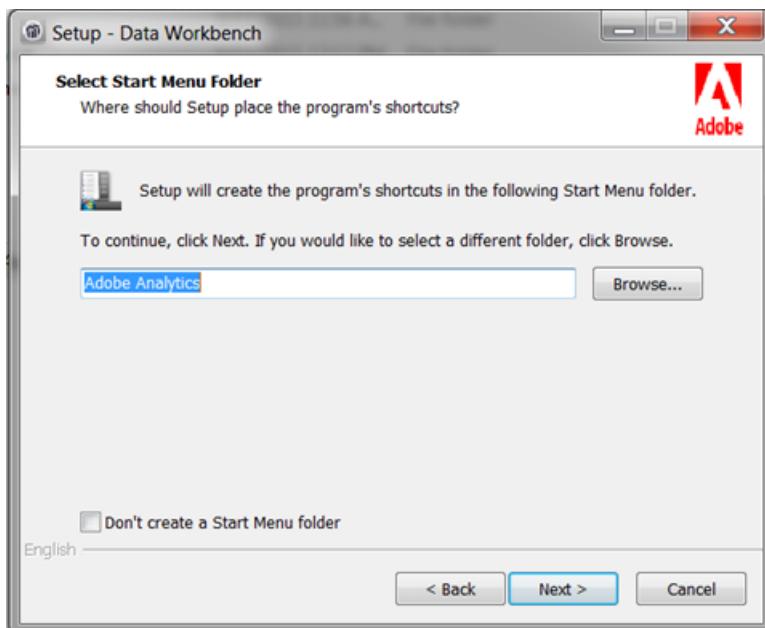
- Data Files (profiles, certificates, trace logs, and user files) are saved by default to:

C:\Users\<username>\AppData\Local\Adobe\Adobe Analytics\Data Workbench\



Important: A generic *Insight.cfg* file without server details will be installed initially. It is recommended that you use the newly installed *Insight.cfg* file and customize it rather than moving a file from a previous installation. Because the path for installing the workstation has changed, the addition of fonts, removal of the *User Folder*, and the removal of the *TraceFileComponent* is recommended.

7. (optional) Select **Custom** to choose the language package and the location of the program and data files.
8. Select location for **shortcuts in the Start Menu**.



Click **Don't create a Start Menu folder** to not install a shortcut on the Windows Start Menu.

9. Click **Next**. A summary of selected file location paths and languages will display. Click **Install**.
10. Locate the **Data Workbench Certificate**.

If the setup wizard cannot find the Data Workbench certificate during installation, it will open a dialog to browse to the location of the certificate (a **.pem** file located by default in the client **Certificates** folder), or click **Skip** to find the certificate after installation.

Click **Install** after locating the certificate.

11. After the setup wizard is complete and Data Workbench installed, click **Finish** to complete setup.

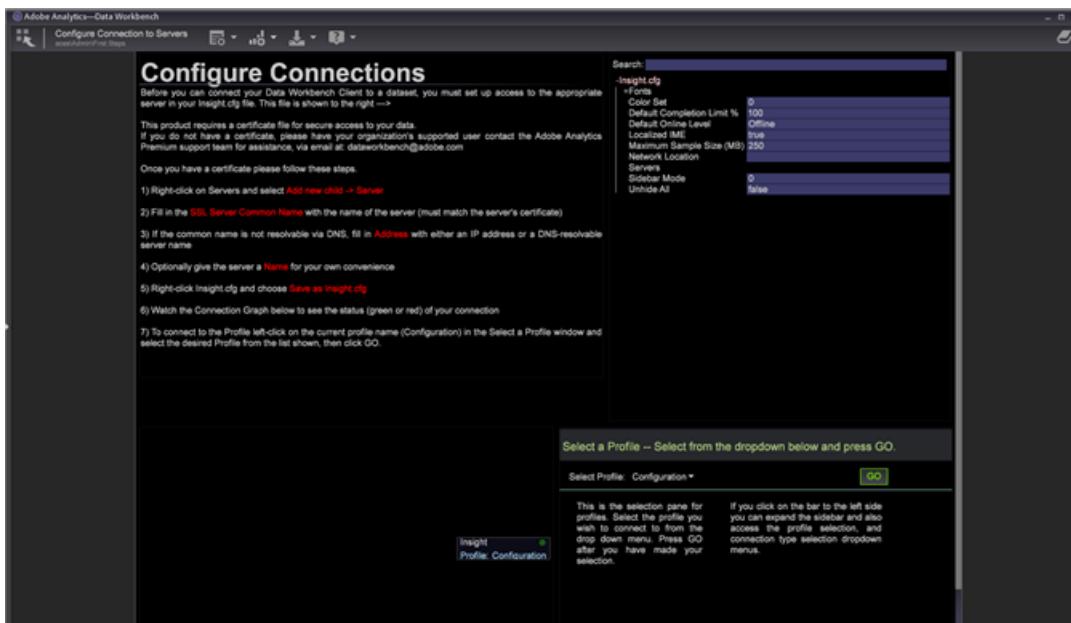


Note: The default log location for the Workstation Set up Wizard at
C:\Users\<userName>\AppData\Local\Temp.

Select the **Launch application** checkbox to open the workbench after setup.

12. Configure connections to servers in **Insight.cfg** file.

After installation of the workstation, the Enhanced Workstation Configuration Experience workspace will open with additional information about *entering server connection information* in the *Insight.cfg* file and an option to select a profile from the drop-down. You can also view the connection status to your servers.



Installation Folders

The Data Workbench folder structure has two installation locations:

- **Program Files**

The **Insight.exe** and supporting client files (**Insight.ini**) are now located by default at
C:\Program Files\Adobe\Analytics\DataWorkbench

- The **Appdata** folder.

Insight.cfg, profiles, certificates, trace logs, and user files are now located by default at
C:\Users\<Winuser>\AppData\Adobe\Analytics\DataWorkbench\

You can set the path for the **Appdata** folder in the **Insight.ini** file:

```
[InitialSettings]
AppDataFolder=C:\Users\mhiatt\AppData\Local\Adobe\Adobe Analytics\Data
Workbench\
Locale=en-us
```

Uninstalling the Workstation

Data Workbench now includes an executable to uninstall the workstation (located by default at **Program Files\Adobe\Adobe Analytics\Data Workbench\unins000.exe**).

Launch and follow steps to remove the Data Workbench Workstation files from your hard drive.



Note: You can launch the **unins000.exe** executable from the folder, using the **Uninstall Data Workbench** shortcut from Start Menu, or from **Control Panel > Program and Features**.

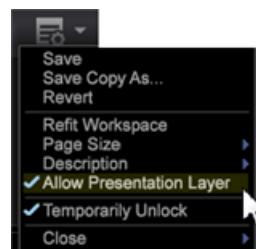
Presentation Layer

The Presentation Layer lets you mark up and annotate your workspace visualizations and then publish with your call-outs and comments. Add text descriptions, graphic objects, callout arrows, color coding, images, and other features in an overlay to add annotations and clarify important data points, and then share with stakeholders.

Add Annotations to your Visualizations:

1. Open a workspace.
2. Click **Allow Presentation Layer**.

Toggle the command in this menu to allow or disallow working in the presentation layer. A checkmark will appear when the presentation layer is allowed.



When you are in the presentation layer, a clickable icon will appear in the right side of the toolbar.



This clickable icon shows that you are in the **Hide Presentation Layer** mode but not able to see annotations.

3. Click icon to toggle to **Show Presentation Layer**.

View annotations in the presentation layer.



This clickable icon shows that you are in the **Show Presentation Layer** mode but not able to edit.



Note: You can also right-click in the workspace to change modes using menu commands.



4. **Edit and add callout features.**

Click icon again to capture a static representation of the workspace visualization and open an overlay to add or edit annotations.

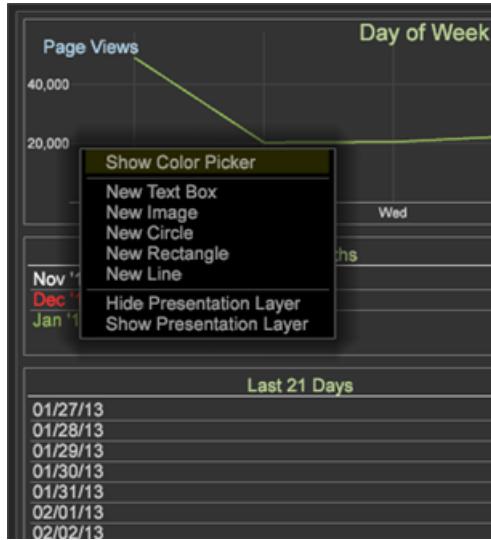


This clickable icon shows that you are in the **Edit Presentation Layer** mode to add and edit annotations.

A gray overlay appears while in edit mode.

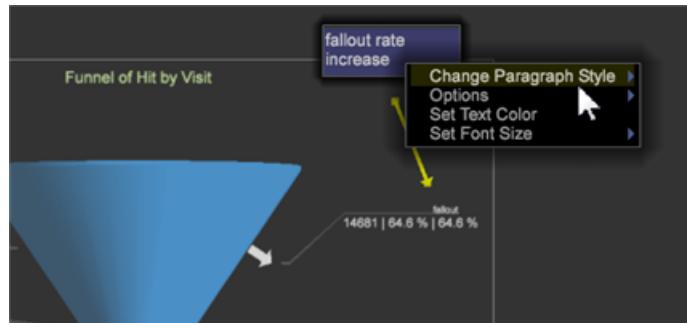
5. **Add callout features.**

Right-click the visualization to add a new text box, arrows, images, and other callout features to highlight and annotate for presentation.



6. Set graphic features.

Select a graphic object and right-click to set options. You can set colors for a box or other graphic object, adjust text settings in a text box, and add arrows or set layering for lines.

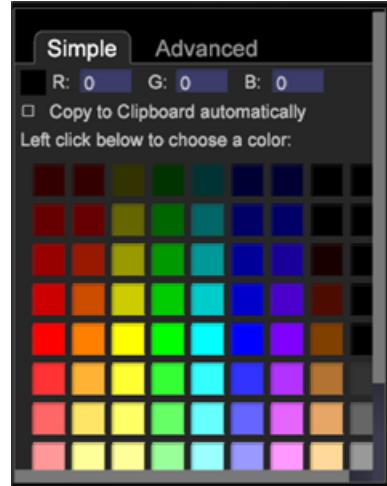


7. Add images to the workspace.

Right-click to add your .png and .jpg images to the overlay and resize.

8. Assign colors to graphic objects.

You can select objects in the presentation layer and assign colors using the color picker.



9. Export and share presentation layer.

After annotating your workspace visualizations, you can export the presentation layer with visualizations and share as a .png graphic file.

Click **Export > Export PNG**.

Metric Dim Wizard

Create Dimensions defined by metric attributes (Metric Dims) using a step-by-step wizard. Then test, preview, and save the new Metric Dim to your Dimensions list.

A Metric Dim converts a metric into a new dimension. For example, a Metric Dim based on a metric of Page Views and level of Visitor will display dimension elements based on the total Page Views for each Visitor. It lets you extend a currently defined metric based on dimension elements to create and save as a new dimension.

Step 1: Select Dimension and Metric

1. Open the Metric Dim Wizard.

In a workspace, right-click and select **Tools > Create Metric Dim**.

2. Name the Metric Dim.

As a default, the Name field will auto-populate based on Level and Metric selections.

3. Select a Dimension Level.

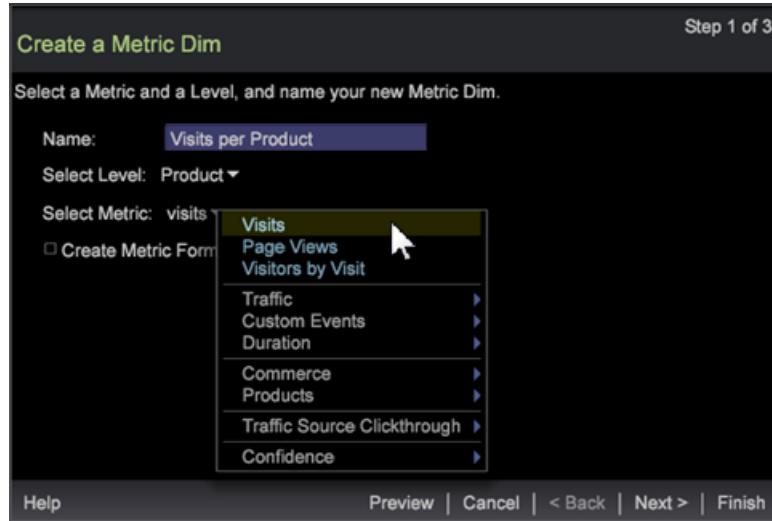
The dimension level is the parent dimension containing all constituent element values to filter input and define a dimension type.

Dimension levels include:

- Clickthrough
- Hit
- Product
- Visit
- Visitor

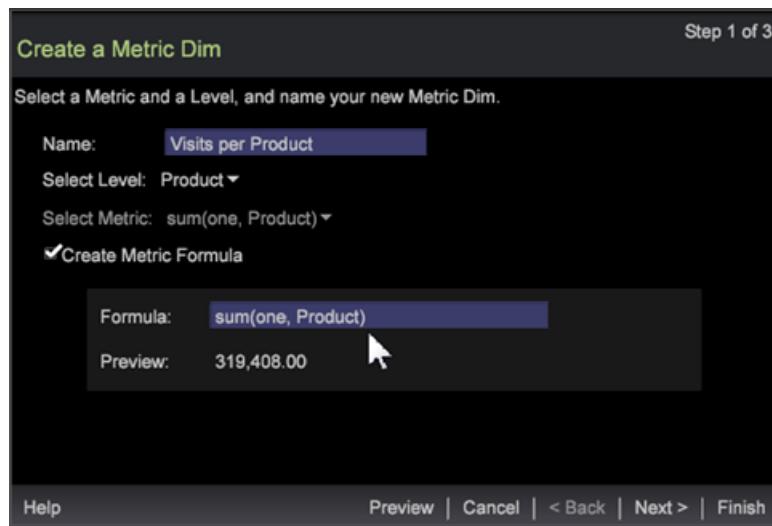
4. Select a Metric.

Select a pre-built metric to extend and save as a metric dim.



5. (optional) Create a Metric Formula.

Click the box to enter a custom metric formula. The calculated Preview value will appear validating the expression.



You can add your own *metric expression* or cut and paste from another metric editor or visualization. Syntax errors, formula errors, undefined filters, and other errors are reported in the wizard.

6. Click **Next**.

Step 2: Format and Set Buckets

1. Select a **Format** for the new metric dim.



The format defines how the metric will be presented when opened in a visualization. These formats are selected [printf standards](#), defined below:

```
%[flags][width][.precision][length][specifier]
%0.2lf = % _ [flags] 0 [width] .2 [.precision] 1 [length] f[ specifier]
```

In the **Preview** field, a value will appear based on the metric and format selected.

2. Add **Bucket Count** expression.

You can define a metric dim with various ranges, or buckets. This returns subsets of elements based on size, such as [0-4], [5-10],...). Elements of the Dimension Level relate to the elements whose range contains the value of metric. See the bucket expression description at [Syntax for Dimension Expressions](#).

3. Click **Preview** to open table of Metric Dim values before saving.



The table details metric values per metric dim.

4. Click **Show in Dimension Menu** to add the newly created dimension to the **Dimension** tab in the **Finder**.
5. Click **Next**.

Step 3: Finish and Save

1. Select to launch the Metric Dim Editor, graph visualization, or table after saving.

Field	Description
Launch Metric Dim Editor	Open the Metric Dim Editor.
Launch Graph	Launch a PNG graphic of the table.
Launch Table	Launch a table in the workspace with values in columns listing values of the new metric dim compared with values of the selected metric.

2. Click **Finish** and save.

A save dialog will open allowing you to save the file. The selected options to view values will open in the workspace.

User Administration of Group Member Access

Administrators can give workstation users the partial ability to manage access control for custom groups.

Self-administration of group member access gives rights to non-administrators to add and delete members in a custom group. The administrator creates a **User List** file and sets up group access in the **Access Control.cfg** file for the new group members.

Accessing the Servers Manager

Setting up the **User List** file and synching it with the **Communications.cfg** file is done in the **Servers Manager** workspace.

1. On the worktop, click the **Admin** tab > **Dataset and Profile** tab.
2. Open the **Servers Manager** workspace.
3. Right-click *>your server name>* in the diagram and select **Files**.

The server files will open in a table with columns *File*, *<server name>*, and *Temp*.

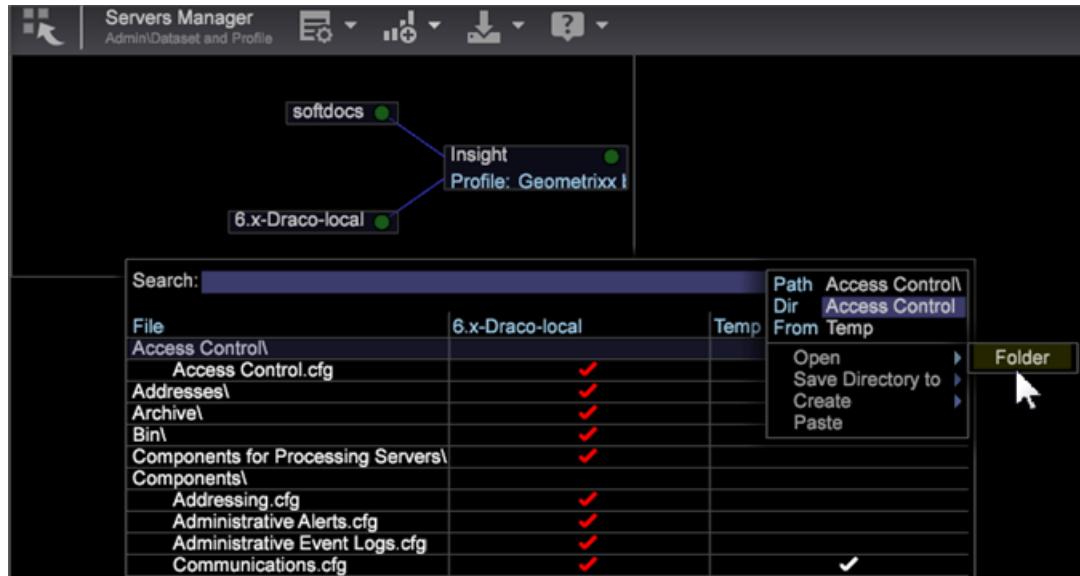
4. **Make Local** by right-clicking in the server column of a server file (for this feature **Access Control** and **Components/Communications.cfg**).

A white checkmark will appear in the **Temp** column. You can edit in the Temp folder. Then right-click the checkmark and **Save To** the server. (It turns red when synched with server).

Create a User List.cfg file

The administrator needs to create a **User List.cfg** file in the **Access Control** folder.

1. Right-click **Access Control** row in the **Temp** column and select **Open > Folder**.



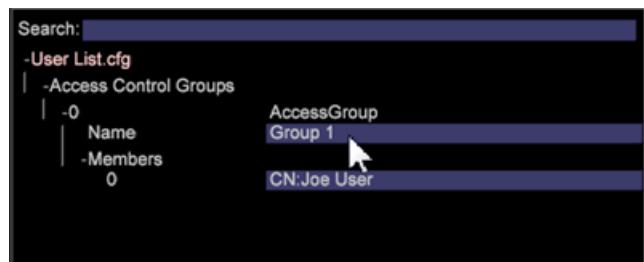
The Access Control folder in the **Temp** folder will open listing a single **Access Control.cfg** file.

2. Add another text file to this folder and name it **User List.cfg** (next to the **Access Control.cfg**).
3. Add the following parameters to the **User List.cfg** file.

The User List file should contain a vector of **AccessGroup** objects, and each **AccessGroup** object should have a name and a vector of strings called **Members**.

```
Access Control Groups = vector: 1 items
0 = AccessGroup:
  Name = string: Group 1
  Members = vector: 1 items
    0 = string: CN:Joe User
```

You can then edit and add users this in the Workstation view of the **User List.cfg** file.



Here's the most basic parameters to add to the **User List.cfg** file. The Members can then be added in the Workstation view.

```
Access Control Groups = vector: 1 items
0 = AccessGroup:
  Name = string:
  Members = vector: 0 items
```

★ Important: As with any **.cfg** file that you manually edit, make sure to use spaces instead of tabs and to pay close attention to the whitespace and syntax. A mistake in this file will cause *Adobe Insight Server* to ignore the User List file.

The **Name** field in each **Access Group** will be referenced within the **Access Control.cfg** file.

 **Note:** Only valid members with directory service prefixes, such as **CN:** or **OU:** are accepted, and these cannot contain wildcard character (*).

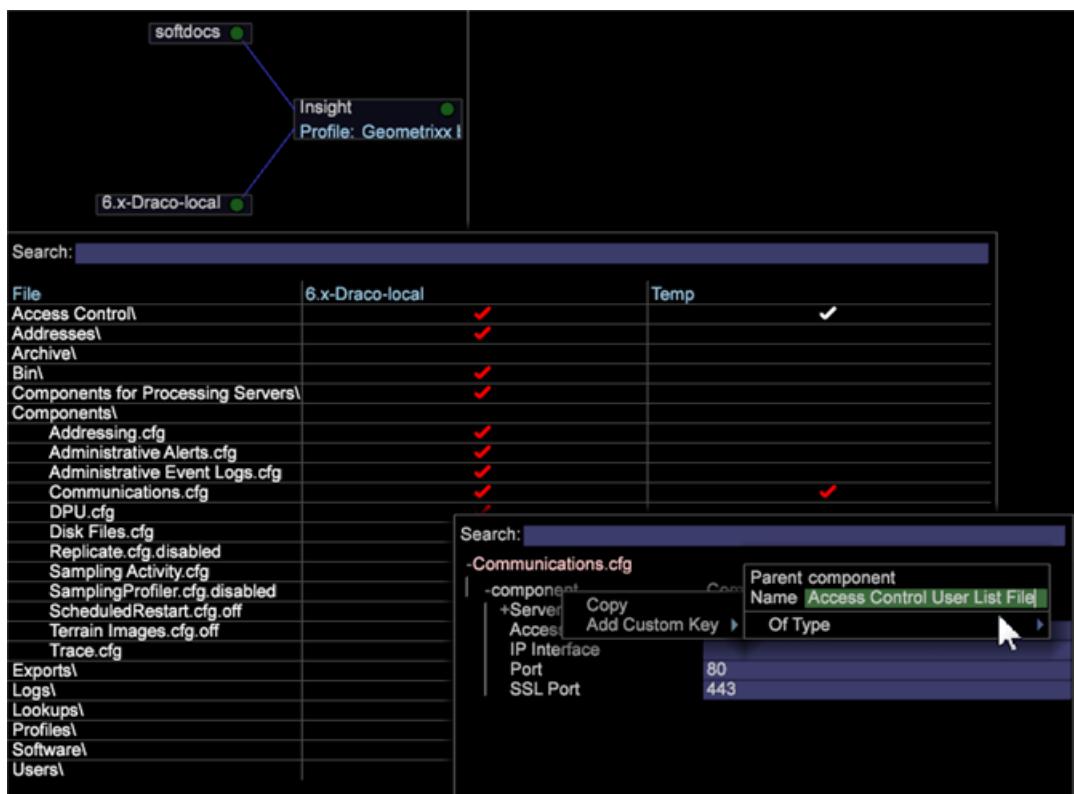
Set up the **Communications.cfg** file

An administrator first enables this feature by opening the **Components > Communications.cfg** file and adding a new key with the name **Access Control User List File**. The string value of this key is the path where this new file will be located.

1. From the server files, click **Components** and right-click the checkmark in the server column. Click **Make Local**.

A white checkmark will appear in the **Temp** column.

2. Right-click the checkmark in the **Temp** column and select **Open > in Workstation**.
3. In the **Communication.cfg** file, right-click **component** and select **Add Custom Key**.



4. Type the **Name** as *Access Control User List File* and set **Of Type** as *String*.

 **Note:** You cannot create the new list file as a Path. To remedy this, you need to save the file, open it in an editor (Notepad), and change "String" to "Path":

Before:

```
component = CommServer:  
    Access Control File = Path: Access Control\\Access Control.cfg  
    Access Control User List File = string: Access Control\\User List.cfg
```

After:

```
component = CommServer:  
    Access Control File = Path: Access Control\\Access Control.cfg  
    Access Control User List File = Path: Access Control\\User List.cfg
```

5. Save the **Communications.cfg** file and (if necessary) save it to the server. This will restart components in the server to make sure you haven't made any mistakes that could prevent the **Communications.cfg** file from being parsed.
6. If your system includes processing servers, modify the configuration file in the **Components for Processing Servers.cfg** file.
7. Right-click **Communications.cfg** and save to server.

The Data Workbench administrator can now confirm that the intended user(s) have access to the user list file and allow the users to manage the group. The user(s) will be able to open the User List file, edit it, and add and remove CN or OU members as needed.

Synch up the **Access Control.cfg** file

The administrator can then edit the **Access Control.cfg** and insert references to the group(s) defined by the *User List* file.

The references to the group(s) should be inserted just like any other member, but with the following syntax:

```
$ (Group Name)
```

Where "Group Name" matches what's defined in the user list file, including white spaces.

```

Search: [redacted]

-Access Control.cfg (modified)
| -Access Control Groups
|   -0
|     Name          AccessGroup
|     Administrators
|   -Members
|     0             OU:Insight Developers
|   +Read-Write Access
|     Read-Only Access
|   +1
|   +2
|   -3
|     Name          AccessGroup
|     Power Users
|   -Members
|     0             $(Power User Group)
|   +Read-Only Access
|   +Read-Write Access
|   +4
|     AccessGroup (Cluster Servers)

```

At this point the Data Workbench administrator can confirm that select group users have access to the user list file. The select users can then open the **User List.cfg** file, edit it, and add and remove CN or OU members as needed.

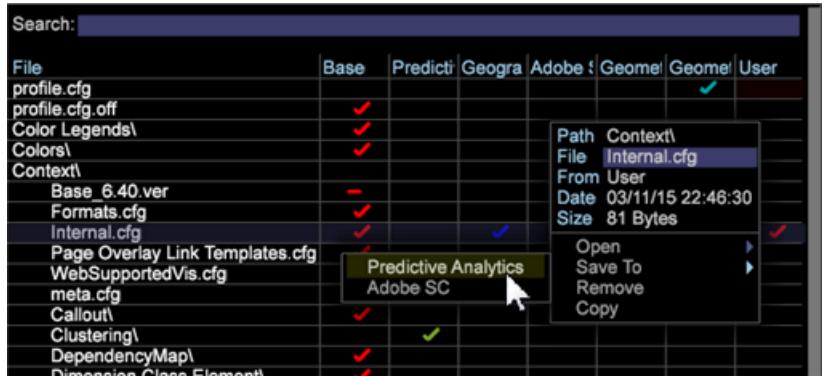
Locking Profiles in the Workstation

The **Internal.cfg** file applied in the Profile Manager prevents changes by users to your custom profiles by the Profile, Dimensions, Reports, Workspaces, Metrics, and Filters managers.

You can prevent profile files from being modified and overwritten when using the managers by saving the **Internal.cfg** file to your custom profile in the Profile Manager. This configuration file prevents users from overwriting multiple files when working in the managers (accessed from the **Admin > Profile** menu).

Locking Profiles in the Profile Manager

1. In the workspace, right-click **Admin > Profile Manager**.
2. In the **Profile Manager**, right-click **Context > Internal.cfg** and **Make Local**.
3. Right-click checkmark in **User** column and save to a *<custom profile>*.



Note: Only changes to profile files by the managers are prevented when saving the **Internal.cfg** to a custom profile in the Profile Manager. You can still save workspaces to the server from the worktop using the **Save to server** command.

Localizing Time Dimensions

Configure the time dimensions to display correctly for the locale.

You can configure the displayed format of time dimensions based on locale in the **Standard Time Dimensions.cfg** file (located by default at **Server/Profiles/<my profile>/Dataset/Transformation/Time/Standard Time Dimensions.cfg**).

For example, in North America you can express the date May 3rd, 2015 as 5/3/15, or %m/%d/%y. However, in other parts of the world this could be interpreted as %d/%m/%y, or March 5th, 2015 due to an ambiguity in the values. To avoid this situation, an administrator might want to change the displayed format to match the expectations of the users in a locale.

1. Override Default Time Dimensions in Standard Time Dimensions.cfg

To enable this feature, the administrator must override the defaults by either editing the existing time dimensions or by creating new time dimensions with additional parameters.

An example of a modified time dimension follows.

The **Format** values for Week, Hour, Day, Month, and Hour of Day are set to the defaults in the example.



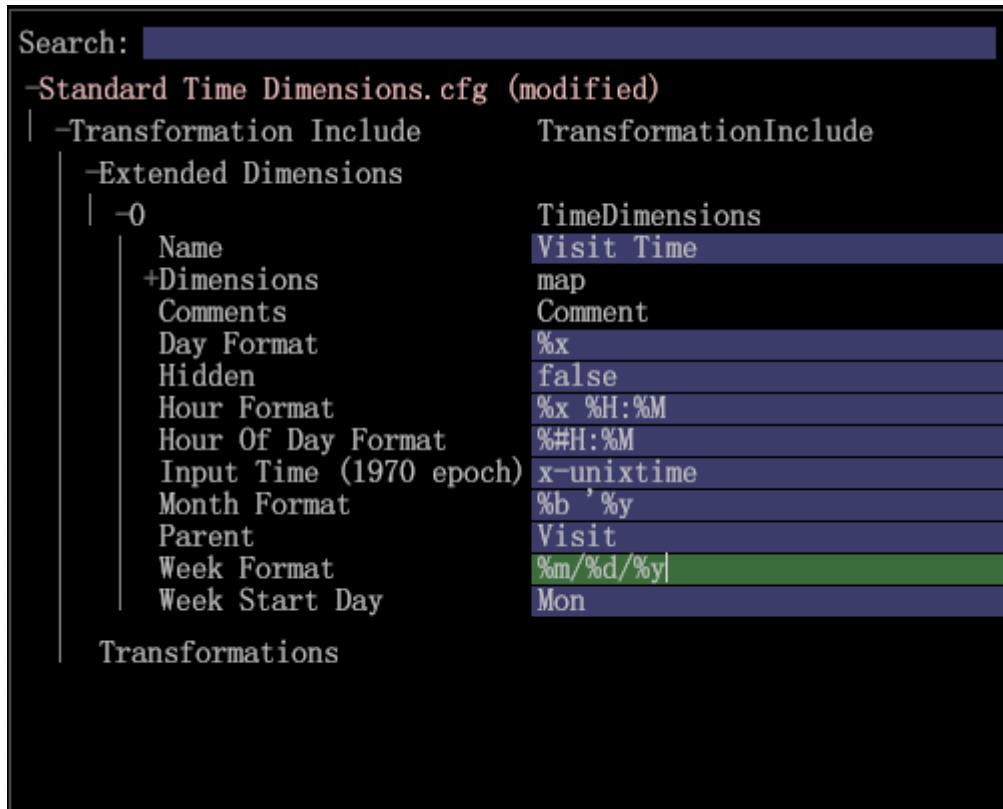
Note: If these lines are omitted, Data Workbench's behavior won't change and the dimension will be compiled using the defaults.

```
Transformation Include = TransformationInclude:  
Extended Dimensions = vector: 1 items  
0 = TimeDimensions:  
Comments = Comment: 0 items  
Dimensions = map:  
    Day = string: Day  
    Day of Week = string: Day of Week  
    Hour = string: Hour  
    Hour of Day = string: Hour of Day  
    Month = string: Month  
    Week = string: Week  
Hidden = bool: false  
Input Time (1970 epoch) = string: x-unixtime  
Week Format = string: %m/%d/%y  
Hour Format = string: %x %H:%M  
Day Format = string: %x
```

```

Month Format = string: %b '%y
Hour Of Day Format = string: %#H:%M
Name = string: Visit Time
Parent = string: Visit
Week Start Day = string: Mon
Transformations = vector: 0 items

```



2. Configure the `meta.cfg` file

Additionally, it's necessary for the package administrator to add these parameters and their defaults to profile's `meta.cfg` file. This allows editing from the workstation.

Here's an excerpt from a configured `meta.cfg` file.

```

dimensions = vector: 6 items
0 = Template:
...
...
5 = Template:
  name = string: Time Dimensions
  value = TimeDimensions:
    Name = string:
    Comments = Comment: 0 items
    Hidden = bool: false
    Week Format = string: %d/%m/%y
    Hour Format = string: %x %H:%M
    Day Format = string: %x
    Month Format = string: %b '%y
    Hour Of Day Format = string: %#H:%M

```

```

Input Time (1970 epoch) = string:
Parent = string:
Week Start Day = string: Mon
Dimensions = map:
    Hour of Day = string: Hour of Day
    Day of Week = string: Day of Week
    Hour = string: Hour
    Day = string: Day
    Week = string: Week
    Month = string: Month

```

Here is an example of a **meta.cfg** file in the workstation:

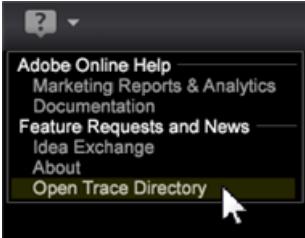
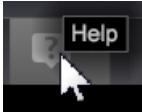


The administrator is then able to go into the **File Manager**, open the file(s) where the time dimensions are configured (e.g., **Standard Time Dimensions.cfg**), and edit them using in the workstation.

New User Interface Features

Data Workbench 6.4 adds new workspace icons, tool tips, splash screens, and <F1> shortcut to the help.

New UI Feature	Description
Open Trace Directory in Help Menu	Access and open the logs files directly by clicking the Help icon on the toolbar .

New UI Feature	Description
	
Type <F1> to open help	<F1> key opens Data Workbench help anywhere in the workstation.
Tool tips over icons	Hover over icons in toolbar to see new tool tips. 
New icons	Updated Adobe Analytics icons for executables, landing pages, splash screen, and other user interface elements. 



Note: You can turn off the icons in the toolbar and return to text names. Add the **Toolbar Icons = bool: false** parameter to the *Insight.cfg* file. False turns off the icons in the workstation user interface and displays menu names in text in the toolbar.

Device Atlas with In-Memory Cache

The **DeviceAtlas.bundle** file now uses an in-memory cache to greatly improve the performance of lookups

By default, Device Atlas will cache up to 100,000 user-agents and their properties. The LRU cache is entirely self-contained inside the **DeviceAtlas.bundle**, so any version of the server capable of using a bundle file will automatically benefit from the increased performance as soon as the new file is loaded.

Modifying the *DeviceAtlas.cfg* file

The maximum size of the LRU cache can be configured by modifying the Cached Elements Limit parameter in the *DeviceAtlas.cfg* file.

Starting with the Data Workbench 6.4 release, you can override the default Cached Elements Limit of "100000" by changing its value and saving the file. The default value has been selected to be sufficient for most needs.

```
component = DeviceAtlasComponent:  
  DeviceAtlas Bundle File = string: Lookups\\DeviceAtlas\\DeviceAtlas.bundle  
  
  Unsyncynchronized Bundle Extraction Path = string: Temp\\DeviceAtlas\\  
  Cached Elements Limit = unsigned int: 100000
```

Data Workbench 6.31 Update

Data Workbench 6.31 provides bug fixes and identifies known issues and their workarounds.

Bug Fixes

- Fixed issue where the **Segment Export** output contains random characters in the export file (which has been identified as random buffer data). This occurs for Segment Exports with both legacy and the new segment export formats.
- Fixed issue where **Add Dimension** in a table was displaying only extended dimensions. Now all standard and extended dimensions are displayed.
- Fixed issue in Simplified Chinese version where the client executable, Insight.exe, if the installation directory was named with non-English characters.
- Fixed issues in the Report Server in the Simplified Chinese version where the email subject and the attachment file names were garbled but the actual output reports were valid.

Known Issues

- Legacy Segment export files output with double quotes even if the export file doesn't contain quotes in the Output Format field.

Workaround: Add these three lines to the .export file. Setting these values will not trigger an MMP integration (as other configuration fields are required) but will bypass unwanted automatic escapes.

```
MMP Configuration = MMPConfiguration:  
    MMP Segment Name = string: UNESCAPE DUMMY  
    MMP Visitor ID Field = string: [Specify a Dimension from the output  
of  
the current export]
```

(The first line has two (2) leading spaces and the next lines four (4). The Dimension from the output of the current export needs to be referenced in the MMP Visitor ID Field.)

Upgrade Instructions

Upgrade instructions for Data Workbench 6.31 are the same as [Data Workbench 6.3 Release Notes](#) on page 43.

New profiles for Data Workbench are located on the **Software and Docs** profile at

```
Profiles - Current\DataWorkBench\  
    English Translated\DataWorkBench_6.31-en-us\
```

File	softdocs1.insight.omniture.com	Temp
Beta Testing\	✓	
Documentation\	✓	
Internal-Only\	✓	
Profiles - Current\		✓
Adobe SC\	✓	
Base 5.3\	✓	
Base 5.4\	✓	
Base 5.5\	✓	
Base 6.0\	✓	
DataWorkBench\	✓	
Chinese Translated\	✓	
English Translated\		
DataWorkBench_6.1-en-us\	✓	
DataWorkBench_6.11-en-us\	✓	
DataWorkBench_6.20-en-us\	✓	
DataWorkBench_6.21-en-us\	✓	
DataWorkBench_6.30-en-us\	✓	
DataWorkBench_6.31-en-us\		
Adobe SC_6.31_en-us.zip	✓	
Attribution-SC_6.31_en-us.zip	✓	
Base_6.31_en-us.zip	✓	
Geography_6.31_en-us.zip	✓	
IP Geo-intelligence_6.31_en-us.zip	✓	
Predictive Analytics_6.31_en-us.zip	✓	
Premium_6.31_en-us.zip	✓	
Site_6.31_en-us.zip	✓	
Japanese Translated\		✓

Data Workbench 6.3 Release Notes

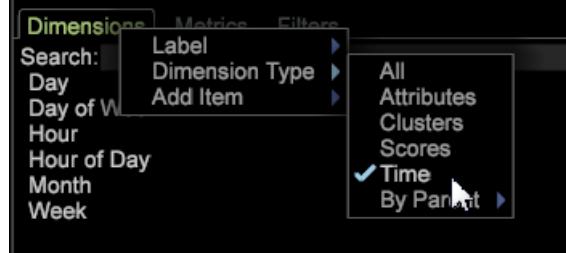
Data Workbench 6.3 release notes include new features, upgrade requirements, fixed bugs, and known issues.

To view previous features and fixes based for each past release, see the [release note archives](#).

New Features

Data Workbench 6.3 includes these new features:

Features	Description
Best Fit Attribution on page 50	Best Fit Attribution provides a machine-learning approach to determine the varying levels of influence that customer interactions contribute to a successful conversion event, and then assigns the appropriate level of attribution to each event. Best Fit Attribution lets you evaluate touches over a window of time before the successful conversion event occurred, and then employs the Data Workbench algorithm to build an attribution model based on your data with results specific to your marketing campaigns and internal workflow.
Profiles and Audiences Export on page 67	Share rich customer segments created in Data Workbench to the Adobe Marketing Cloud. Integrating with the Master Marketing profile enables the Adobe Marketing Cloud and other Adobe Analytic capabilities the opportunity to leverage rich audience segments created in Data Workbench. This requires that you have the Marketing Cloud configured and running. Integrating with the Master Marketing Cloud lets you run predictive features such as clustering or propensity scoring, and then push out segments to the larger Adobe Marketing Cloud or other products such as Adobe Target and Adobe Experience Manager .
Segment Export with Custom Headers on page 70	Added industry-standard format options for segments for direct integration with other capabilities, such as R , without additional file manipulation. This provides a more seamless workflow and quicker analysis.
Clustering 2.0 on page 72	Includes a new KMeans++ algorithm (KMeans is currently supported) that uses a faster approach to finding centers for an expedited cluster-generation process.
Trend Lines on page 73	Present a very visual and easy-to-interpret depiction of the data.
Regression Analysis Graph on page 64	Provides the ability to compare the impact of one factor to another directly within the analyst workflow.
Pie Chart Update on page 74	Updates to the Pie Chart visualization lets you use default colors identified in a legend, or set colors based on the color chart.
Chord Visualization on page 60	The Chord Visualization provides another view of the Correlation Matrix .
Query String Grouping on page 76	If you have many fields with custom evars, props, and variables, during log processing you can build a name value pair to combine fields in a report.
Latency Analysis on page 77	The Latency visualization lets you analyze latent customer behavior within a set number of days before or after an event occurred after a campaign or other event type.
Time dimension	In the Finders panel, you can now right-click on the Dimensions tab and select Dimension Type > Time . A list of time dimensions will display in the search results.

Features	Description
	
Lock feature	The new Lock feature displays an icon in the toolbar when a workspace is locked. You can unlock the workspace by clicking the Add menu and then clicking Temporarily Unlock .
Logical Operators and new Metric features in the Filter Panel	AND/OR logical operators were added to the Filter Panel , allowing you to join or add metrics when filtering data. As you change metrics, the Filter Percentage adjusts and displays accordingly.
Keyboard Shortcuts	New keyboard shortcuts in Data Workbench allow you to navigate across the main worktop and individual workspaces using the arrow keys. In addition, the toolbar in the workspace is now displayed on the worktop window.
Windows 8.1 support	Microsoft Windows 8.1 64-bit is <i>now supported</i> for client installation.

Upgrade Requirements and Recommendations

New profiles for Data Workbench are located on the **Software and Docs** profile at

```
Profiles - Current\DataWorkBench\
    English Translated\DataWorkBench_6.31-en-us\
```

File	softdocs1.insight.omniture.com	Temp
Beta Testing\	✓	
Documentation\	✓	
Internal-Only\	✓	✓
Profiles - Current\		
Adobe SC\	✓	
Base 5.3\	✓	
Base 5.4\	✓	
Base 5.5\	✓	
Base 6.0\	✓	
DataWorkBench\	✓	
Chinese Translated\	✓	
English Translated\	✓	
DataWorkBench_6.1-en-us\	✓	
DataWorkBench_6.11-en-us\	✓	
DataWorkBench_6.20-en-us\	✓	
DataWorkBench_6.21-en-us\	✓	
DataWorkBench_6.30-en-us\	✓	
DataWorkBench_6.31-en-us\	✓	
Adobe SC_6.31_en-us.zip	✓	
Attribution-SC_6.31_en-us.zip	✓	
Base_6.31_en-us.zip	✓	
Geography_6.31_en-us.zip	✓	
IP Geo-intelligence_6.31_en-us.zip	✓	
Predictive Analytics_6.31_en-us.zip	✓	✓
Premium_6.31_en-us.zip	✓	
Site_6.31_en-us.zip	✓	
Japanese Translated\	✓	

Upgrade Server:



Note: If you have customized profiles that take precedence over the default files provided in the Base package, then you will need to update these customized files:

-
- **Update the Meta.cfg file** (`E:\..\Profiles\<your custom profile>\Context\meta.cfg`) to set updated password encryption for the File System Unit (FSU server), and to add entries for the Name Value Pair tranformations to take advantage of [Query String Grouping](#) on page 76.

1. Open the `meta.cfg` file on the FSU.
2. Change the data type for **Proxy Password** from "string" to "EncryptedString" in the *Workstation Configuration* section.

```
Proxy User Name = string:  
Proxy Password = EncryptedString:      (from Proxy Password = String)  
Use Address File = bool: true
```

3. Add new entries to enable the new Name Value Pair transformations: `BuildNameValuePair` and `ExtractNameValuePairs`.

Open a workspace and right-click **Admin > Profile Manager**.

Under **Context**, click the **meta.cfg** file in the **Base** column and then click **Make Local**. From the User table column, right-click and select **Open > in Workstation**.



- In the new window, click **metadata** and add acceptable children templates.

+90	meta
-91	meta
-acceptable children	
-0	Template
-value	Column
Column Name	
Field Name	
name	Name Value
path	BuildNameValuePair:Input Columns
-92	meta
-acceptable children	
-0	Template
-value	Column
Column Name	
Field Name	
name	Name Value
path	ExtractNameValuePair:Output Columns
+93	meta

- Open **transformation** and add new templates.

-transformations	
+0	Template
-1	Template
-value	BuildNameValuePair
Name	
Comments	Comment
Condition	AndCondition
Delimiter	
Input Columns	
Output	
name	BuildNameValuePair
+2	Template
+3	Template
+4	Template
+5	Template
+6	Template
-7	Template
-value	ExtractNameValuePair
Name	
Comments	Comment
Condition	AndCondition
Delimiter	
Input Field	
Output Columns	
name	ExtractNameValuePair
+8	Template

-
- **Update for Fast Merge improvements.** Add parameters or change values to the following configuration files to take advantage of speed improvements in Data Workbench during a transformation.

- **Communications.cfg** (E:\Server\Components\Communications.cfg)

```
18 = SourceListServer:  
    URI = string: /SourceListServer/  
    Listing Interval = int: 10 (new)
```

- **Disk Files.cfg** (at E:\Server\Components and E:\Server\Components for Processing Servers)

```
Disk Cache Size (MB) = double: 1024 (from double: 256)  
Disk Cache Read Limit (MB) = double: 768 (new)
```

- **Log Processing Mode.cfg** (E:\Server\Profiles\<your profile>\Dataset\Log Processing Mode.cfg)

```
(changed)  
Batch Bytes = int: 268435456  
Cloud Bytes = int: 268435456  
Real Time FIFO Bytes = int: 268435456  
  
(new)  
Cache Bytes = int: 32000000  
Fast Input Decision Ratio = double: 200  
Fast Input FIFO Bytes = int: 268435456  
FIFO Hash Mask = int: 16383  
Fast Merge Buffer Bytes = int: 536870912  
Slow Merge Buffer Bytes = int: 268435456  
Fast Merge Fan In = int: 64  
Key Cache Size Logarithm = int: 21  
Max Seek = int: 512  
Output Old Buffer Bytes = int: 536870912  
Overflow FIFO Bytes = int: 67108864  
Paused = bool: false
```



Note:

- To take advantage of the Fast Merge improvements, make sure you have at least 8 GBs of RAM per DPU.
- The Save Interval, the duration between dataset saves, might need to be reduced if the new Log Processing Mode parameters are applied. Because the Transformation will complete faster, the dataset state file will contain more data between Save Intervals compared to prior versions.

It is recommended to reduce the Save Interval setting to 1800 seconds (30 minutes)—the default Save Interval is 3600 seconds (1 hour). (If this line needs to be added to the Log Processing Mode.cfg, note that there are two (2) leading spaces).

```
Save Interval (sec) = int: 1800
```

- It is recommend that the new Log Processing Mode parameters not be applied to a dataset that is configured with a **CrossRows** using the **All** operation. With these settings, the memory usage for this operation can become too great for the environment.

- **Adobe Target with DWB integration update.** A new export file, ExportIntegration.exe, replaces the existing TnTSend.exe file on the Insight Server (E:\Server\Scripts\TnTSend.exe). This new export file supports both **Adobe Target** integration and coordination with the new Master Marketing Profile (MMP) and **Adobe Audience Manager**.

You will need to update the following commands for Adobe Target exports.

```
Command = string: TnTSend.exe
to
Command = string: ExportIntegration.exe
```



Note: This will only affect exports created prior to version 6.3.

You can also try the following to employ the old export process:

- Create a new Test And Target Export in the workstation.
- Modify the old Test and Target export found in Server/Profiles/<your profile>/Export .

- **Update the Adobe SC profile.** Changes to the Exclude Hit.cfg file require a field to be declared in the associated Decoding Instructions.cfg file.



Note: If your Adobe SC profile includes a customized Decoding Instructions.cfg file, you will need to include a DelimitedDecoder parameter to your customized file.

```
0 = DelimitedDecoder:
  Delimiter = string: \t
  Fields = vector: x items
  ...
  5 = string:
Changed to:
  5 = string: x-hit_source
```

Adding the DelimitedDecoder field allows you to take advantage of feature updates and to avoid possible Log Processing problems resulting from these updates.

Upgrade Client:

- **Update your client from the server.**

Once your server has been updated, your client can update automatically if the Insight.cfg file is configured properly:

1. Edit the Insight.cfg file.

```
Update Software = bool: true
```

Then **Save**.

2. Exit and launch the client.
3. Connect to the profile.

The client will automatically upgrade to Data Workbench 6.3.

4. Exit out of the client.
5. Edit Insight.cfg

- Change Proxy Password = string:

```
to Proxy Password = EncryptedString:
```

Remove the value of the previous Proxy Address and Proxy Password.

- **Save**.

6. Launch the client.
7. Edit Insight.cfg.

-
- Enter Proxy Password for all the servers and Save.
 - Enter the Proxy Address for all the servers and Save.



Important: The Proxy Address and Proxy Password must be entered and saved from within the client.

8. Connect to the profile.



Note:

- Follow the exact upgrade sequence in order to avoid an account lockout. If the account is locked, please perform all the required changes in the exact sequence listed, save your work, and exit out of the client. Wait for the lockout to release (about 45 minutes), then launch the client again.
- The password modification should be performed in the client only due to the fact that the passwords are saved in Windows Credential Vault.

• **Recommendation: New Windows Aero Themes.**

Upgrade the look of your client application using Windows Aero Themes.

• **Recommendation: Fonts for Chinese and Japanese versions:**

Chinese:

- Arial
- SimSun

Japanese:

- MS Gothic
- Meiryo
- MS Mincho
- Arial
- SimSun



Note: *SimSun* can be used for Chinese and Japanese. If attempting to write in half-byte characters in Japanese, you also need to include *MS Mincho*. To enable these fonts in *Insight.cfg*, you can add these parameters.

```
0 = string: Arial  
1 = string: SimSun  
2 = string: MS Mincho
```

These fonts should be listed in the workstation configuration file: *Insight.cfg*.

Upgrade to Adobe Analytics Premium

To run **Best Fit Attribution** in Data Workbench, you need to receive new certificates from Adobe ClientCare for your Client, Server, and Report Server (.pem files) to support Adobe Analytics Premium. Each of the new certificates will have this parameter:

```
Product = Premium
```

The Premium Package is available for download on **Software and Docs** under the **Getting Started** tab on the **Profiles and Lookup** files workspace. Navigate to **Profiles - Current\DataWorkBench\<language>\DataWorkBench_6.30-en-us\Premium_6.30_en-us.zip**.

Once the **Premium** profile is loaded on your Server, you will need to add a *Premium* parameter to your custom **Profile.cfg** file. This allows your custom profile to include the menus, visualizations, and workspaces as part of Adobe Analytics Premium.

Fixed Bugs

- Fixed issue where the **Density Map** visualization was missing largest elements.
- Fixed issue in **Density Map** where area of elements was not portraying the proportion of the metric value.
- Fixed issue where dragging metric from **Finders** panel to metric legend outside of the metric column would delete the legend from the workspace.
- Fixed issue where **Print Workspace** using **Sidebar** and **Both** options will not include the Copyright info in the printed page.

Known Issues

- Users of **AMD Radeon™ graphics cards** should update to the latest graphics drivers. Some early versions of the driver claim they support OpenGL 3.2 but are inconsistent.
- Output generated by **Segment Export** configuration without a header declaration can result in a bogus header appearing at the beginning of the file that conflicts with the first set of rows.
- **Add Dimensions** is showing only the *Extended Dimensions*. The workaround is to use the **Finders** tool to drag dimensions to tables.
- When 3D Scatter Plot Visualization includes callouts, the zoom might display plots outside the border of the visualization. To work around this issue, zoom the 3D Scatter Plot first and then add callouts to your visualization.
- Using Workstation in Remote Desktop session will crash when renaming workspaces.
-
- **Legacy Segment export files output with double quotes** even if the export file doesn't contain quotes in the Output Format field.

Workaround: Add these three lines to the .export file. Setting these values will not trigger an MMP integration (as other configuration fields are required) but will bypass unwanted automatic escapes.

```
MMP Configuration = MMPConfiguration:  
    MMP Segment Name = string: UNESCAPE DUMMY  
    MMP Visitor ID Field = string: [Specify a Dimension from the output  
of  
                                the current export]
```

(The first line has two (2) leading spaces and the next lines four (4). The Dimension from the output of the current export needs to be referenced in the MMP Visitor ID Field.)

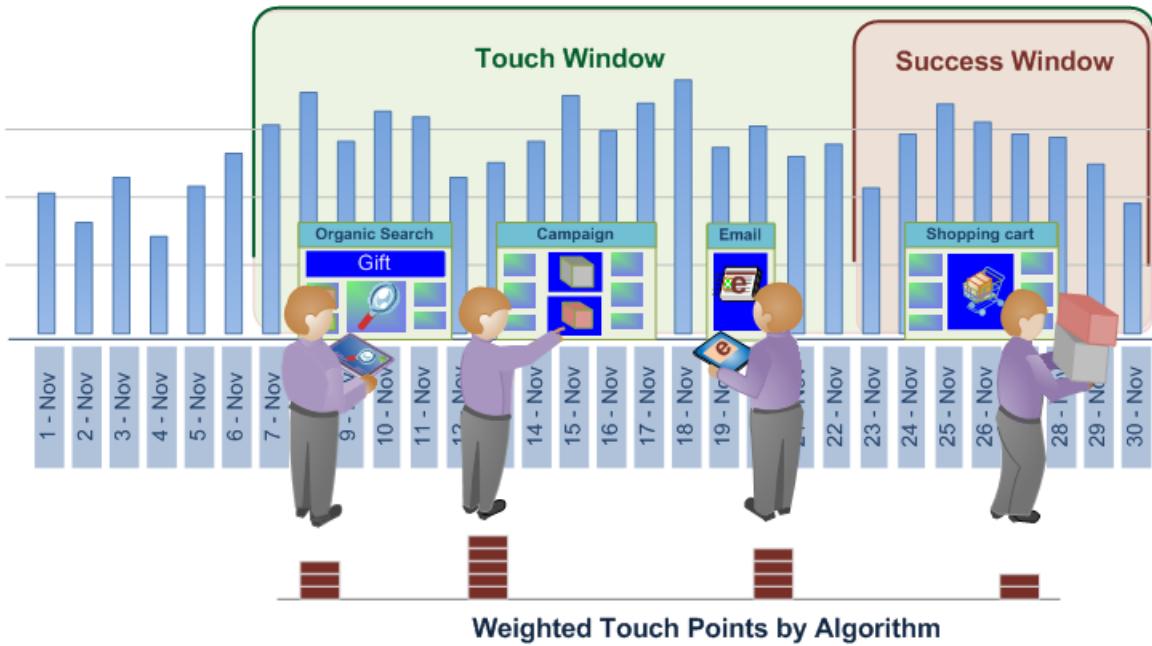
Data Workbench 6.3 features

Data Workbench 6.3 includes the following features.

Best Fit Attribution

Best Fit Attribution is a machine-learning approach to assigning attribution values across the different channels of a successful conversion event. Data Workbench automatically evaluates contributions to success across a window of time per channel, and then builds an attribution model based on your customers' actual interaction patterns.

Best Fit Attribution lets you compare the interactions, or touches, that contributed to a successful sale, email sign-up, or other performance indicators. The attribution analysis automatically assigns weight to the most important touches and provides an attribution model per channel based on your data and responsive to your market and internal protocols.



For example, if a customer visits your site through an organic search, then engages with a campaign, and then signs up for an email, **rules-based Attribution** would identify the first touch or last touch, or evenly distribute success attribution across all touch points using preset attribution models. Where rules-based attribution is defined by the user, the Best Fit attributes sets values through an algorithm by calculating the probability of a conversion as a function of the observed touch points.



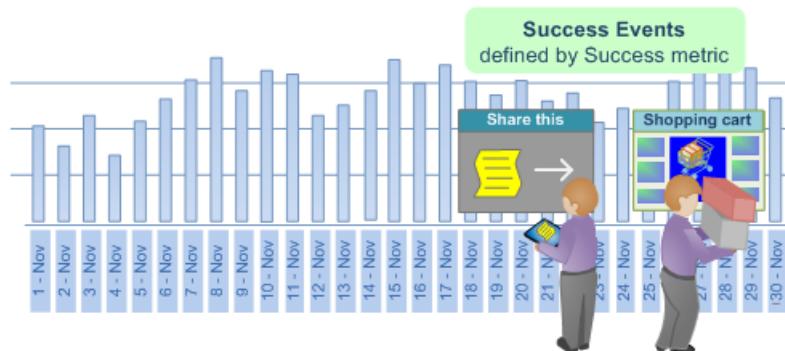
Note: To run **Best Fit Attribution** in Data Workbench, you need to update your server certificate (. pem file) to support Adobe Analytics Premium. You also need to add **Premium** to your custom **Profile.cfg** for the client and receive new certificates from Adobe ClientCare for Server and Report Server. See [Upgrade Requirements and Recommendations](#) on page 44 for Data Workbench 6.3.

Basic Setup

See [Build a Best Fit Attribution Model](#) on page 55 for step-by-step instructions.

Set the Success metric

Define a metric representing a success event.



The Success Metric is often *Orders*, although you can leverage Data Workbench to define a very complicated success metric in conjunction with the Success Window.

Set the Touch metric (optional)

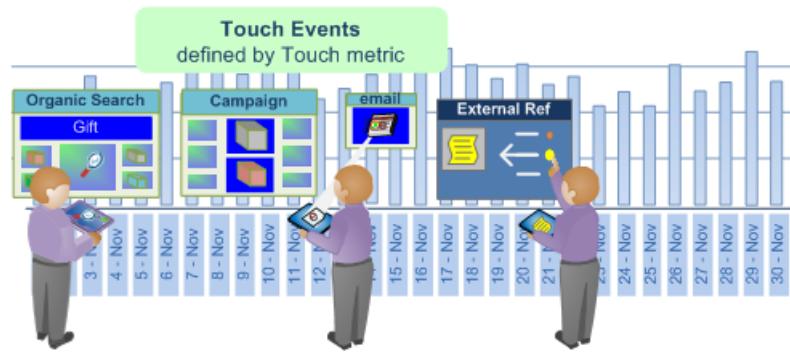
Identify interactions to track that led to a successful conversion, then set the Touch metric over which attribution will be calculated.



Note: Setting a Touch metric is only required if you are using it to derive Channel Metrics from drag and drop Dimension elements instead of using existing Channel metrics.

If you do not have a metric defined for campaigns or channels, but do have dimensions representing channels, the Best Fit Attribution can build them for you automatically based on the Touch metric.

For example, with the Touch Metric set as *Hits*, and given a dimension called *Media Type* with elements that include *Email*, *Press Release*, *Print Ad*, and *Social Media*, the visualization will generate Channel metrics of the form *Hits* where *Media Type* = *Email* when you drag and drop the element(s) onto the visualization.

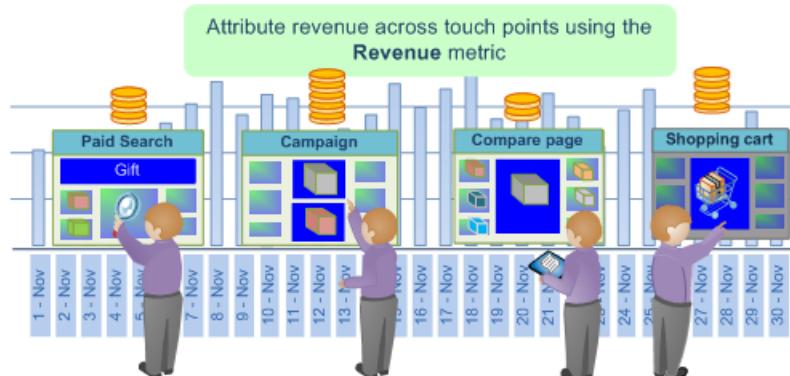


The Touch metric then determines the allocation of attribution scores to identify marketing interactions considered influential for success, allowing you to qualify marketing touches for the population identified in the Success window. You can set metric such as *Page Views* or *Hits*, or use customized touch metrics specific to your needs.

In many cases, the Touch window should include the Success window to evaluate a long lead time in the sales cycle.

Set the Revenue metric.

You can opt to identify revenue across touch points by setting an appropriate revenue metric. If specified, the model will display the distribution of revenue over the input channels.



You can set a revenue metric with currency data types to allocate success across all top touch points defined and analyzed. This metric breaks down the final sales revenue and allocates based on the weighting allocated by the algorithm.

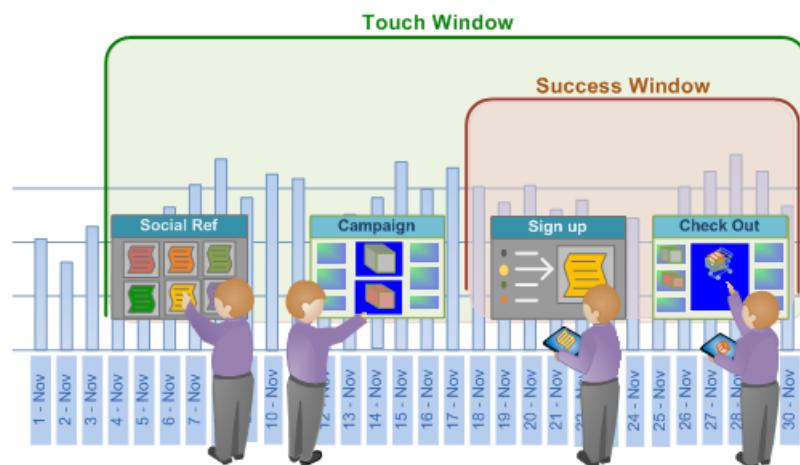
Set the Success and Touch Windows.

The Success window defines the population to examine and the period for successful events, allowing you to indicate the windows of time and breadth of population to consider for the analysis through a workspace selection. The **Success** window defines the period and population to examine for success events. The **Touch** window specifies the historical time period to examine for channel interactions leading up to the success events.



Note: Setting a Touch Metric is only required if you are trying to build Success metrics automatically by dragging dimension elements onto the visualization.

You can set a day, month, year, or any available time frame to constrain your evaluation of success and touch events across the sales cycle or for specific audiences entering your site. Creating windows to limit attribution allows you to focus your analysis on the relevant periods of time for your specific needs.



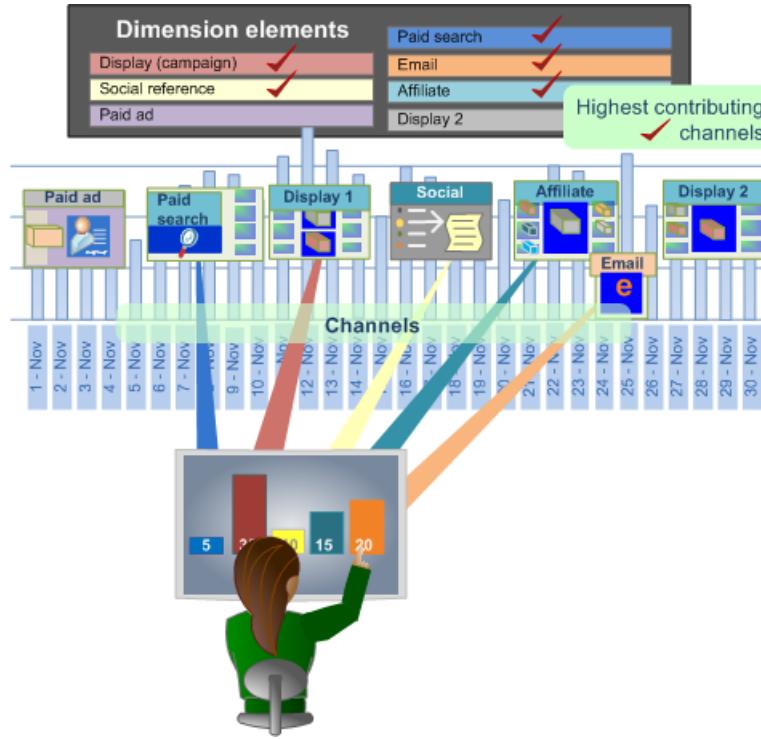
In many cases, you will want the Touch window to include the Success window to let you extend your analysis over a long lead time based on your sales window. Or you can track and analyze touches separate from the success event.

Select the Channels.

When entering channels you have two choices.

Add the Touch Metric and add Dimension Elements to the Channels

In many cases, you will want to break down the top touch points by dimension elements to define specific channels. Based on the element values, Best Fit Attribution will automatically select the top performers and rank them according to percentage and display them in a chart visualization.



An attribution model will be built by drawing on the visitors who interacted during your Success window and examining the channel Touches during the Touch window that did or did not result in a successful event.

Breaking Down by Channels

When entering channels you have two options:

- Add a **Touch Metric** and then add **Dimension Elements** for the Channels.
or
- Create metrics that filter for the channel elements that you want to evaluate.

Option 1: Add a Touch Metric and add Dimension Elements for Channels.

This is the easier approach. Best Fit Attribution creates the metrics automatically to evaluate for attribution. In the example below the Touch Metric is **Hits** and Channels are: **Display Campaigns**, **Email Campaigns**, and **SEM Campaigns**.

Using this method, Best Fit Attribution creates a metric in the background for evaluating the attribution across the channels (but you never see the auto-generated metric and they are not saved). In the example below, three metrics are created where Hits is filtered for each of the three channels (e.g., *Display Campaigns*, *Email Campaigns*, and *SEM Campaigns*). This is the easiest because you let the Best Fit Attribution create the metrics for you.

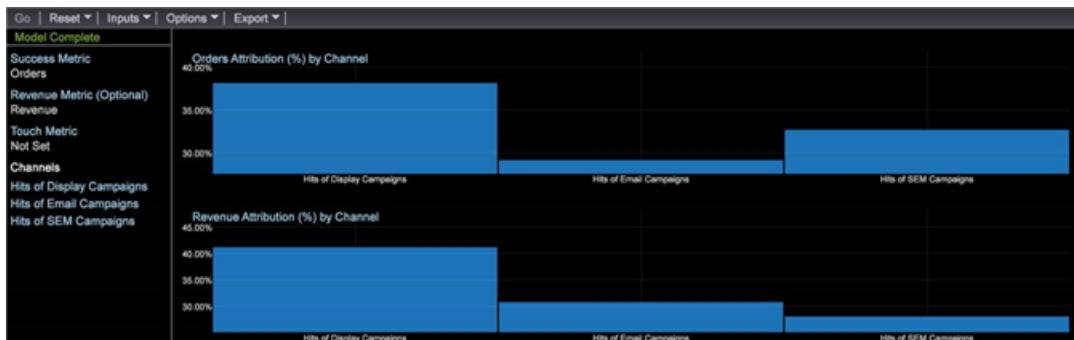


Option 2: Create a Metric.

In the second option, you create and save the metrics for the channels that you want to evaluate by filtering a specific channel. An example of such a metric is shown below.

Editing	Hits of SEM Campaigns.metric
Name	Hits of SEM Campaigns
Formula	sum(One, Hit)[MarketingChannel = "SEM Campaigns"]
Preview	6,614,522

Then, instead of entering a Touch Metric and Dimension Elements for the Channels you can click on the menu bar in the visualization and select **Inputs > Add Channel** and then select the metrics that you created.



See the example of the second method below. You can see that the results of both options are identical.

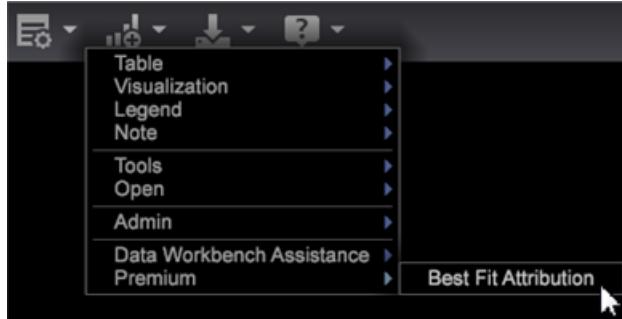
Build a Best Fit Attribution Model

Open Best Fit Attribution from the **Premium** menu and follow these steps to build a Best Fit Attribution model.

See an overview of [Best Fit Attribution](#) on page 50.

1. Open **Best Fit Attribution**.

Open a workspace and click **Premium > Best Fit Attribution**.



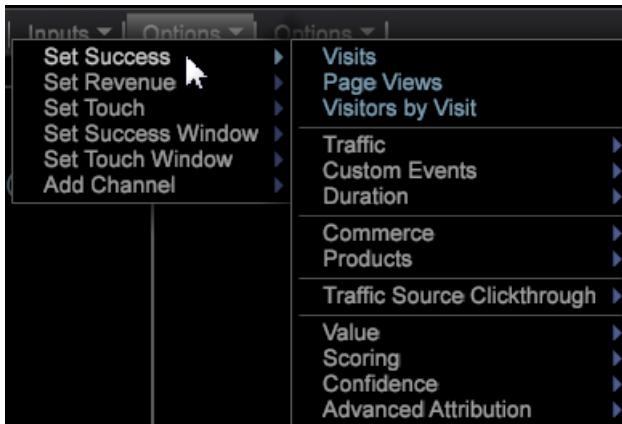
Note: Best Fit Attribution is an Adobe Analytics Premium feature that requires you to enable Premium in your Profile. It requires you to update your certificate and add the Premium profile to your profile.cfg file. See [upgrade instructions](#) for DWB 6.3.

2. Set the **Success** metric.



Note: You can either drag a metric from a **Finder** table to the left pane of the Attribution visualization, or select from the **Inputs** menu.

Click **Inputs > Set Success**. The metric menu will open.



Select a metric that identifies a successful conversion.

3. (optional) Set the **Revenue** metric.

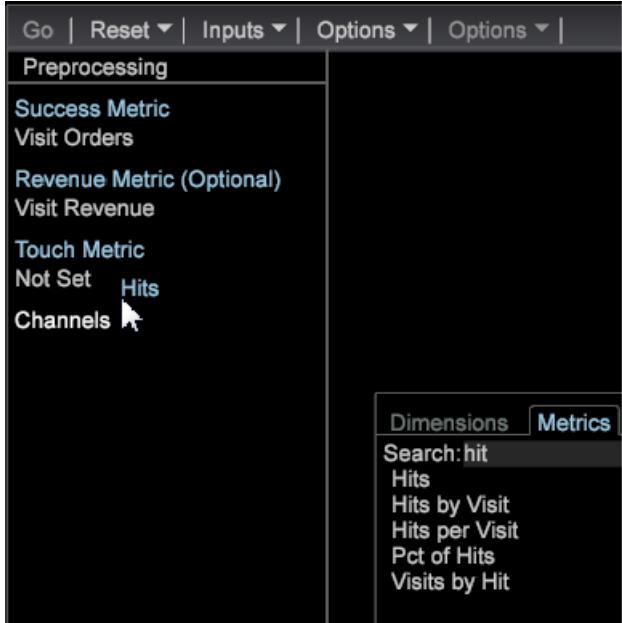
Set a metric to evaluate revenue across the conversion process.

4. Set the **Touch** metric.



Note: Setting a Touch Metric is only required if you are trying to build Success metrics automatically by dragging dimension elements onto the visualization.

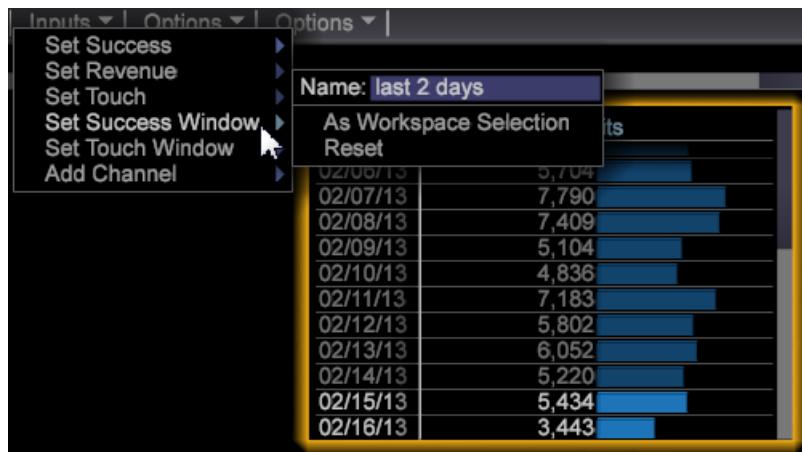
Click the **Inputs** menu and select **Set Touch**, or drag a metric from the Finder.



This will be used to derive channel metrics when dimension elements are used as inputs.

5. Set a **Success** window.

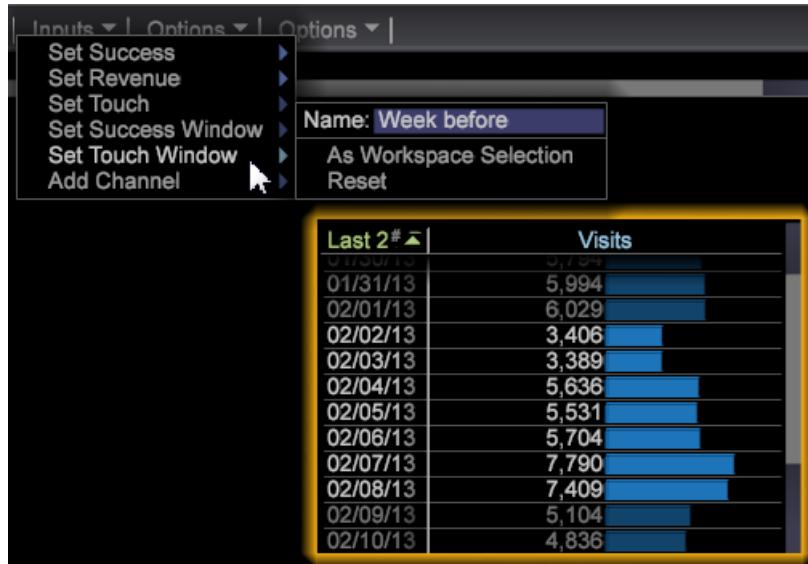
Click **Inputs > Success Window**. Select a date range from a table and then name the Success window. Click **Workspace Selection** and the selected dates will be assigned as the range of time for the Success metric.



Note: Since the Success window is a workstation selection, you can include any dimension(s) to your Success window.

6. Set a **Touch Window**.

Click **Inputs > Touch Window**. Select a date range from a table and then name the Touch window. Click **Workspace Selection** and the selected dates will be assigned as the range of time for the Success metric.



By default, the **Touch** window will be set to the same time period as the **Success** window.

7. (optional) Set a Training Filter.

You can also specify a **Training Filter** in the workspace to filter visitor data.



Note: In setting both the Success and Touch windows, you can apply the Training filter to the current workspace selections to further limit your data.



Note: The training set is always drawn from visitors who satisfy the Success window. By filtering using the Filter Editor, you can create a subset of visitors reported in the Success window.

8. Specify channel metrics that represent touches.

Either drag metrics to the visualization, or choose them from the **Inputs > Add Channel** menu. If you do not already have metrics defined for campaigns or channels, but do have dimensions representing channels, the visualization can build them for you automatically with the specification of a Touch metric.

For example, with the Touch metric set to **Hits**, and given a dimension called **Media Type** with elements that include things like **Email**, **Press Release**, **Print Ad**, and **Social Media**, the visualization will generate Channel metrics of the form **Hits where Media Type = Email** when you drag and drop the element(s) onto the visualization.

9. Press **Go**.

The Best Fit Analysis process will run, and a chart will display attributions per channel based on the selected inputs.



Note: Right-click **Model Complete** on the completed analysis to see statistics for the attribution model.



When complete, a graph will display an attribution model calculated per channel, and a distribution of the *Revenue* metric (if set). The model can be saved internally or exported to other systems.



Note: **Streaming**, **Online** and **Offline** modes produce different effects when building an attribution model based on the latency of the data being evaluated. In Streaming mode, the detail **Model Complete** message will display. In Online and Offline modes, the detail **Local Model Complete** will display.

Options menu

The **Options** menu provides advanced features to set up and display Best Fit Attribution analysis.

Options menu	Description
Set Training Filter	The Training Filter is used with the Success Window to filter the population when building the attribution model. This will provide a subset of data that includes only the visitors that you want to analyze.  Note: Experienced users can also leverage the flexibility of filters to focus beyond the time line of the Success and Touch Windows. For example, in addition to selecting a time range, you can select a set of <i>Referring Domains</i> to only examine the attribution for users from those domains.
Show Complex Filter Description	Displays the filter code for the Training Filter, Success Window, and Touch Window.
Save Model	Saves the current attribution model for future use.
Load Model	Opens a previously saved attribution model.

Options menu	Description
Presentation View	Hides the top menu bar for presentation.
Options > Advanced includes features to set the training set size and specify the approach to take in the case of a class imbalance.	
Advanced > Training Set Size	<p>Sets the training set size.</p> <p> Note: The default training size is Large for 250,000 visitors.</p> <ul style="list-style-type: none"> • Tiny = 50,000 • Small = 75,000 • Normal = 100,000 • Large = 250,000 • Huge = 500,000
Advanced > Class Balance	Identifies and defines the number of input records to generate for a class imbalance issue based on dataset size.

Reset and Remove options	Description
Reset Model	From the Reset menu, select Reset Model to clear the visualization but keep input metrics.
Reset All	From the Reset menu, select Reset All to clear the visualization and the input metrics.
Remove	Right-click on any input and select Remove to clear the metric from the selected input.
Remove All	Right-click on <i>Channels</i> and select Remove All to clear all input metrics.

Chord Visualization

The Chord visualization allows you to show both the proportion and correlation between metrics, displaying larger chords as an indication of a stronger correlation.

The Chord visualization lets you see identify correlations between metrics, allowing you to add and easily evaluate possible correlations. It also provides another view into any previously built [Correlation Matrix](#). Using the Chord visualization, you cannot identify a positive or negative correlation between the metrics—only that a correlation exists. In certain cases, determining a direct or inverse relationship can be identified by applying counter metrics.

1. Open the Chord visualization.

In the workspace, right-click **Visualization > Predictive Analytics > Chord**.

2. Select a Dimension from the menu.

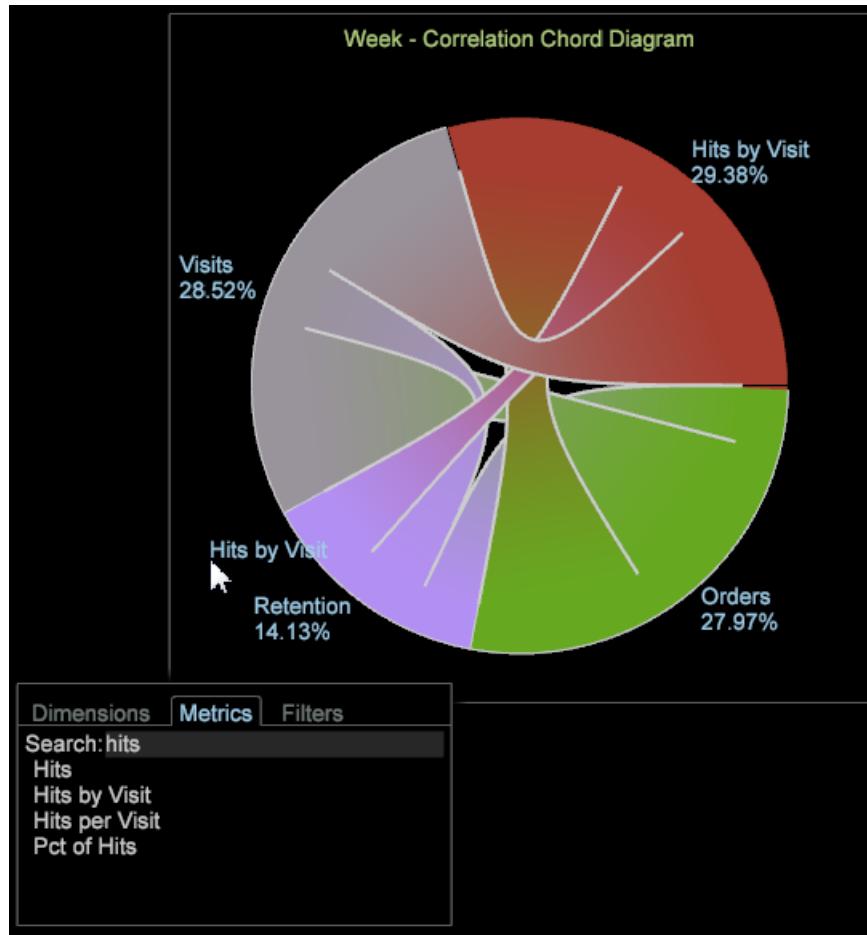
A blank visualization will open allowing you to select a dimension. The dimension name will appear at the top of the blank chord visualization.



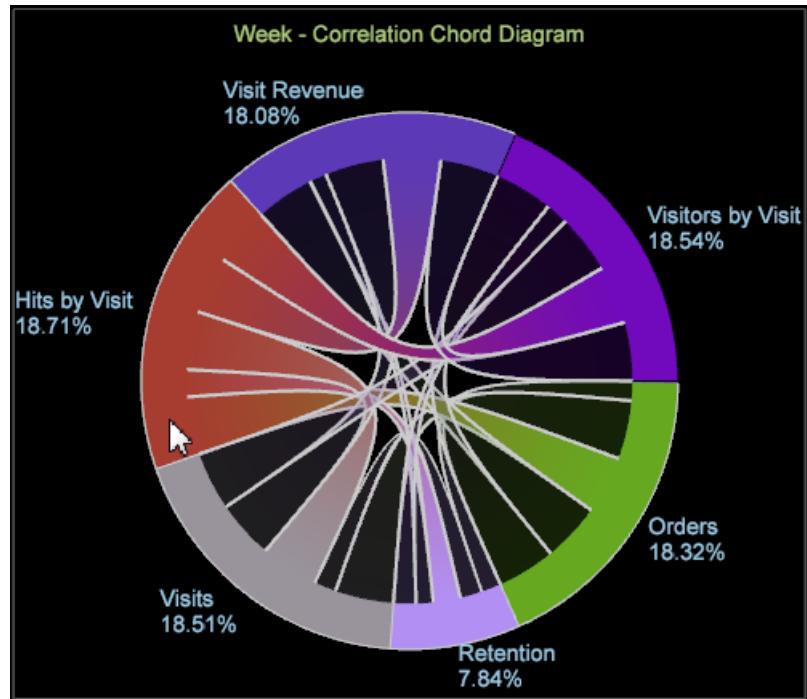
Note: If you already have a Correlation Matrix open in the workspace, you can also render it as a Chord visualization.

3. Choose metrics to correlate.

Drag metrics from the **Finder** by clicking **Ctrl-Alt** to drag metrics from the table to the chart. After two or more metrics are selected, the chart will automatically refresh and begin displaying correlation data. Continue adding metrics as needed to correlate data points.

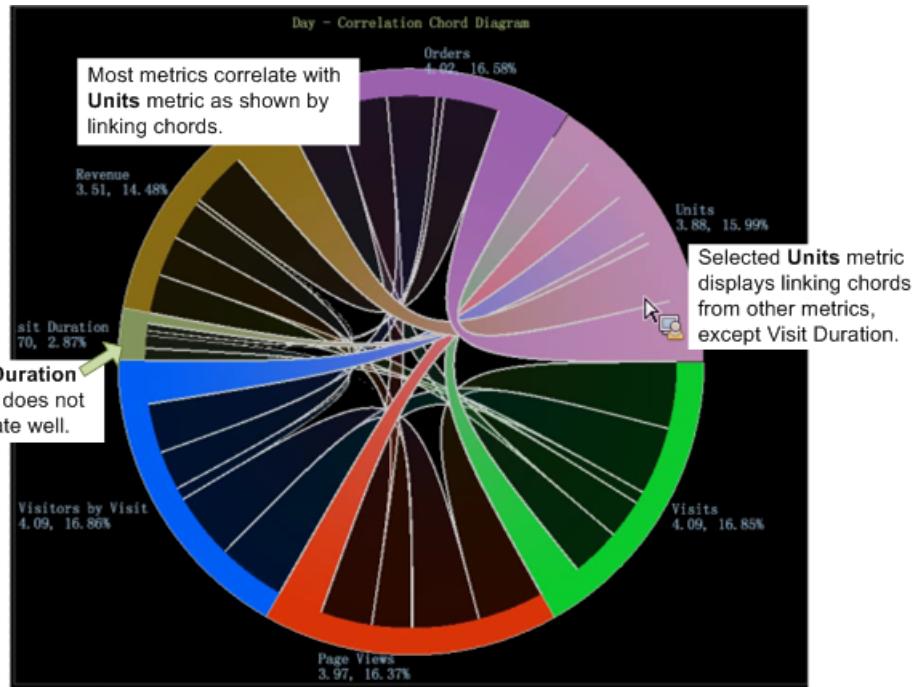


The Chord visualization displays the proportion of the whole represented by the area of each segment. Continue to add metrics as need to identify and investigate significant relationships.

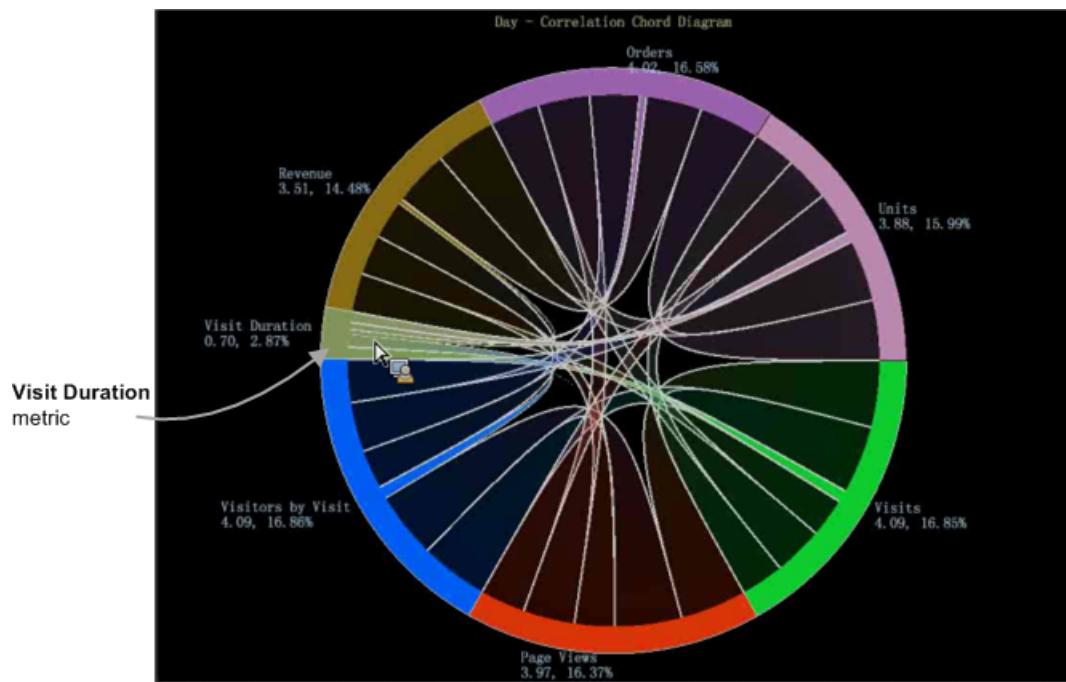


4. View the Chord visualization.

Hover over each metric in the visualization to see relationships. In the example, you can see a correlation between Units and most other metrics (except for the **Visit Duration** metric).

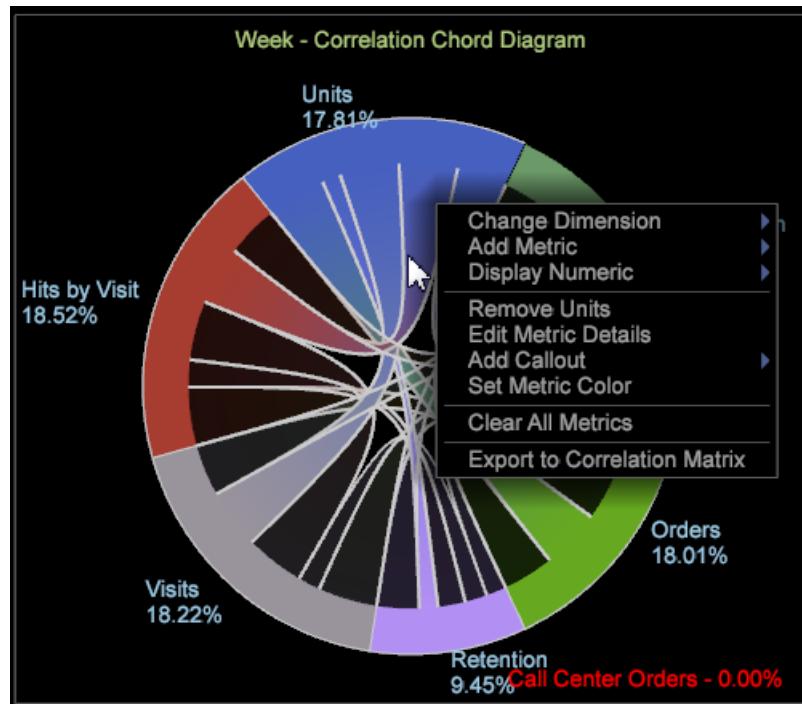


When you hover over the **Visit Duration** metric on the Chord visualization, you can see there is very little or at most weak correlation between all other metrics.



5. Change Settings.

Right-click the Chord visualization to open a menu to change the dimension, display the dimensions as absolute numbers or as percentages, remove the selected metric or all metrics, edit colors and details, and export values to a Correlation Matrix.



Regression Analysis Graph

The Bar Graph in Data Workbench now includes a regression comparison for multiple metrics across multiple graphs.

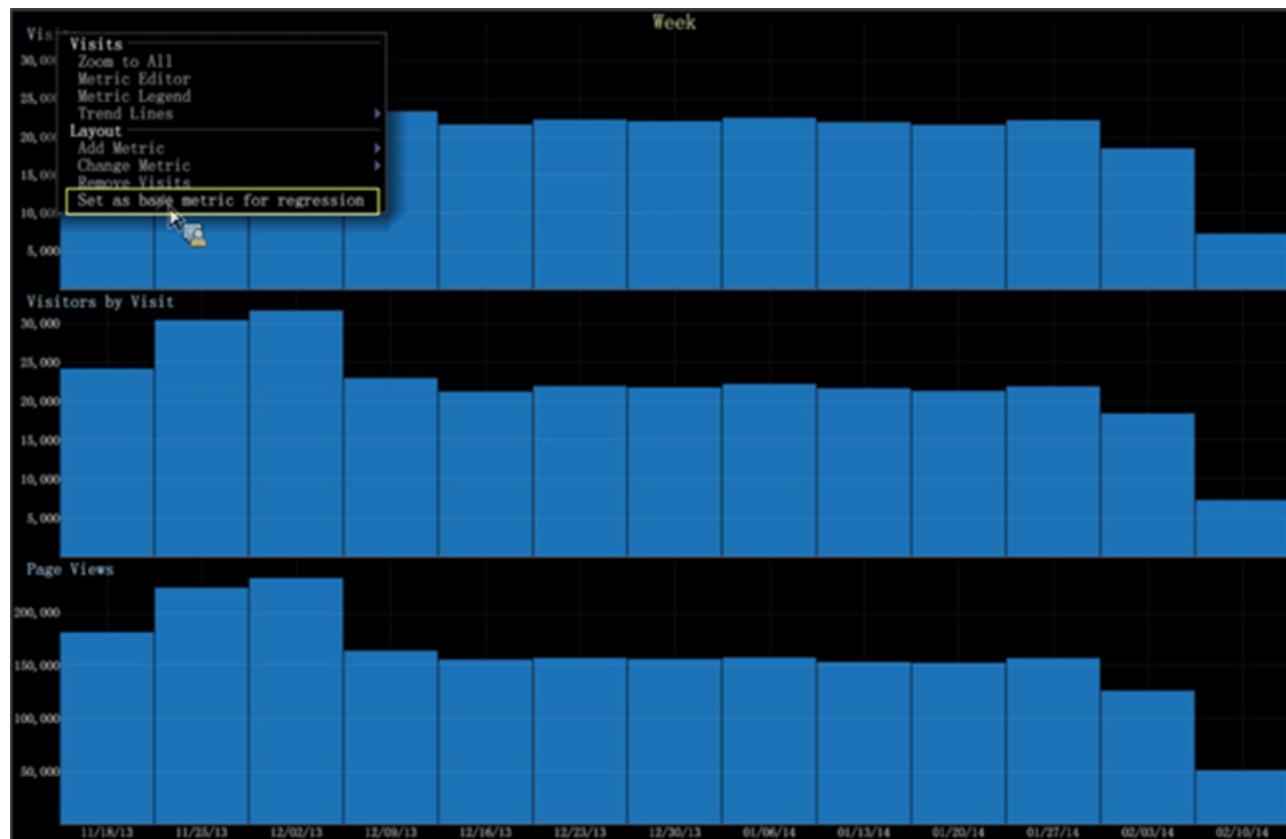
Bar graphs in Data Workbench let you regress metrics in one graph to metrics in another graph. If you have multiple graphs, you can compare a metric (as the independent variable) to a graph evaluating other metrics (as dependent variables). This lets you determine the strength of the relationship between one dependent variable (the metric established first) and a series of other changing metrics (regressions with the established dependent metric).

The regression analysis on a graph visualization allows analysts to perform "what-if" scenarios. For example, if visits increase to this level, what impact will this increase have on revenue?

Setting up Regression Analysis

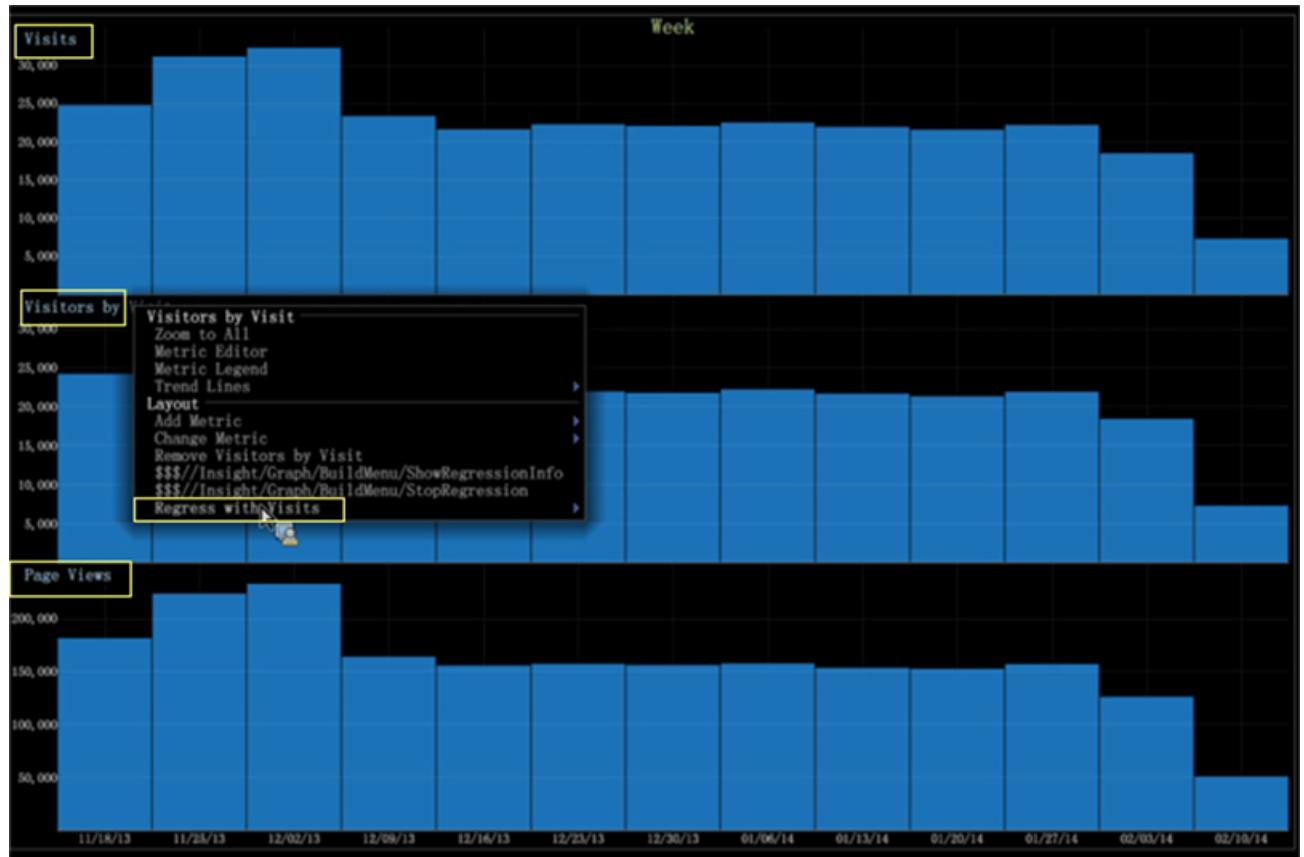
1. Select graph as a dependent metric for a regression comparison.

Right-click on the graph and select **Set as base metric for regression**.



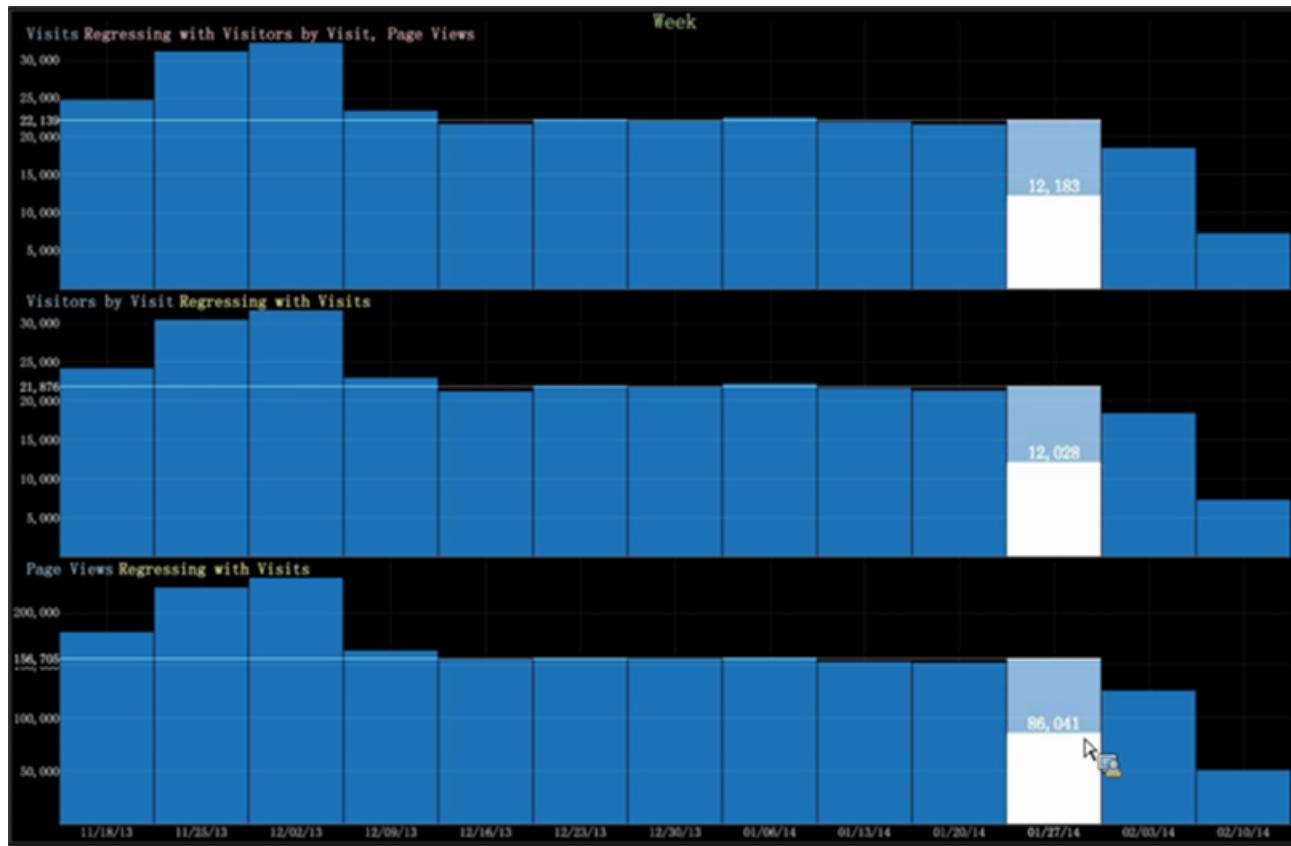
2. Set other metric graphs as independent variables.

Right-click metric and select **Regress with <base metric name>** for other metrics.

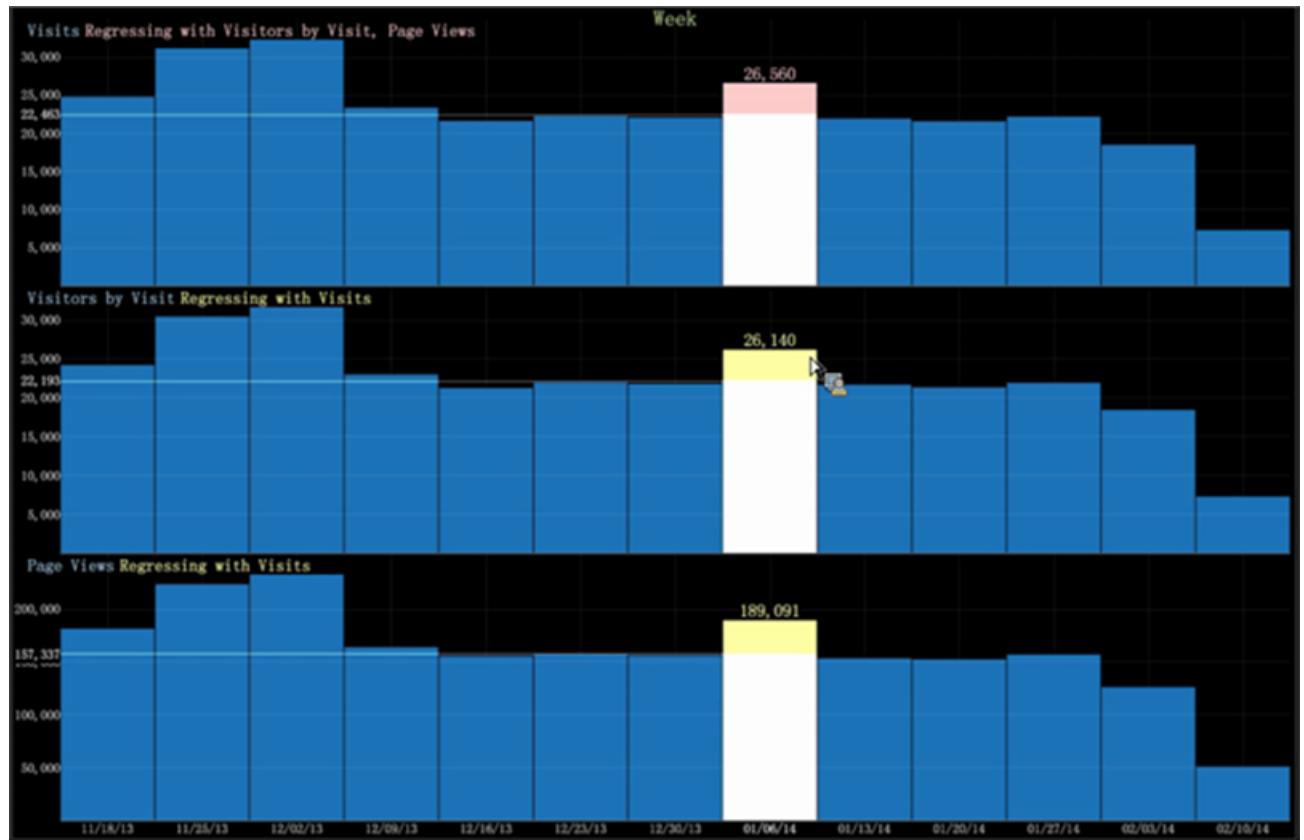


3. View regression by right-clicking on the graph to move the bar up and down.

If you right-click on the graph for a specific value, you can then see the regression ratios for each metric based for upward or downward values.



For example, if my Page Views decrease to 86,041, then the other metrics will have these values: Visits at 12,183 and Visitors by Visit at 12,028.



If Visitors by Visits values increase to 26,141, then the other metrics will be Visits at 26,560 and Page Views will be at 189,091.

Profiles and Audiences Export

Data Workbench lets you export files to integrate with Profiles and Audiences as part of an integrated Adobe Experience Cloud.

Profiles and Audiences is part of the [Experience Cloud Identity Service](#), a core service of the Adobe Experience Cloud. The Profiles and Audiences export allows audiences to be shared across the Experience Cloud using a unique Experience Cloud ID (ECID) that is assigned to every visitor and then used by Audience Manager. The ExportIntegration.exe application E:\Server\Scripts is employed to generate both Profiles and Audiences, and Adobe Target exports.

Configuring the FSU Server to use Profiles and Audiences

1. Access your FSU server.
2. Open the MMPExport.cfg file. (Server/Admin/Export/MMPExport.cfg).
3. Enter values in the all fields as required. For example:



Note: MMP/AAM integration relies on Amazon's s3 bucket for data transfer.

The s3 information required for MMP (s3) transfer can be obtained from Audience Manager team.

```
Sample MMPExport.cfg
MMP Export Configuration = MMPExportConfiguration:
s3 Bucket = string: aws_bucket_for_mmp
s3 Object Directory = string: test/files/
s3 Region = string: us-east-1
```

```

s3 Access Key = string: ZZKI62005YBA
s3 Secret Key = string: ioqwa3OpNE5
data Provider Name = string: 895
client ID = string: mcprofile2-test
client Secret = string: saea1287617212987q
username = string: mmptest
password = string: pass
numRecordsPerChunk = int:
numThreads = int:
maxRetriesOnSendFailure = unsigned int:

```



Note: The `MMPExport.cfg` file also lets you take all records, split them into sets, and create chunks of records. The chunks of records are then exported to Amazon S3. Three mandatory parameters are required to create chunks of records: `numRecordsPerChunk`, `numThreads`, and `maxRetriesOnSendFailure`.

Definition of Parameters

Parameter	Definition
<code>s3 Bucket</code>	The AWS S3 bucket where the export is transferred to.
<code>s3 Object Directory</code>	A path to save s3 files. This supports sub-directories. ★ Important: Space and multibyte characters are not allowed in the path and will create errors in the export. (The hyphen is allowed).
<code>s3 Region</code>	The AWS s3 Region where the export is sent to. Ex. us-east-1
<code>s3 Access Key</code>	AWS s3 Access Key
<code>s3 Secret Key</code>	AWS s3 Secret Key
<code>data Provider Name</code>	This will be the folder name that is used for storing segments and traits in AAM respectively. This should be unique per customer.
<code>client ID</code>	This is a unique client ID provided to a customer when provisioned for MMP.
<code>client Secret</code>	This is a unique client secret provided to a customer when he/she is provisioned for MMP.
<code>username</code>	MMP username
<code>password</code>	MMP password
<code>numRecordsPerChunk</code>	Determines the chunk size in terms of number of records. The implementation clips the user specified value to min = 1000 records (~50 KB chunks) and max = 50000 records (~2.5 MB chunks). A default value of 10000 is used in case the user does not specify this configuration property.
<code>numThreads</code>	Determines the parallelism of the chunk sending part. It accepts a value between 1 to 24 threads, and its default value is 12 threads.
<code>maxRetriesOnSendFailure</code>	Determines the number of retry attempts to be made in case of chunk send failures. Default value is 0 specifying no retries. Sleep interval of 2 seconds is used between retries.

Generating MMP export from the client

1. From the client, open a workspace and right-click **Tools > Detail Table**.
2. Add **Level**.
3. Right-click the header and select **Add Attributes**.
4. Right-click the header and select **New Master Marketing Profile Export**.



5. Expand **Query**.



6. Expand **MMP Configuration**.

-
7. (required) Enter the **MMP Segment Name** and **MMP Visitor ID Field**. These parameters cannot be left empty.
 8. The **MMP Segment Name** should match the Segment ID defined in the MMP.
 9. The **MMP Visitor ID** is the attribute column defined in step 4 that corresponds to the **Visitor ID**.
 10. Once these fields are entered, you can save the export by right-clicking the header for the export and choose **Save as "User\export"**.
 11. Open **Admin > Profile Manager** and save the export to the profile.

If all data is entered correctly, this will generate an export file in the FSU (Server/Exports) and it will also transfer the export to the AWS using the information in `MMPExport.cfg`. The log for this is provided in `Server/Trace/. eg., MMP-102014-133651- [Segment Export Name] .log`

```
Query = SegmentExportQuery:
Command = string: ExportIntegration.exe
Command Arguments = string: \"%file%.cfg\" \"%file%\""
Filter = string:
Level = string: Page View
MMP Configuration = MMPConfiguration:
MMP Segment Name = string: 12345
MMP Visitor ID Field = string: Tracking ID
Oneshot = bool: true
Output Fields = vector: 3 items
0 = ColumnDefinition:
Column Name = string:
Field Name = string: Tracking ID
1 = ColumnDefinition:
Column Name = string:
Field Name = string: PID
2 = ColumnDefinition:
Column Name = string:
Field Name = string: SID
Output File = string: MMPTest.txt
Output Format = string: %1%\t%2%\t%3%\r\n
Schedule End Time = string:
Schedule Every = string:
Schedule Start Time = string:
Time Limit (sec) = double: 1800
```

Configuration Details	Description
MMP Segment ID	Required. This is an identifier you would define first in Audience Manager.
MMP Visitor ID Field	Map the MCID.

Segment Export with Custom Headers

Create custom column export headers for your segment export files to add easily understood descriptions for exported segments. This export feature also lets you output as TSV and CSV files.

New functionality has been added to Segment Export, including the ability to export with a header, or in CSV and TSV formats.

You can create column headers for your export files.

Creating a New Segment Export

1. Open a workspace and right-click **Tools > Detail Table**.
2. Right-click and select **Add Level > Extended** > Choose an item.
3. Right-click title and select **Add Attribute**.

Select a dimension from the menu.

4. Right-click title and select **Add Metric**.

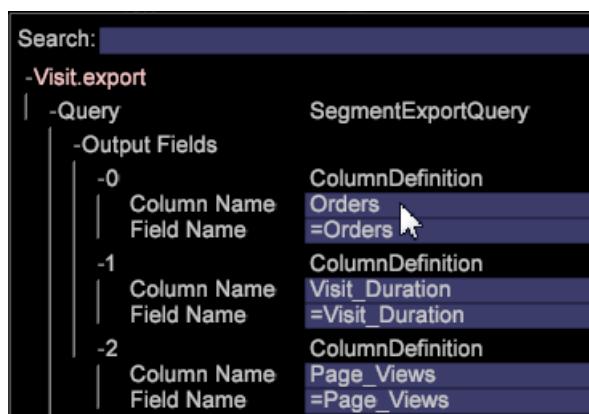
Select a metric from the menu.

5. Right-click title and select **New Segment Export**.



New Segment Export with Header automatically populates the Column Name with the name of the metric.

New Segment Export requires you to set a custom name.

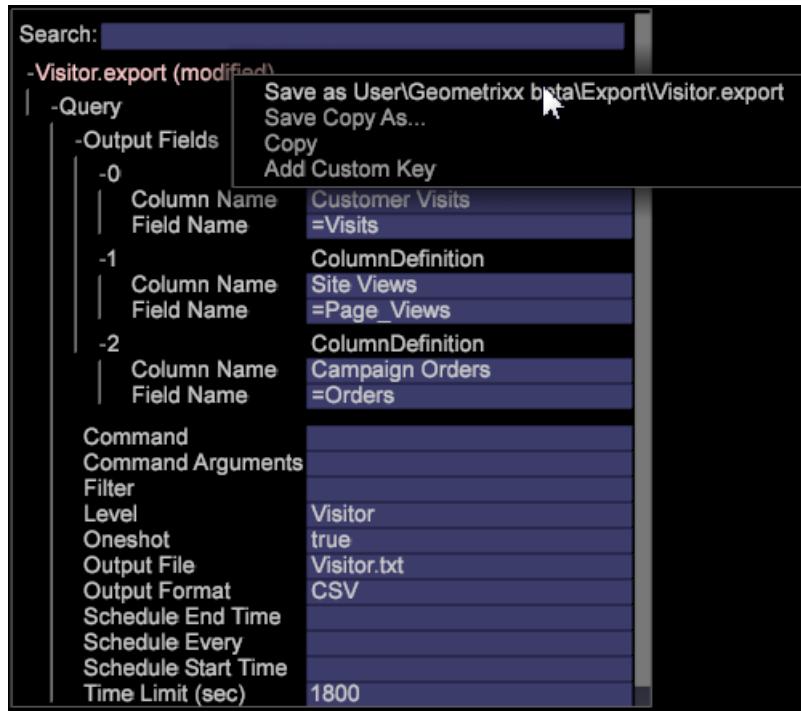


Note: The Column Name field cannot be left empty or the header will not be present.

6. Right-click and name the segment and then click **Save Export File**.

An export window will open.

7. Right-click the export name and click **Save as <export filename>**.



- Right-click Admin > Profile Manager > Expand Export. Find the export file you just created and save it to an existing profile.



Clustering 2.0

The Cluster Builder now includes a KMeans++ algorithm (only the KMeans algorithm was previously supported) that uses a faster approach to finding centers for an expedited cluster-generation process.

KMeans Algorithms

In the *Cluster Builder*, you can now select **Options > Algorithm** to select algorithms when defining clusters.

- KMeans.** This algorithm uses canopy clustering to define the centers of the cluster.

- **KMeans++**. This algorithm expedites cluster building when running against large sets of data.

KMeans++ is an improved implementation of KMeans clustering algorithm because it provides better initialization of initial k centers. (The original KMeans algorithm chooses initial centers randomly.) KMeans++ selects the first center randomly. The remaining k-1 centers will be chosen one by one based on the distance a data point is to the closest existing center. The furthest data points have a better chance to be chosen as a new center than nearby data points. After the initial center is chosen, the procedure is performed exactly the same as the original KMeans clustering.

The workflow for KMeans++ is exactly the same as the workflow for KMeans clustering, except that you need to select **Options > Algorithm > KMeans++** in the cluster builder.



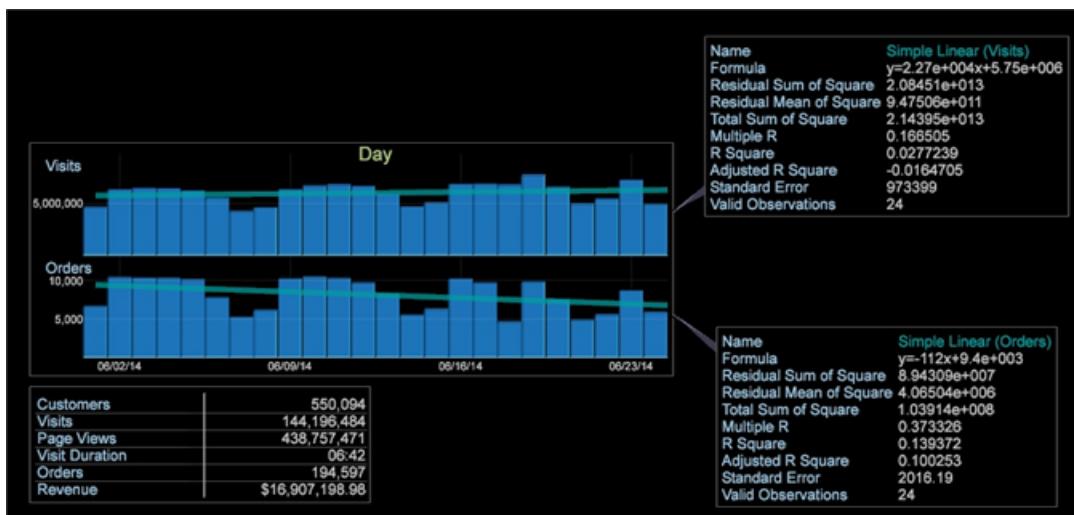
Note: Each DPU runs its own KMeans++ procedure on its own data portion. If the DPU has enough available memory (the ratio is configurable in the PA Server.cfg file), then the data of those involved variables will be brought into memory. The remaining k-1 initial center selection and converging iterations all happen in memory, which is faster than the previous KMeans clustering.

Trend Lines

Trend lines lets you overlay graphs to compare and interpret data.

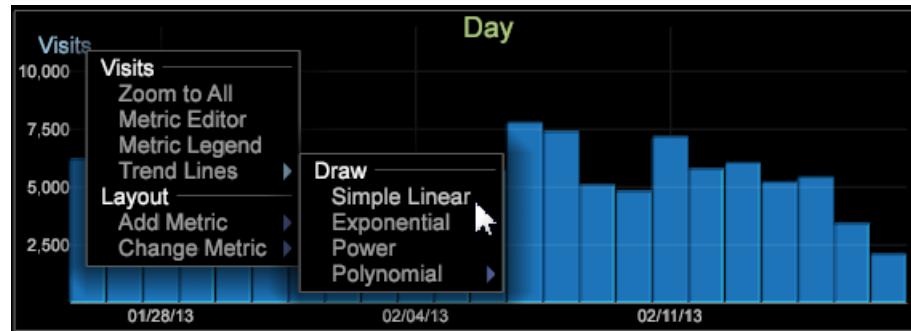
Like the [Scatter Plot](#) visualization, you can now set trend lines on a graph visualization to display the rate of change based on linear, exponential, power, or polynomial lines. The Trend Line feature allows you to overlay trend lines on a graph, most commonly over a Time dimension.

For example, in this graph comparison, we can see that Visits are trending up, but Orders are trending down.



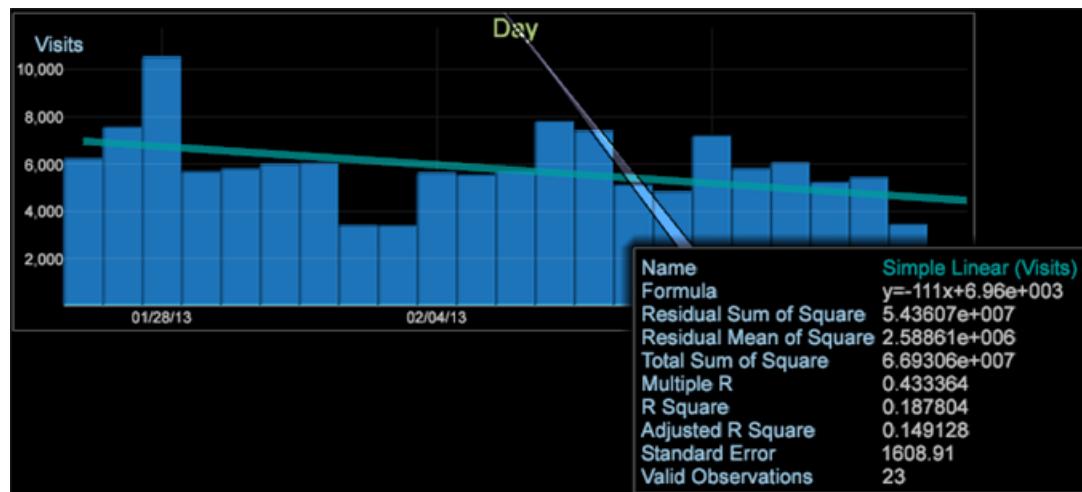
To add a Trend Line

1. Open a graph and right-click the metric name in the upper left corner.
2. Click **Trend Lines** and select from the options.



You can select the trend line to appear over the graph as **Simple Linear**, **Exponential**, **Power**, or **Polynomial**. Polynomial will create a polynomial regression trend line. Simple Linear will create a trend line as the rate of change along the regression line. Exponential calculates a trend line as $y = b \cdot \exp(a \cdot x)$ and Power as $y = b \cdot x^a$.

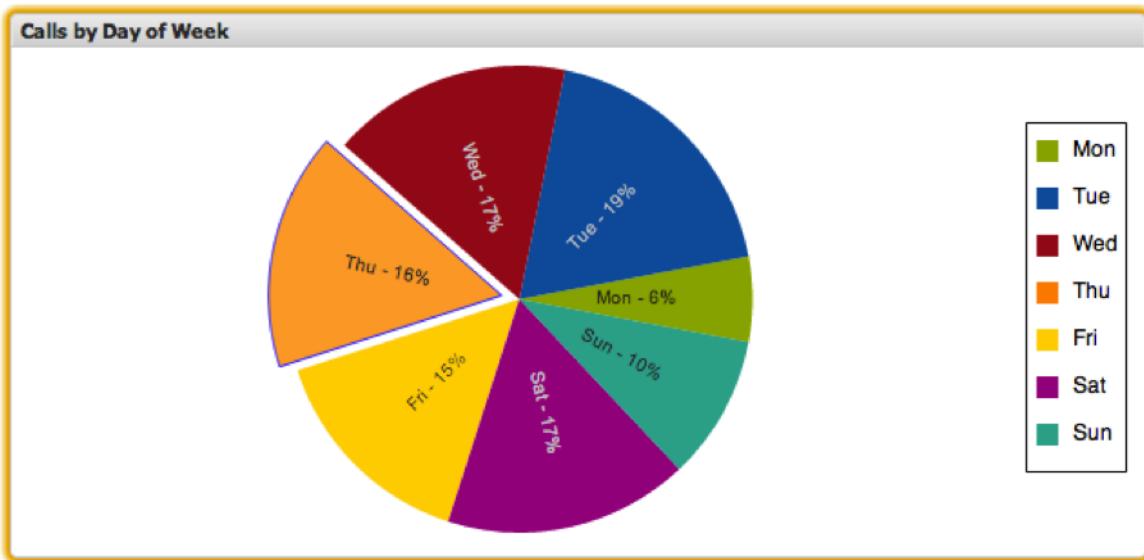
The trend will be calculated and rendered on the graph, and a callout will open displaying detailed information of the trend equation.



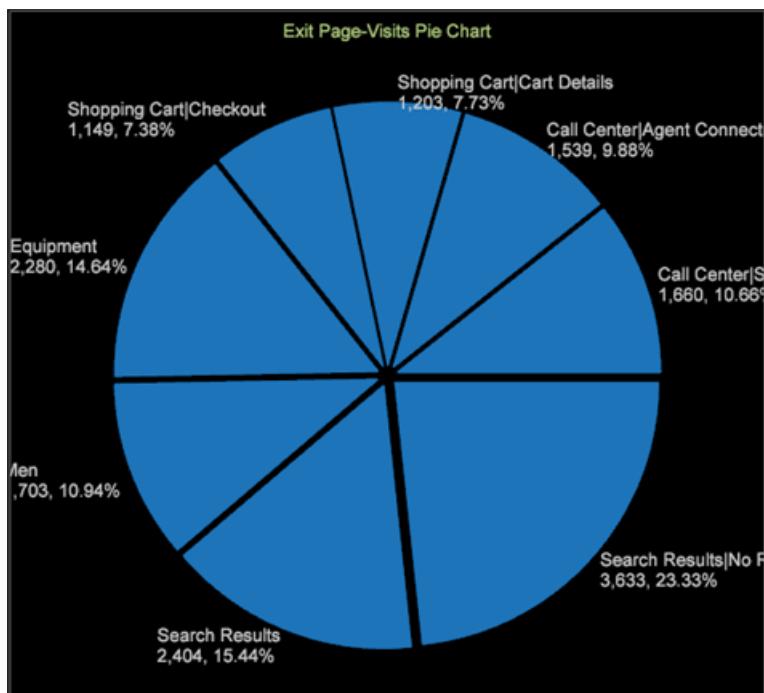
Pie Chart Update

Updates to the Pie Chart visualization lets you use default colors identified in a legend, or set colors based on the color chart.

When you open a pie chart, the colors for data sets are set by default with each identified in the legend.

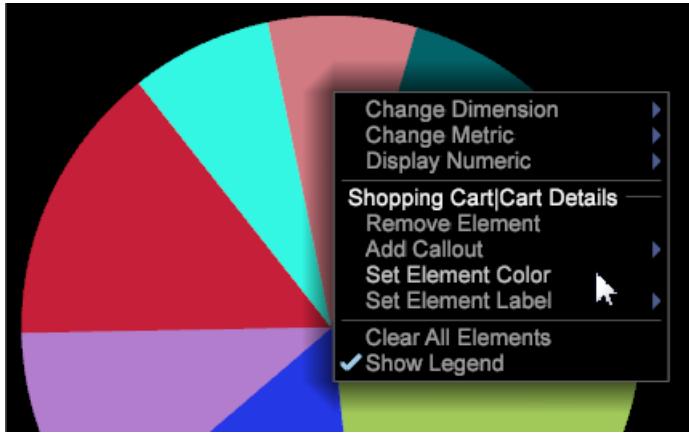


The legend can be toggled on and off by right-clicking the chart and selecting **Show Legend**. The result is a pie chart without color codes and elements identified in callouts for each section.



Customizing your Pie Chart

Right-click the pie chart to open a menu to modify your pie chart.



The menu allows you to change your dimension elements, metrics, and numeric display, as well as customize chart colors and labels.

Query String Grouping

Query String Grouping lets you integrate a large number of fields together.

Query String Grouping is specific to each profile, but works well in transformations as shown in this example:

1. Create the pairs you wish to bundle by adding a custom configuration file (E:\...\Dataset\Log Processing\SC Fields.cfg) and then adding the Transformation Type *BuildNameValuePair* as a parameter.

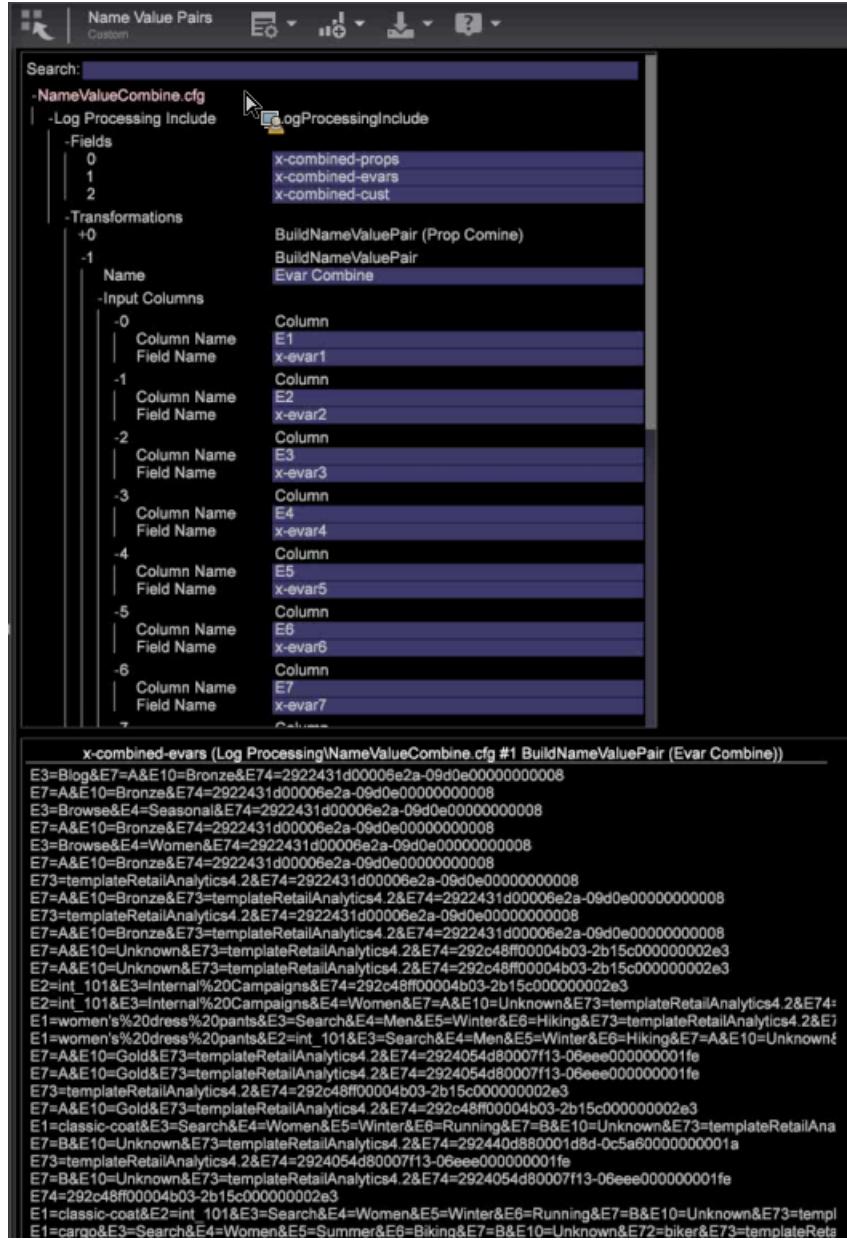
```
2 = BuildNameValuePair:
    Comments = Comment: 0 items
    Condition = AndCondition: 0 items
    Delimiter = string:
    Input Columns = vector: 1 items
    0 = Column:
        Column Name = string: e100
        Field Name = string: x-cust100
        ... (all the fields you wish to build)
        Name = string: Custom Events
        Output = string: x-event-list
```

2. Create a new file for extracting the condensed data into the fields you wish to use by adding a custom configuration file (E:\...\Dataset\Transformation\SC Fields Transformation.cfg) and then adding the Transformation Type *ExtractNameValuePair* as a parameter.

```
2 = ExtractNameValuePair:
    Comments = Comment: 0 items
    Condition = AndCondition: 0 items
    Delimiter = string:
    Input Field = string: x-event-list
    Name = string: Custom Events
    Output Columns = vector: 1 items
    0 = Column:
        Column Name = string: e100
        Field Name = string: x-cust100
        ... (all the fields you wish to extract)
        Name = string: Custom Events
        Output = string: x-event-list
```

Other Uses

If you have many fields with custom evars, props, and variables, during log processing you can build a name value pair to combine fields in a report. For example, you can build named-value pairs into combined fields to reduce the tempDB file size.



Latency Analysis

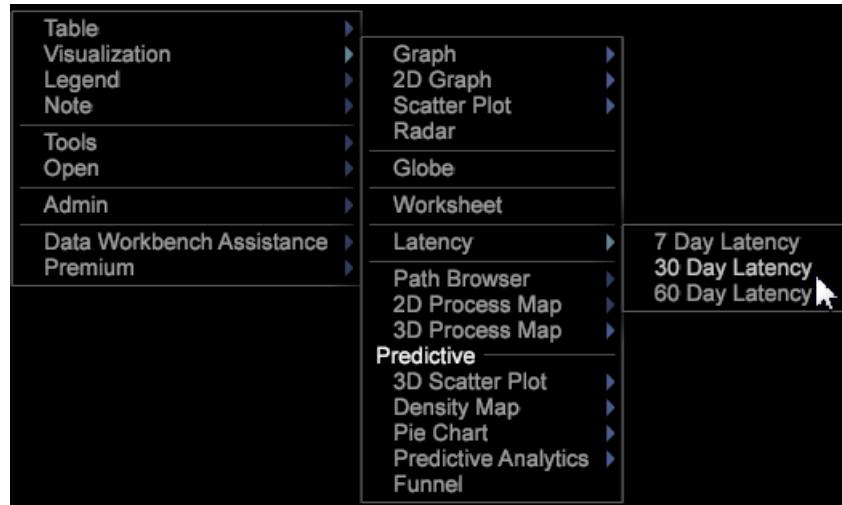
The Latency visualization lets you analyze latent customer behavior within a set number of days before or after an event occurred after a campaign or other type of event or period of time.

The **Latency** visualization allows you to set a metric to identify behavior before, during, or after an event occurred to determine its effect on customer behavior. For example, you can identify the effectiveness of a marketing campaign by viewing the revenue a week after an event occurred. Or you can show customer behavior a week previous to the event as a baseline to see an event's effect on behavior.

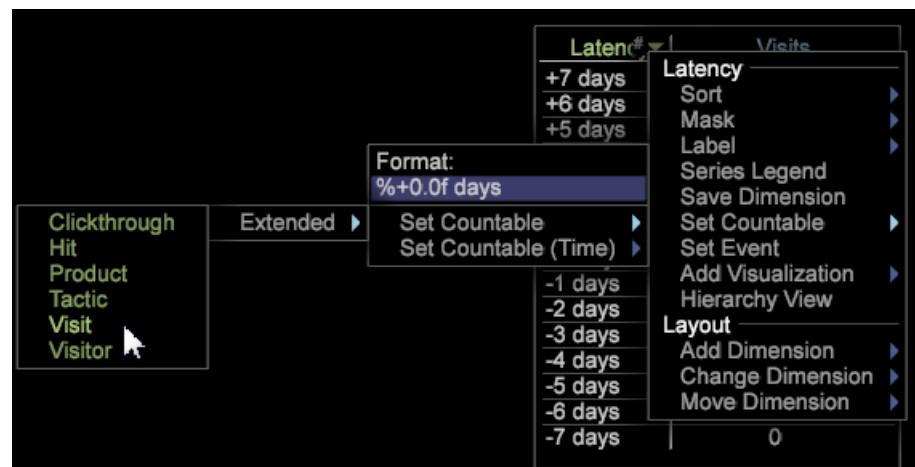
Using the Latency visualization, you can change the Time dimension from day, hour, week, or other time period, to visit, click through, hit, or other countable dimension.

Setting up Latency Analysis

1. Open a workspace and right-click **Visualization > Latency**.
2. From the menu, select the number of days to analyze for latency events.



3. Right-click on the **Latency** label to open menu.
- Using the **Set Countable** option, set a countable dimension such as visit, clickthrough, hit, and others.
- Using the **Set Countable (Time)** option, set a time dimension such as day, hour, week, day of week, and hour of day.



4. Modify the Latency visualization.

Right-click on the **Latency** header to select menu options to change countables, dimensions, to re-label, or to select or change other settings.

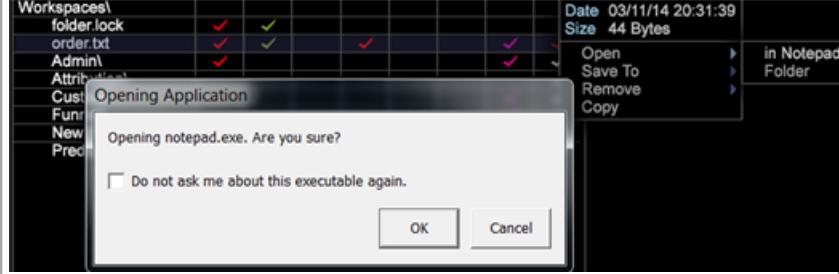
Data Workbench 6.21 Update

Data Workbench 6.2.1 provides new features and bug fixes.

New Features

Data Workbench 6.2.1 includes these new features:

Features	Description
Vertical Scroll Bars in Text and Wrapping Text features.	<p>Text boxes now have vertical scroll bars and wrapping text.</p>  <p>The "local sample" is a compact version of the dataset that is stored on your Data Workbench Client machine. It provides both instant approximate answers for your queries, as well as the capability to work offline.</p> <p>Visual Server is configured by default to provide Data Workbench with a 256MB sample of data. Some machines, especially those with less than 1GB of RAM, may perform poorly with so much data. In these cases, we recommend decreasing the Maximum Sample Size (MB) setting in the Insight.cfg file to 100:</p> <p>1) Change the Maximum Sample Size (MB) setting to the new value.</p>
Sorting Thumbnails on the worktop	<p>Names sorted on the worktop are now insensitive to the case of the character, sorting alphabetically AaBbCcDd instead of ABCDabcd .</p> 
Searching dimensions based on the parent dimension.	<p>In the Finders panel, you can now right-click on the Dimensions tab and click select Dimension Type > By Parent. A list of top-level countable dimensions will display. When you select one of these parent dimensions, a list of its subordinate dimensions will appear in the search results.</p>

Features	Description
	
Dialog asking to open an external application.	<p>You will now be presented with a dialog box the first time you try to open an external application in Data Workbench.</p> <p>For example, if you open a text file in Notepad, you will get the following message.</p>  <p>This will also create a local file called <code>InsightSES.dat</code> placed in the client install folder.</p>
Changing Toolbars to Buttons	<p>You can opt out of using the new toolbar icons provided in Data Workbench 6.2. by changing the <code>Toolbar Icons</code> argument in the <code>insight.cfg</code> file to <code>false</code>.</p> <pre data-bbox="649 1220 1488 1248"><code>Toolbar Icons = bool: false</code></pre> <p>You will need to restart the client for the change to take effect.</p>
Reset options in Propensity Scoring and Decision Trees updated	<p>In the Propensity Scoring (Tools > Predictive Analytics > Propensity Score) and the Decision Tree (Visualizations > Predictive Analytics > Classifications > Decision Tree Builder) visualizations, you now have two reset options:</p> <ul style="list-style-type: none"> Reset Models—Clears out the model but maintains the settings and inputs. Makes the Go button selectable. Reset All—Resets all settings (as in previous design).

Bug Fixes

- The **Browsers** and **Operating Systems** lookup files will not be updated within the legacy **Traffic** profile (for example, `Lookups\Traffic\Browsers.txt`). Instead, configuration of the **Traffic** profile will utilize the `DeviceAtlas` bundle (`Lookups\DeviceAtlas\DeviceAtlas.bundle`) to provide this configuration information.
- Data Workbench 6.2.1 will be the last release to provide a download of the 32-bit client application. All future client application downloads will be 64-bit and continue to require Windows 7 or newer. Memory limitations

of the 32-bit application are addressed with the introduction of the 64-bit application beginning with the 6.1 release.



Note: The 32-bit version of the Data Workbench client application may experience potential issues related to memory limitations when running predictive models using the clustering and scoring features.

Data Workbench 6.2 Release Notes

Data Workbench 6.2 release notes include new features, upgrade requirements, bug fixes, and known issues.

To view previous features and fixes based for each past release, see the [release note archives](#).

New Features

Data Workbench 6.2 includes these new features:

Features	Description
Data Workbench Client UI Updates on page 83	Data Workbench 6.2 includes new user interface features: <ul style="list-style-type: none">• New bookmarks panel• New Icons in Workspace on page 84• Drag Workspace Views on page 85• Quick Keys to Change Workspace Views on page 85
Decision Trees	Decision trees are a predictive analytics visualization used to evaluate visitor characteristics and relationships. The Decision Tree Builder generates a decision tree visualization based on a specified positive case and a set of inputs.
Finders on page 92	Use Finder panels in Data Workbench to select metrics, dimensions, and filters. These panels provide search support, sorting options, and drag and drop capabilities.  Note: With the new Finder panels, you can export a list of your Dimensions, Metrics, and Filters to an MS Excel spreadsheet.
Binary Filter in the Correlation Matrix on page 169	The Binary Filter has been updated with new features, requiring you to rebuild any Correlation Matrix with a Binary Filter built in previous versions.
Density Map	The density map is a visualization that displays elements as shaded rectangles within a square map.
Attribution Profile	To quickly analyze attribution values (events to attribute responsibility for a successful conversion or sale), Data Workbench provides a rules-based Attribution profile with features for the Architect to set up the Attribution reports and the Analyst to run the reports.
Analytics Reports on page 107	Report templates standardize Adobe Analytics' reports for users of the data workbench who utilize the Adobe SC profile. These reports are identical to reports employed in Marketing Reports & Analytics (formerly SiteCatalyst).
3D Scatter Plots	A 3D Scatter Plot graphs the elements of a data dimension (such as Days or Referral Site) on a three-dimensional grid where the x, y, and z axes represent various metrics.

Bug Fixes

- Updated the Visual Site lookup file to address search engine changes to the query search term.
- Fixed inaccurate error message, "Failed to import workspace", when importing a workspace in the client workstation even though the import was successful.

-
- Workstation connection error displaying "412 Configuration Conflict" message is now replaced with user friendly message that identifies system action.
 - The "post" command can now be executed in Report Server.
 - Fixed user interface errors in client user interface for Simplified Chinese.
 - Adobe Analytics updated the data feed that powers Data Workbench to take advantage of the Master Marketing Profile that integrates with the Adobe Marketing Cloud. All Data Workbench users were required to prepare their environment for this transition by April 21, 2014.

The master marketing profile was introduced to provide a complete view of customers across Adobe Analytics. This new service is available within the Adobe Marketing Cloud to drive further value across analytics tools to start establishing the foundation for these features within Analytics. The new Marketing Cloud visitor identifier will be added to the data feed, along with other enhancements and improvements to adapt to the new data feed and global visitor identifier.

- When importing a workspace, an error message is displayed even though the import was successful.

Upgrade Requirement

- The Attribution profile is configured for users who have implemented the Adobe SC profile to employ the Analytics (SC/Insight) data feed. By default, the Marketing and Conversion events are employed as the default interactions evaluated in the rules-based models. See [Deploying the Attribution Profile](#) for additional information.
- For users of the Adobe SC profile upgrading to Data Workbench 6.2, if you are not using the default configurations, verify that the x-bot_id value in the SC_Fields.cfg file is being decoded properly and that the x-bot_id field is listed properly in the Decoding_Instructions.cfg and the Exclude_Hit.cfg files. This will only be an issue if you have modified the configuration file from the default configuration.
- If you have deleted unused fields in the Dataset > Log Processing > SC_Fields.cfg file for the Adobe SC profile, you will need to update to accommodate updated field values used for the Attribution profile (see [Deploying the Attribution Profile](#)).

Known Issues

- When 3D Scatter Plot Visualization includes callouts, the zoom may display plots outside the border of the visualization.

Workaround: Zoom the 3D Scatter Plot first and then add callouts to your visualization.

- Dragging metric from Finders panel to Metric Legend outside of the metric column will delete the Metric Legend from the workspace.

Workaround: Users that wish to drag metrics to the Metric Legend should drop in the first column (metrics column).

- Print Workspace using Sidebar and Both options will not include the Copyright information on the printed page.
- Using Workstation in remote desktop session will crash when renaming workspaces.
- (In Simplified Chinese version) Actual report outputs are valid in Report Server but email subject lines and attachment file names are garbled.
- (In Simplified Chinese version) When using word wrap in the Worksheet visualization, localized words are not being wrapped correctly resulting in junk characters added to the string.
- (In Simplified Chinese version) Unable to launch Insight.exe if the installation directory is named with non-English characters.

Workaround: Keep default names or rename using only English characters in the folder path to launch executables.

Data Workbench 6.2 features

Data Workbench 6.2 includes the following features.

Upgrade Requirement

- The Attribution profile is configured for users who have implemented the Adobe SC profile to employ the Analytics (SC/Insight) data feed. By default, the Marketing and Conversion events are employed as the default interactions evaluated in the rules-based models. See [Deploying the Attribution Profile](#) for additional information.
- For users of the Adobe SC profile upgrading to Data Workbench 6.2, if you are not using the default configurations, verify that the x-bot_id value in the SC_Fields.cfg file is being decoded properly and that the x-bot_id field is listed properly in the Decoding_Instructions.cfg and the Exclude_Hit.cfg files. This will only be an issue if you have modified the configuration file from the default configuration.
- If you have deleted unused fields in the Dataset > Log Processing > SC_Fields.cfg file for the Adobe SC profile, you will need to update to accommodate updated field values used for the Attribution profile (see [Deploying the Attribution Profile](#)).

Data Workbench Client UI Updates

Data Workbench 6.2 includes new user interface updates to the bookmarks panel, new icons in the workspace toolbar, the ability to drag the workspace within a screen, new quick keys, and updates to the pie chart visualization.

[New Bookmark Features](#) on page 83

[New Icons in Workspace](#) on page 84

[Drag Workspace Views](#) on page 85

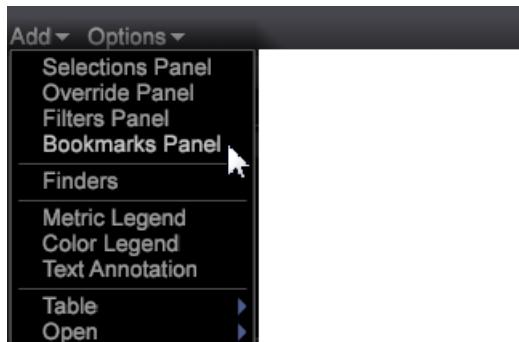
[Quick Keys to Change Workspace Views](#) on page 85

New Bookmark Features

You can now bookmark significant workspaces to quickly move between visualizations and reports employed in your workflow.

Working with Bookmarks

1. Bookmark a workspace by clicking the Bookmark icon  in the upper right corner of the toolbar.
2. Click **Add > Bookmarks Panel** in the left pane to open a list of bookmarks.



3. To open a bookmarked workspace, click a workspace name in the **Bookmark Panel**.



The selected workspace will open. When you click another bookmarked workspace, the previous workspace will close and the newly selected workspace will open, allowing you to quickly navigate through your workflow.

To delete a bookmark:

- In the Bookmark Panel, right-click and select **Remove <bookmark title>** to delete a selected bookmark, or select **Clear All Bookmarks** to delete all bookmarks.
- You can also right-click on the workspace in the thumbnail view within the worktop and select **Clear Bookmark**.



Important:

- 25 bookmarks can be saved.
- If you add a bookmark and then move the location of the workspace, the bookmark will be invalid and must be deleted from the Bookmark Panel and reset.

New Icons in Workspace

Data Workbench 6.2 now replaces the text in the workspace with icons. You can still hover over and see the tool tip message identifying the icon, including **File**, **Add**, and **Export**.



A new **Help** icon is added to access the documentation and other knowledge centers, including the following links:

Documentation links	Description
Marketing Reports & Analytics	Open to the Adobe Marketing Reports & Analytics help page.

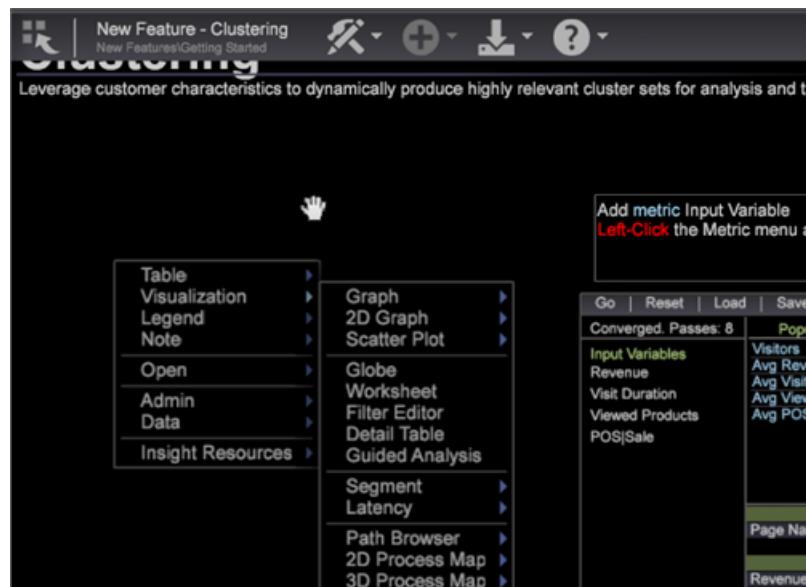
Documentation links	Description
Idea Exchange	Open to the Idea Exchange login . This online portal allows users to provide update changes and enhancement ideas to data workbench. These customer-focused ideas can then be voted on by all users.
Help	Open the Data Workbench documentation . You can also press <F1> to open help within a workspace.
About	Open to identify the client version of data workbench.



Note: You can also press <F1> to open the documentation from a workspace.

Drag Workspace Views

If a workspace is larger than the viewable screen, you can move the view to see all elements within the workspace. You can click in the background (outside of the visualizations and tables) and drag the screen to move the viewable area within the workspace. The cursor will change to a hand icon when dragging the view within the workspace frame.



Quick Keys to Change Workspace Views

New quick keys let you resize and refit workspaces between window and full page views. See the [Quick Reference](#) for more keyboard quick keys.

Commands	Quick Keys	Combined menu commands
Full screen view. Workspace fills the screen and refits to the new size.	Ctrl plus Ctrl + (on keypad) <i>or</i> Ctrl Shift + (on keyboard)	<ul style="list-style-type: none"> File > Page Size > Fill Screen <i>followed by</i> File > Refit Workspace

Commands	Quick Keys	Combined menu commands
Window view. Workspace displays in a standard window view and refits to the new size.	Ctrl minus Ctrl -	• File > Page Size > Standard <i>followed by</i> • File > Refit Workspace

Decision Tree Builder

Decision trees are a predictive analytics visualization used to evaluate visitor characteristics and relationships. The Decision Tree Builder generates a decision tree visualization based on a specified positive case and a set of inputs.

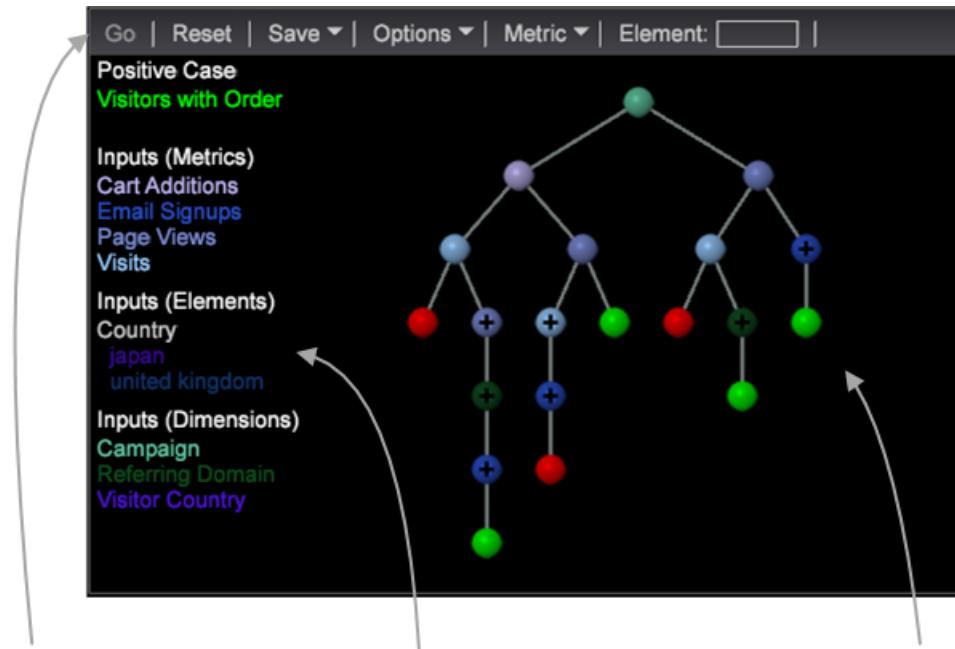
A Decision Tree is a binary classifier with a set of rules (or filters) identifying visitors who satisfy specific rules based on a positive case. A decision tree sets rules to classify visitors who satisfy (or do not satisfy) this positive case. These rules generate a tree map to provide a level of confidence to meet these positive case results.

A Decision Tree is built by examining inputs at each level and choosing the one that provides a maximum gain of information at a specified split point. Split points for each variable-level generates two sets:

- Values less than or equal to the split point, and
- Values greater than the split point.

Use decision trees to

- Perform meaningful analysis and interpretation in less time.
- Employ automated segment generation.
- Quickly make inferences from a model based on a large amount of data.



Toolbar and Menus	Input Listing	Tree Display
The toolbar includes buttons and menu commands for the Decision Tree, including features to set the Positive Case and add Input Listings.	This area displays the inputs into the tree model. They are color coded to match nodes in the Tree Display area.	This area displays the tree model with leaf nodes color-coded based on its prediction: green for a True prediction of the Positive Case, and red for a False prediction.

<p>Like other visualizations, the Element box lets you drag and drop Dimension and Elements, although you can also drag directly from the Finders pane.</p> <p>For additional information, see Decision Tree Options on page 90.</p>	<p>Right-clicking on an input allows you to remove the input from the model and reset. If you hover over a tree node, it will display the split conditions along the branch to that node and the prediction at that node with its confidence value.</p>	<p>The split nodes are color coded to the inputs that match their selection condition. Hovering over a node displays information about the split and expands the inputs listing to display the split points along the branch and the distribution of the training set.</p> <p>Nodes below a threshold are not displayed by default. Click on an expandable node (indicated by a + symbol) to explore a branch. Click on the root node to return to the full tree display.</p>
---	---	---

Building a Decision Tree

Set up a Decision Tree by identifying a positive case and adding metric and dimension inputs to evaluate the data and explore the decision tree.

Follow these steps to build a decision tree.

1. Open a new workspace.

After opening a new workspace, you might need to click **Add > Temporarily Unlock**.

2. To open the Decision Tree Builder, right-click **Visualization > Predictive Analytics > Classification > Decision Tree Builder**.

3. Set a **Positive Case**.

You can define a positive case for a decision tree by selecting dimensions in a Finder or dimension elements in a table, or by designing a filter in the Design Filter. In fact, the positive case can be a combination of multiple selections in the workspace including filters, dimensions, elements, and all types of Data Workbench visualization values.

- **Design and Apply a Filter** as a positive case. Right-click in the workspace and select **Tools > Filter Editor** to design and apply a filter.
- Add **Dimensions** as a positive case. In the workspace, right-click and select **Tools > Finders** (or select **Add > Finders** in the left pane). Type a dimension name in the **Search** field and then select a dimension.
- Add **Metrics** as a positive case. Right-click and select **Tools > Finders** or select **Add > Finders** in the left pane to open a Metrics table. Select a metric as your positive case.
- Add **Dimension Elements** as a positive case. Right-click in the workspace and select **Table** to open dimension elements, then select from the dimension elements to set your positive case.

4. Click **Options > Set Positive Case**.

This sets the positive case and lets you name it. The name will appear under the **Positive Case** heading in the workspace.



Note: When you set the positive case the Decision Tree uses the current workspace selection, which can be defined as the Visitors (or whatever top-level countable is defined, but in most cases Visitors) that match the current selection within the workspace. These combine as a single filter for a single positive case (not multiple positive cases).

Clicking **Set Positive Case** when there is no selection will clear the positive case.

5. (optional) Select **Set Population Filters** to define the visitor population to be classified.

If no population filter is applied, then the training set is drawn from all visitors (default is "Everyone").



Note: Click the **Show Complex Filter Description** to view the filtering scripts for the Positive Case and Population Filter.

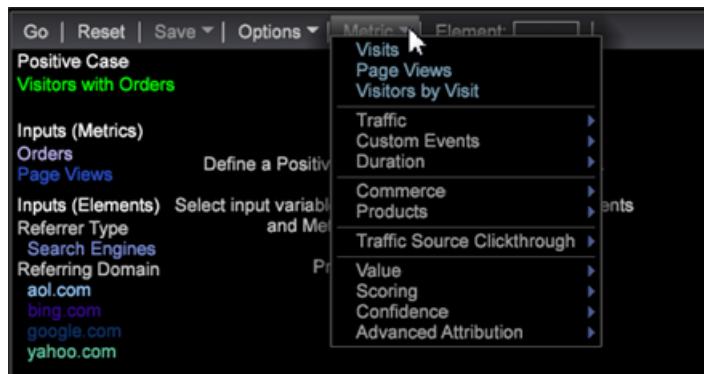
6. Add **Metrics, Dimensions, and Dimension Elements** as inputs.

You can select inputs by dragging and dropping from the Finder panels or from tables for individual dimension elements. You can also select from the **Metrics** menu in the toolbar.

- Add **Metrics** as inputs.

Select Metrics from the toolbar. Press **Ctrl + Alt** to drag one or more metrics to the Decision Tree Builder.

The metric will appear in the **Input (Metrics)** list as an input with unique color-coding.



- Add **Dimensions** as inputs.

In the workspace, right-click and select **Tools > Finder** and type the dimension name in the **Search** field. Press **Ctrl + Alt**, select a dimension, and drag the dimension to the Decision Tree Builder.

The dimension will appear in the **Input (Dimensions)** list with a unique color-coding.

- Add **Dimension Elements** as inputs.

In the workspace, right-click and select a Dimension table. Select Dimension Elements, press **Ctrl + Alt**, and drag the selected elements to the Decision Tree Builder.

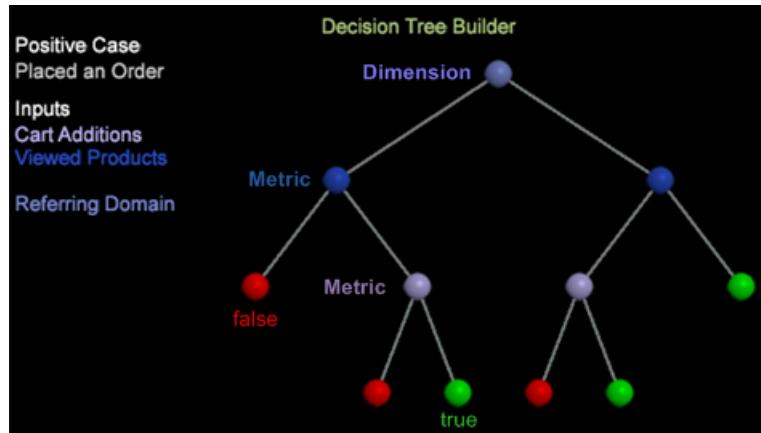
The dimension elements will appear in the **Input (Elements)** list with a unique color-coding.



Important: You can select up to a maximum of fourteen inputs to be evaluated. An error message will appear if too many inputs are added.

7. Select **Go** from the toolbar.

The decision tree will build based on the selected dimensions and metrics. Simple metrics such as Cart Additions will build quickly, while complex dimension such as Visit Duration with multiple data points will build more slowly with a percentage of the completion displayed as it converts. The tree map will then prune and open for user interaction. The dimension and metric inputs will be color-coded consistent with the node names.



The leaf node displays as green (true) or red (false) if the tree has been pruned and if there is a prediction of **True** or **False** following the pruned branches.

Attention: The training sample is pulled from the dataset for the tree builder to use. Data Workbench uses 80 percent of the sample to build the tree and the remaining 20 percent to assess the accuracy of the tree model.

8. Verify accuracy using the **Confusion Matrix**.

Click **Options > Confusion Matrix** to view the Accuracy, Recall, Precision and F-Score values. The closer to 100 percent, the better the score.

The Confusion Matrix gives four counts of accuracy of the model using a combination of values:

- Actual Positive (AP)
- Predicted Positive (PP)
- Actual Negative (AN)
- Predicted Negative (PN)



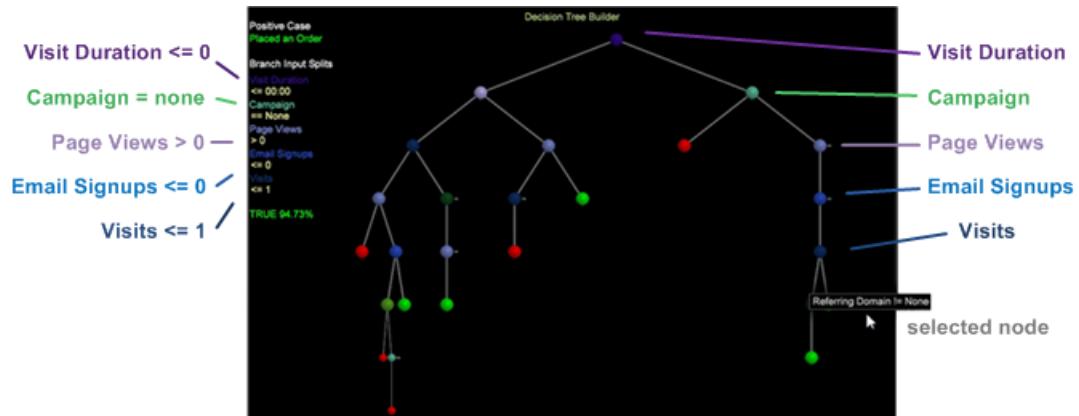
Tip: These numbers are obtained by applying the resulted scoring model of the 20 percent testing data withheld and already known as the true answer. If the score is greater than 50 percent, it is predicted as a positive case (that matches the defined filter). Then, Accuracy = $(TP + TN) / (TP + FP + TN + FN)$, Recall = $TP / (TP + FN)$, and Precision = $TP / (TP + FP)$.

9. Explore the decision tree.

After generating a decision tree, you can view the path of the prediction and identify all visitors who meet the defined criteria. The tree identifies the input split for each branch based on its position and color-coding. For example, if you select the Referring Domain node, the nodes leading to that split are listed by color-code to the left of the tree.

You can make selections of the leaf nodes to select branches (rule sets) of the decision tree.

For this example: If the visit duration is less than 1, no campaign exists, at least one page view exist, no email signups, and there was at least one visit. The projections on this meeting criteria and placing an order is **94.73** percent.



Decision Tree interaction: You can select multiple nodes on the tree using the standard **Ctrl-click** to add, or **Shift-click** to delete.

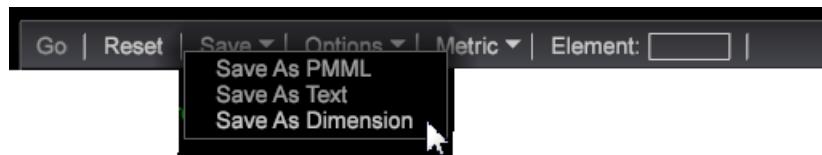
Color-coded nodes: The color of the nodes matches the color of the input dimensions and metrics as assigned by Data Workbench.

Bright green and red nodes at the leaf-level of a pruned branch predicts the node as True or False.

● Bright green	Identifies that the node equals true and that all conditions are met.
● Bright red	Identifies that the node equals false and not all conditions are met.

10. Save the Decision Tree.

You can save the Decision Tree in different formats:



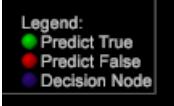
- Predictive Markup Language (**PMML**), an XML-based file format used by applications to describe and exchange decision tree models.
- **Text** displaying simple columns and rows of true or false, percentages, number of members, and input values.
- A **Dimension** with branches corresponding to predicted outcome elements.

Decision Tree Options

The Decision Tree menu includes features to set the positive use case, filters, leaf distribution options, confusion matrix, and other advanced options.

Toolbar buttons	Description
Go	Click to run the decision tree algorithm and display the visualization. This is grayed-out until there are inputs.
Reset	Clears inputs and decision tree model and resets the process.
Save	Save the Decision Tree. You can save the Decision Tree in different formats: <ul style="list-style-type: none"> • Predictive Markup Language (PMML), an XML-based file format used by applications to describe and exchange decision tree models.

Toolbar buttons	Description
	<ul style="list-style-type: none"> Text displaying simple columns and rows of true or false, percentages, number of members, and input values. A Dimension with branches corresponding to predicted outcome elements.
Options	See table below for Options menu.

Options menu	Description
Set Positive Case	Defines the current workspace selection as the model's Positive Case. Clears the case if no selection exists.
Set Population Filter	Defines the current workspace selection as the model's Population Filter and will be drawn from visitors who satisfy this condition. The default is "Everyone."
Show Complex Filter Description	Displays descriptions of the defined filters. Click to view the filtering scripts for the Positive Case and Population Filter.
Hide Nodes	Hides nodes with only a small percentage of the population. This menu command displays only when the decision tree is displayed.
Confusion Matrix	<p>Click Options > Confusion Matrix to view the Accuracy, Recall, Precision and F-Score values. The closer to 100 percent, the better the score.</p> <p>The Confusion Matrix gives four counts of accuracy of the model using a combination of values:</p> <ul style="list-style-type: none"> Actual Positive (AP) Predicted Positive (PP) Actual Negative (AN) Predicted Negative (PN) <p> Tip: These numbers are obtained by applying the resulted scoring model of the 20 percent testing data withheld and already known as the true answer. If the score is greater than 50 percent, it is predicted as a positive case (that matches the defined filter). Then, Accuracy = $(TP + TN) / (TP + FP + TN + FN)$, Recall = $TP / (TP + FN)$, and Precision = $TP / (TP + FP)$.</p>
Display Legend	<p>Allows you to toggle a legend key on and off in the Decision Tree.</p>  <p>This menu command displays only when the decision tree is displayed.</p>
Advanced	Click to open Advanced menu for in-depth use of Decision Tree. See table below for menu options.

Advanced menu	Description
Training Set Size	Controls the size of the training set used for the model building. Larger sets take longer to train, smaller sets take less time.
Input Normalization	Allows the user to specify whether to use the Min-Max or the Z Score technique to normalize inputs into the model.
SMOTE Over-Sampling Factor	When the Positive Case does not occur very often (less than 10 percent) in the training sample, SMOTE is used to provide additional samples. This option allows the user to indicate how many more samples to create using SMOTE.
Leaf Class Distribution Threshold	Allows you to set the threshold assumed for a leaf during the tree building process. By default, all members of a node must be identical for it to be a leaf (prior to pruning stage).

Finders

Use Finder panels in Data Workbench to select metrics, dimensions, and filters. These panels provide search support, sorting options, and drag and drop capabilities.

A Finder panel can be opened in the left sidebar or within a workspace.



Dimensions Finder	Metrics Finder	Filters Finder
A list of all dimensions in your query model.	A list of all metrics in your query model.	A list of all filters created for your organization.
A screenshot of the Dimensions Finder panel. It shows a list of dimension names: Search:page, Clickthrough Landing Page Name, Entry Page, Entry Page eVar24, Exit Page, Internal Search Page, Next Page, Organic Search Landing Page URL, Page, Page Event Link Name, Page Event Link URL, Page Event Types, Page Name, and Page Type.	A screenshot of the Metrics Finder panel. It shows a list of metric names: Search:vis, Visitor Visits, by Visit, per Visit, of Visitors, of Visits, v Visitors, sum Visits, t Duration, t Orders, t Revenue, t Units, tors, and tors by Visit.	A screenshot of the Filters Finder panel. It shows a list of filter names: Search:gain, PC filter, and Single Access Filter.

To open a Finder:

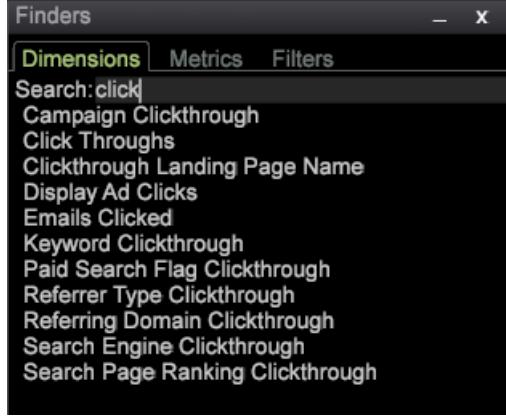
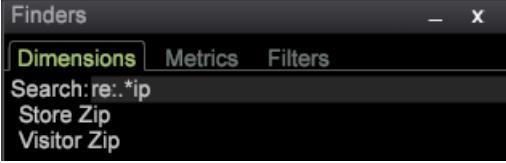
- Right-click in a workspace and select **Tools > Finder**.

The Finder pane with tabs for Metrics, Dimensions, and Filters will open in the workspace.

- Right-click in the left sidebar and select **Add > Finder**.

The Finder pane will open in the left panel.

The **Finder** includes the following features:

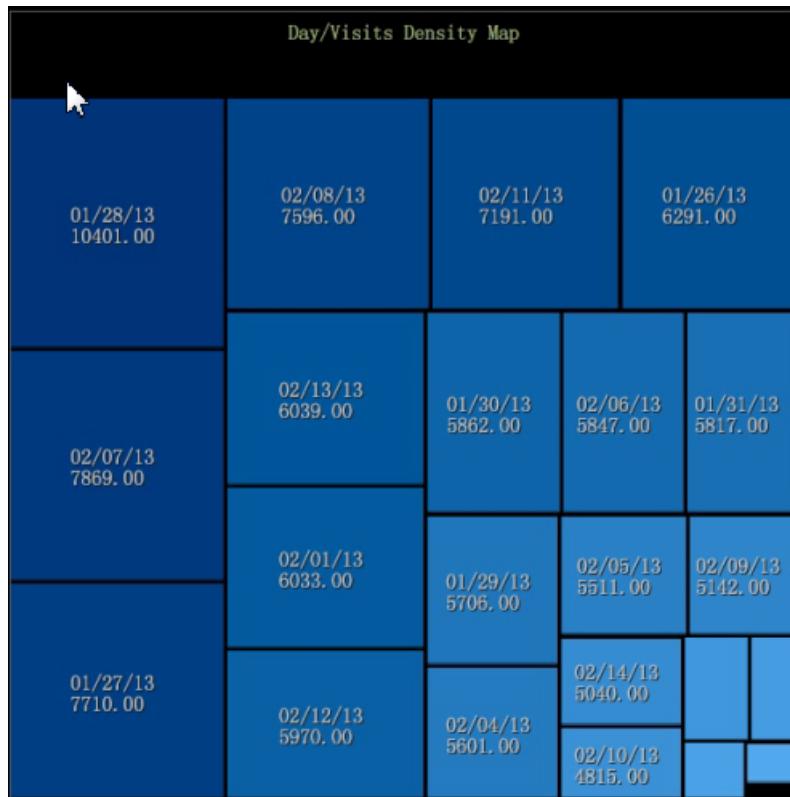
Finder Features	Details
Drag and drop	<p>You can drag and drop dimensions or metrics from the panel to a visualization in the workspace to change the dimension or add new metrics.</p> <ol style="list-style-type: none"> 1. Hold down the <Ctrl> and <Alt> keys and select the dimension or metric from the Finder panel. 2. Drag a new dimension from the pane and drop it to the visualization to change or add dimensions. 3. To add metrics, drag a new metric from the pane and drop it on the metric header of the selected visualization. <p>This will work for all relevant visualizations, including tables, visitor cluster, correlation matrix, scatter plots, and the 2-D bar graph (depending on the axis).</p>
Search	<p>A Search box in the Finder panels lets you filter names for Dimensions, Metrics, and Filters.</p> <ul style="list-style-type: none"> • Pattern matching (simple glob search). Start typing the name of a required dimension, metric, or filter entity in the Search field and only matching strings contained anywhere in the name will be filtered and displayed in the Finders pane. <p>For example, enter:</p> <p>Search:click</p> <p>You could get the following results in the Dimensions Finder:</p>  <p>Standard pattern matching lets you use the wildcard characters, such as . (dot), "?" , and "*" (star).</p> <ul style="list-style-type: none"> • Regular expressions. More complex regular expressions are also supported for added search capability. Add the prefix "re:" before your search term (no spaces) to interpret as a regular expression. <p>For example, enter:</p> <p>Search:re.*ip</p> <p>You could get the following results in the Dimensions Finder:</p> 

Finder Features	Details
	For in-depth search information, see regular expressions .
Dimension Type	In the Dimension tab, you can right-click on the tab heading to sort by the type of dimension.
	 <ul style="list-style-type: none"> • Attributes—Dimensions built based on characteristics of the visitor, products, geography, time, video, and other attributes. • Clusters—Dimensions built within the cluster builder. • Scores—Dimensions built within the propensity scoring.
Label	In each tab, you can right-click and select Label to rename the Finder pane.
	 <p>The default Dimensions, Metrics, and Filters labels can be changed to a tab name that meets your organization's conventions.</p>
Add Item	In each tab, you can right-click and select Add Item to open a table and manually add Dimensions, Metrics and Filters.
Finders bar	Right-click in the Finders bar in the left sidebar to open a menu for additional features.
	
Close	Right-click in the Finders bar and select Close to close a Finders pane.
Save	Save the list locally by right-clicking in the header bar and selecting the Save option.
Export	You can export a list of selected dimensions, metrics, or filters from the Finder panel by right-clicking in the Finders bar and selecting Export from the menu. Add a name and export to Microsoft Excel.
Copy	Copy a list of Dimensions, Metrics, or Filters. You can copy as a file or as a graphic in Dark Background, Light Background, or Monochrome.
Minimize	Minimize the Finder pane. Only the Finders bar will appear.
Borderless	Displays a pane with no border lines for Finders in the workspace (but not in the left sidebar).

Density Map

The density map visualization displays elements as shaded rectangles within a square map.

The sizes of the rectangles are dependent on element values, where larger values are represented by rectangles of larger area. Similar to a pie chart, this visualization allows you to quickly see which elements constitute the greatest percentage of the selected dimension.



To create a density map:

1. Open a new workspace.

After opening up a new workspace, you may need to click **Add > Temporarily Unlock**.

2. Click **Visualization > Density Map**.
3. Select a **Dimension** from the menu.

For example, select **Time > Days**.

In contrast, selecting **Time > Hours** would give you more elements with smaller values displaying as smaller rectangles.



Note: You will want to pick a dimension with multiple elements per your needs. The current limit is 200 of the largest elements for each dimension.

4. You can change dimension views by opening **Visualization > Table** and selecting across elements from the table to display in the map.



The map will respond to selections from the table.

5. Hovering over small elements will display their name and value in text that appears near the mouse cursor.
6. Mask elements by right-clicking and selecting **Mask**, then choose an option.



To display all masked nodes, select **Unhide All**.

7. Spotlight elements by right-clicking and selecting **Spotlight**, then choose an option. Spotlighting lets you highlight and dim elements in a range.
8. Add a color legend to the workspace. You can identify values in the map using the color legend.

You can add a color legend to the workspace and the nodes will change color based on the additional dimension of data.

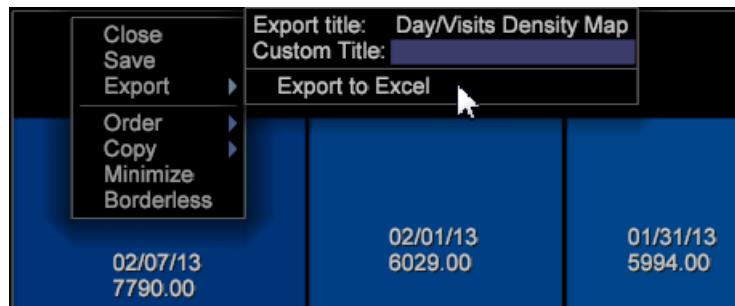
9. Change the dimension or metric by right-clicking the map title and selecting from the menu.



10. Add callouts by right-clicking a cell and selecting **Add Callout**. You can select from different types or visualizations from the menu.



11. As in all visualizations, you can right-click above the title bar for basic commands to Close, Save, Export to Microsoft Excel, Order, Copy, Minimize, and Borderless to display a visualization without a border.



12. The Density Map lets you select and deselect multiple elements similar to other visualizations:

- Left-click to select an element.
- Ctrl + click to select multiple elements.
- Shift + click to deselect an element.

- Right-click within selected elements to open a menu. Then choose **Deselect** or **Deselect All** to clear selected elements.

Additional Options

Right-click the Density Map to open a menu with these options:

Option	Description
Add Callout	<p>Add a text or graphic as a callout in the visualization to further identify or describe an element.</p> <p>You can also select a blank Metric Legend, Table, Line Graph, or Scatter Plot based on the selected element in the Density Map. You can then add metrics and dimensions to these blank visualizations as needed.</p>
Mask	<p>Masking options let you hide selected elements. Right-click to display Mask options.</p> <p>Hide This Element—Choose this option to mask a single element that you have selected.</p> <p>Hide Selected—Choose this option to mask multiple elements that you have selected.</p> <p>Show Top— Choose this option to display only the top 100, 50, 25, or 10 top elements based on the values in the Density Map.</p> <p>Show Bottom—Choose this option to display only the bottom 100, 50, 25, or 10 top elements based on the values in the Density Map.</p>
Spotlight	<p>Spotlighting lets you highlight and dim elements in a range. Right-click to open a menu of options.</p> <p>Show Top— Choose this option to highlight only the top 100, 50, 25, or 10 top elements based on values in the Density Map.</p> <p>Show Bottom—Choose this option to highlight only the bottom 100, 50, 25, or 10 top elements based on values in the Density Map.</p>
Deselect Deselect All	Select these commands to deselect the current element, if selected, or deselect all elements that are selected.

Attribution Profile

Using the new rules-based Attribution profile in Data Workbench, you can quickly analyze attribution events and assign responsibility leading up to a successful conversion defined by you. The Attribution profile comes complete with the information necessary for your data Architect to set up and extend its features, and includes pre-built workspaces for your Analyst to jump right in and start analyzing.

The Attribution profile allows you to gain a new perspective on the relationships between your marketing efforts and a successful customer lead generation or sales conversion. The Attribution profile helps you qualify interactions that should receive allocation of credit for realized revenue or participation downstream in the customer journey. It helps identify the impact of your marketing efforts and costs by allowing you to quickly analyze attribution events, and then assign responsibility for first or last touches or other events leading to a successful sale.

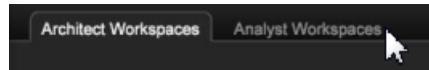


Important: The Attribution profile is configured for immediate use by users who have implemented the Adobe SC profile that uses the Analytics (SC/Insight) data feed. By default, the Marketing and Conversion events are employed as the default types of interactions evaluated in the provided rules-based models.

See [Deploying the Attribution Profile](#) on page 101 and [Attribution Models](#) on page 103 for additional information.

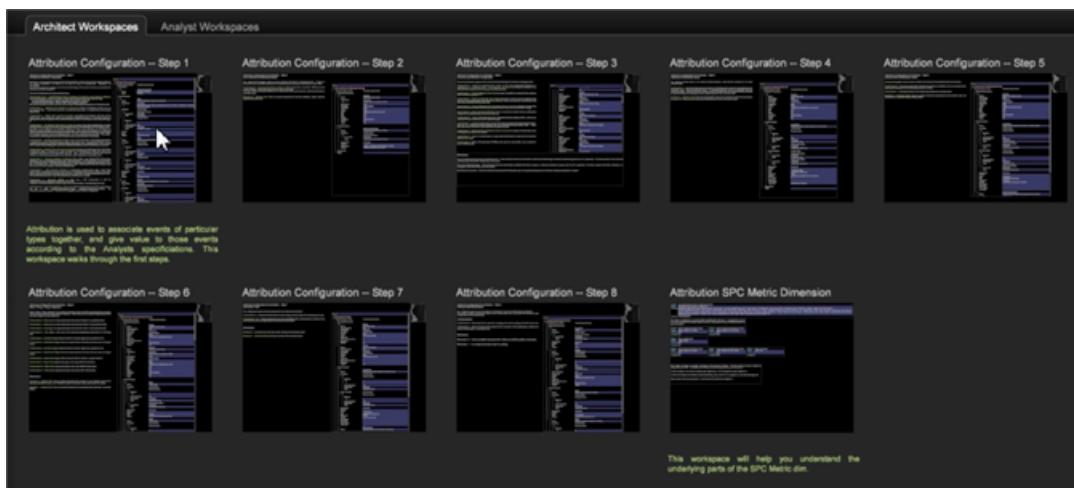
Architecture and Analyst Workspaces

Within the Attribution profile, you have Architect and Analyst workspaces defined on separate tabs in the workbench.



Architecture Workspaces

Within the **Attribution** tab, click the **Architect Workspaces** tab to open workspaces specifically designed to set up your configuration files for basic attribution modeling.



The Architecture tab includes workspaces to step through each of the configuration files in the profile dataset folder. For example, **Attribution Configuration - Step 1** lets you identify the Attribution values within the Transformation section of the `profile.cfg` file.

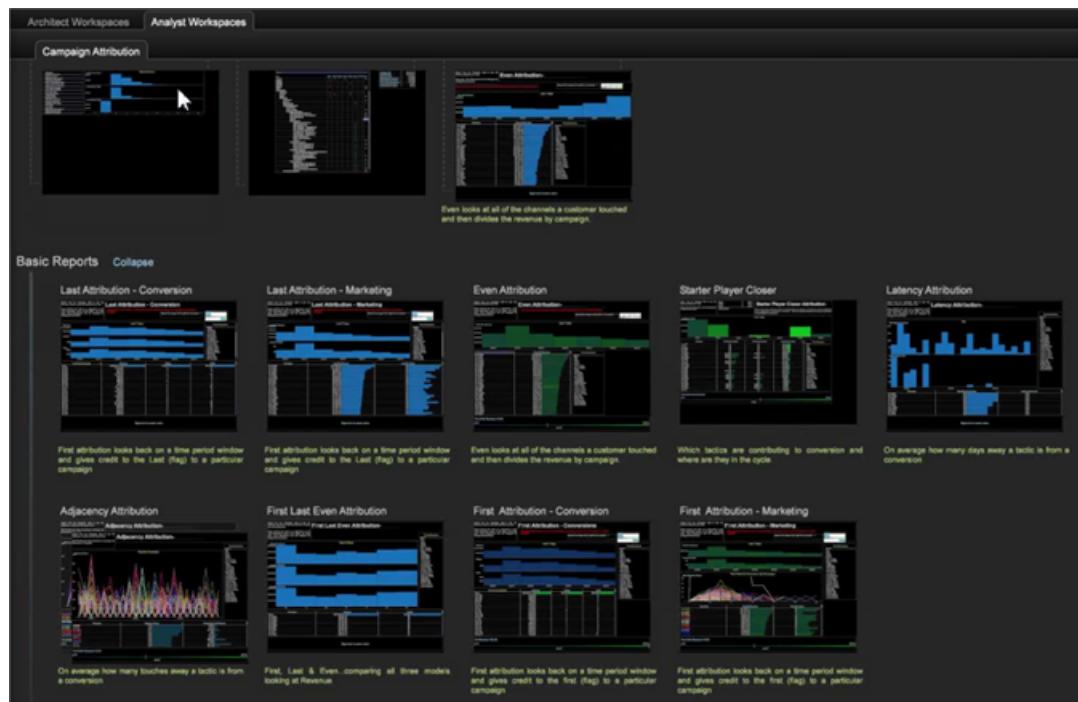
A screenshot of the 'Attribution Configuration for Architects - Step 1' workspace. The main area contains text and tables describing the configuration of Attribution parameters. On the right, there is a search sidebar titled 'Search:' with a list of configuration file names such as 'profile.cfg', 'Client.cfg', 'Cluster.cfg', etc. The search bar has the text 'Attribution' entered.

Analyst Workspaces

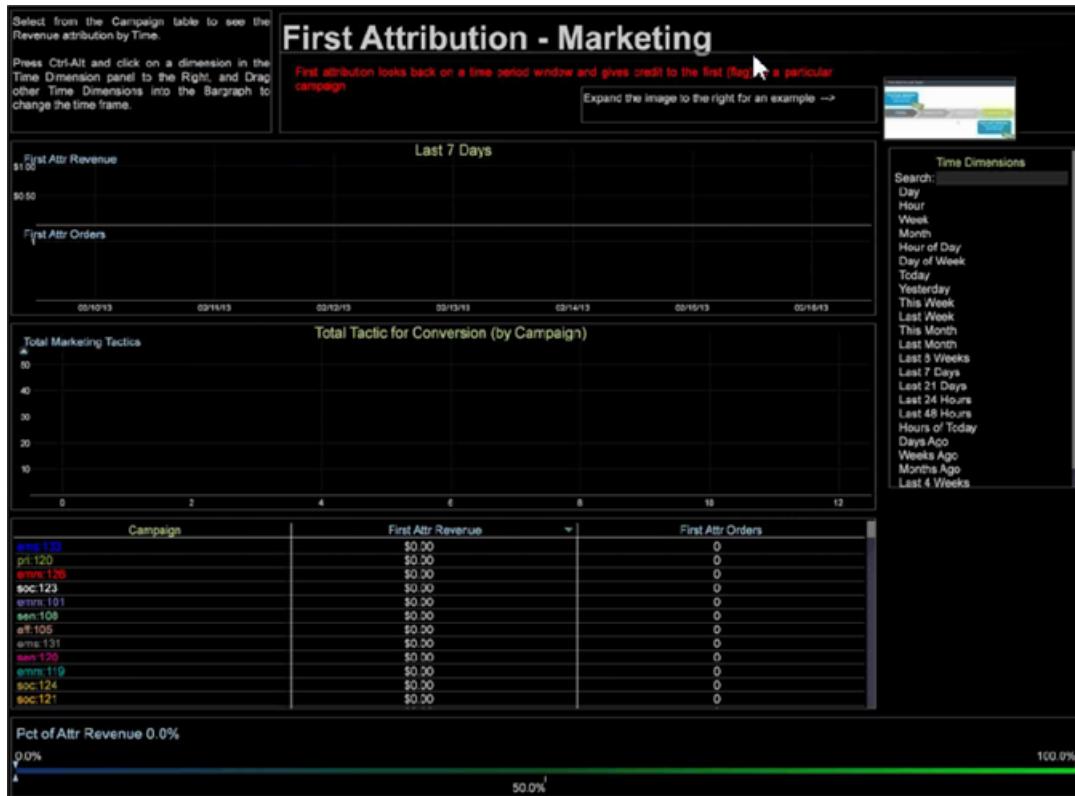
Click the **Analyst Workspaces** tab to open workspaces prebuilt analysis utilizing the dimensions and metrics provided with the Attribution profile.

These workspaces are organized into four categories:

1. **Basic Reports** expose a single model within a workspace.
2. **Comparative Reports** extended the analyses by presenting multiple models within a single view.
3. **Investigation Reports** expands the reporting templates to present the attribution models in different formats. This section also introduces and exposes the position-based weighting ratios.
4. **Pathing Reports** provide visibility into the customer's marketing journey with multiple pathing visualizations to fully explore and express the process flows and interaction paths



The Analyst tab includes workspaces pre-configured with reports. For example, **First Attribution** lets you select from the **Campaign** table to see the **Revenue** attribution based on **Time**.



Deploying the Attribution Profile

The Attribution profile is an inherited, ready-to-drop-in profile. In combination with the Adobe SC profile and Analytics (SC/Insight) data feed, the profile can be deployed to quickly expose new attribution models across digital channels.

After saving the Attribution profile to the primary server, there are two additional steps necessary to integrate it into the current profile within the `Profile` directory: (1) Set up the `Profile.cfg` file, and (2) Declare the Required Fields.

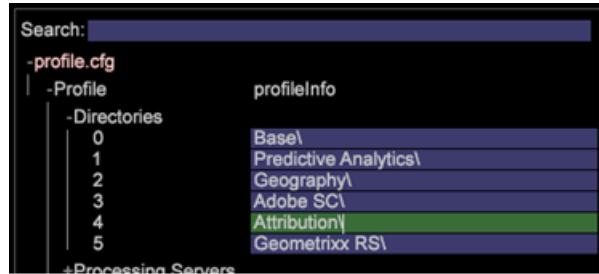
Setting up `Profile.cfg` file

Like all profiles, the Attribution profile needs to be added to the `profile.cfg` file. Because the Attribution profile depends on the Adobe SC profile, the Adobe SC profile needs to be listed first in the configuration file before the Attribution profile.



Note: These steps will require a re-transformation of the dataset.

1. Open the `profile.cfg` file in your custom profile folder. (Open in `server\Profiles\(\custom profile name)\profile.cfg`.)
2. If the Attribution profile is not listed in the configuration file, add it to the list.



3. Make sure the **Attribution** string is listed below the **Adobe SC** profile string.
4. Save the updated **profile.cfg** file and then save it to server from the Profile Manager.

Declaring the Required Fields

The Attribution profile takes predefined fields and with a series of transformations exposes those fields in new and useful ways through extended dimensions. To provide the most immediate value the Attribution profile depends on fields available with the Adobe SC profile.

Default Variables	Field Name and Decoder Position (Adobe SC)
Campaign	x-campaign, #199
Marketing Channels	x-va_closer_detail, #162 x-va_instance_event, #163
Order event	x-order, #206 x-purchaseid, #200
Revenue	x-revenue, #205
Units	x-units, #204

1. Verify that these fields are declared in the Decoder Group used to define the Adobe Analytics data source. The default decoder group is provided under **Dataset\Log Processing\Decoding Instructions.cfg**.
2. Verify that these fields are declared in the **Fields** section of the **SC_Fields.cfg** file. This file can be located under **Dataset\Log Processing\SC_Fields.cfg**.

Attribution Additions and Troubleshooting

The Attribution profile added a configuration file, **0a_Marketing_Channels.cfg**, which copies the value of the **x-va_closer_detail** into a new field called **x-marketing-channel**, when the **x-va_instance_event** field matches "1". Both **x-va_closer_detail** and **x-va_instance_event** are decoded by default, and passed from decoding in the installed packages available when you update to version 6.2.

The **x-marketing-channel** field is then used in the Simple dimension called Marketing Channel.



Important: If you have altered your profiles by removing previously unused fields that are now being used, you will want to verify that the **x-va_closer_detail** and **x-va_instance_event** fields are being decoded and passed through for use.

If fields are missing, then you will get a message in your detailed status:

`x-va_closer_detail` is not available

or

`x-va_instance_event` is not available

Attribution Models

Seven different attribution models are provided to use with the Attribution profile to help illustrate and quantify the customer-marketing journey.

First and Last model



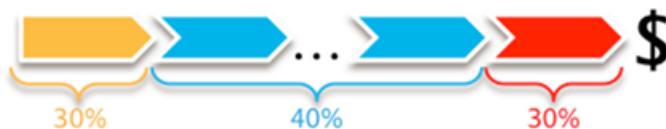
The first attribution models are the traditional **First** and **Last** touches. Understanding the first step into converting a sale or the last touch before checking out and buying a product provides an understanding of campaign types—for awareness, the First touch, and for call-to-action, the Last touch.

Even model



The view into the marketing engagement is expanded with the introduction of **Even**. Every qualifying marketing interaction is given an equal share of the subsequent order and revenue generated from the conversion.

Starter, Player, and Closer model



And moving beyond an equal allocation the **Starter, Player, & Closer** model provides a position based weighting scheme with adjustable weighting. The weights can be explored while performing analysis.

Consider this example: The starter (or first) gets 30 percent of the converted revenue, closer (or last) gets 30 percent, and all the players (those in-between) share an equal portion of the remaining 40 percent.



Note: It's fairly common to assign the allocation back to marketing interactions based on either revenue or occurrence of the order. These next two models attribute different characteristics to the marketing interactions to expose other dynamics of success.

Adjacency model



The Adjacency model provides visibility into the marketing position away from the conversion, answering the question: Is the marketing channel typically the 1st (closest), 5th, or 10th marketing channel interaction away from success?

Latency model

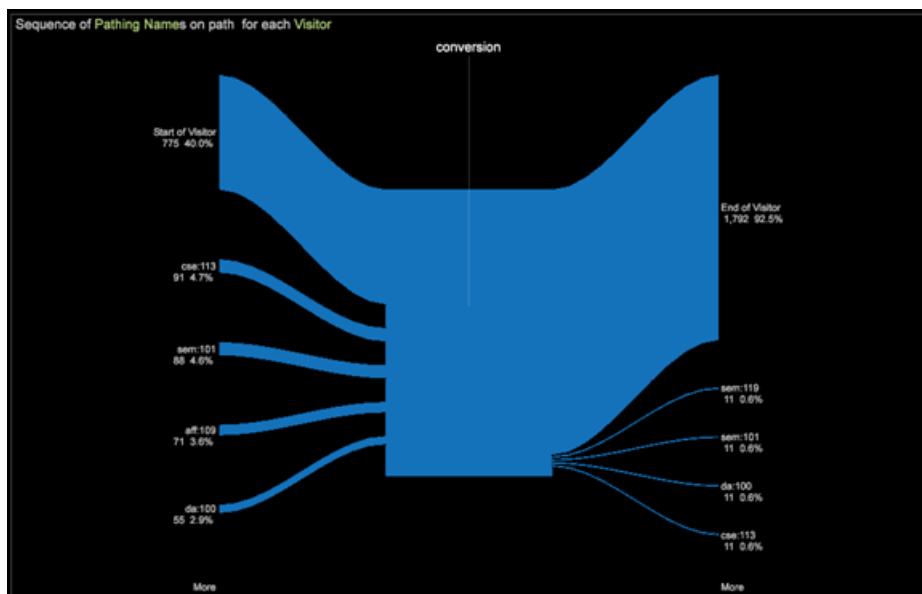


The **Latency** model helps describe the lag time between the interaction with the marketing channel and the occurrence of success. This is useful in presenting awareness versus call-to-action type campaigns and knowing which levers the business can pull to get a more timely response from a target audience. (A value of zero (0) days means the marketing interaction occurred the same day as the success.)

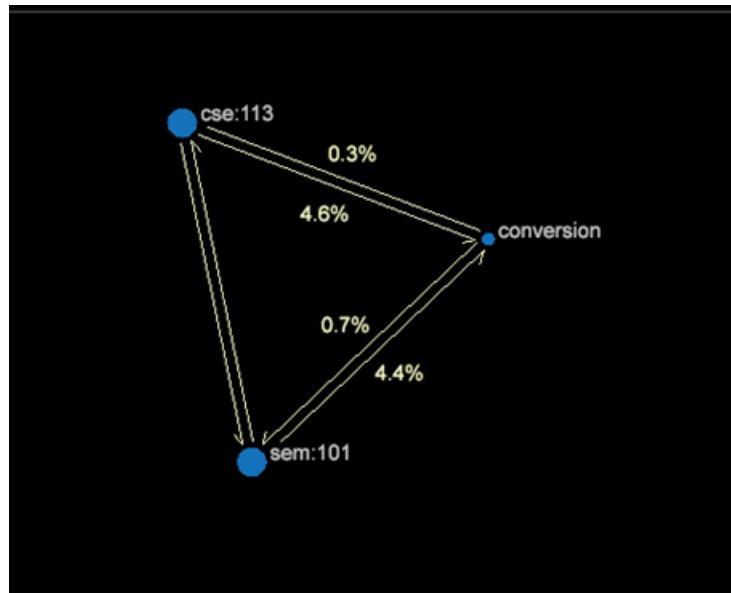
Pathing model

The Pathing model provides different approaches for exploring the customer's engagement with marketing and his or her successful conversion, establishing the relationships between marketing interactions within the customer journey. Explore the process maps to understand high-level flows between supporting marketing channels and success. Evaluate direct sequential marketing interactions with the path browser.

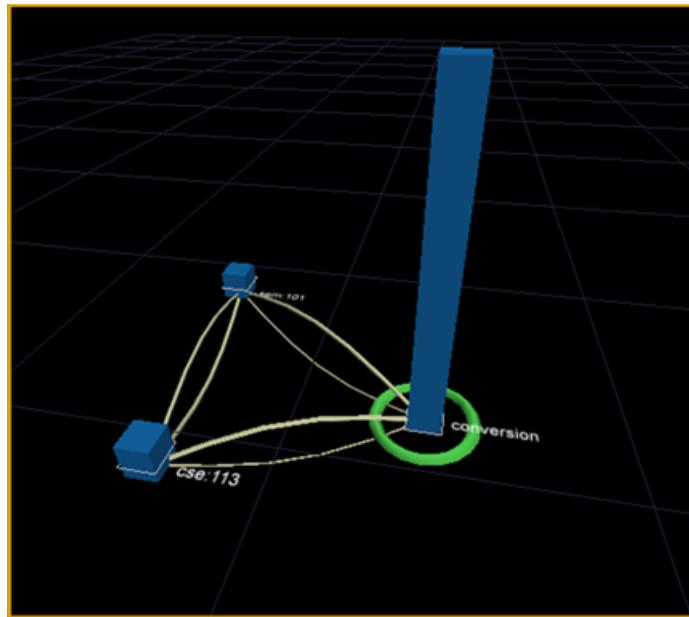
Path Browser visualization



2D Process Map visualization



3D Process Map visualization



3D Scatter Plots

Navigation title:

A 3D Scatter Plot graphs the elements of a data dimension (such as Days or Referral Site) on a three-dimensional grid where the x, y, and z axes represent various metrics.

Like the [Scatter Plot 2D](#), this visualization is useful when trying to understand the relationship between large numbers of disparate items employing different metrics.

To employ the 3D Scatter Plot visualization:

1. Open a new workspace.

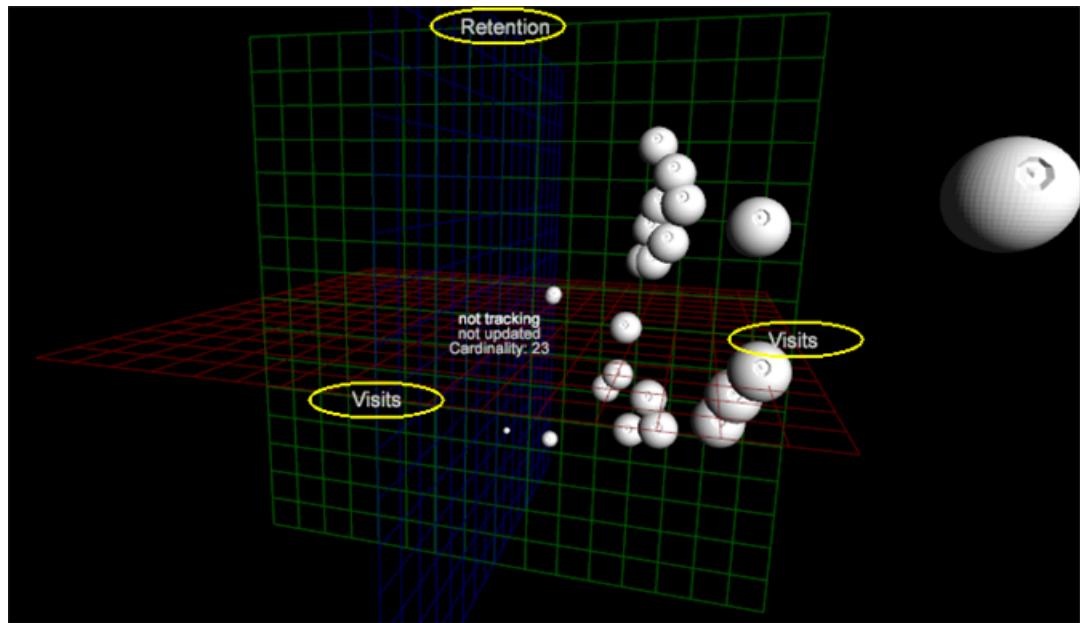
After opening a new workspace, you may need to click **Add > Temporarily Unlock**.

2. Right-click and select **Visualization > 3D Scatter Plot**.

A menu listing **Dimensions** will open.

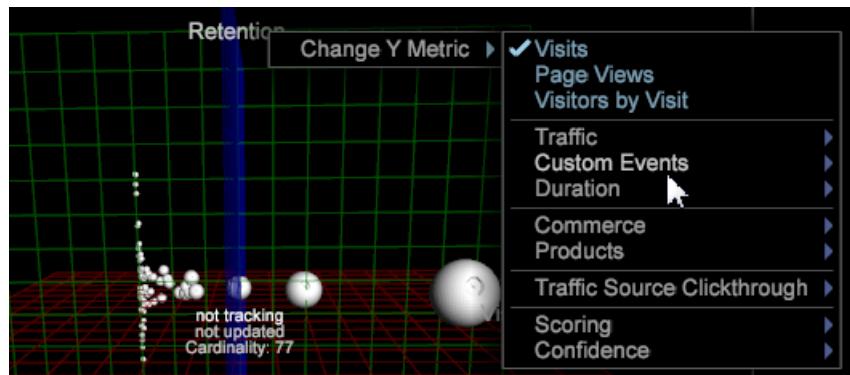
3. Select a dimension for the query.

The 3D Scatter Plot will open the default metrics for that dimension.



Selecting the **Days** menu displays the following 3D Scatter Plot with these default metrics on the following axes: **x=Visits**, **y=Retention**, and **z=Visits**.

4. Change metrics. Right-click on the metric label in the x, y, or z axis and select **Change Metric**. Then select a different metric for the selected axis.

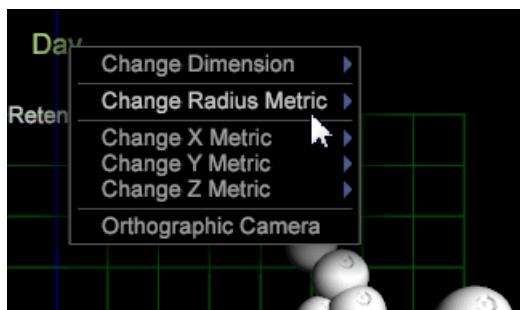


Important:

- Drag a metric to one of the three axis labels and drop it to change the selected axis to the dropped metric.

- Drag a metric anywhere else on the visualization and drop it to change the radius metric for that axis.
 - Drag a dimension to anywhere on the visualization and drop it to change the dimension for the visualization.
5. Change the Radius Metric. Right-click the title at the top of the page (titled after the selected dimension) and select **Change Radius Metric**.

The radius metric defines the size of the plotted point based on the metric selection. The relative position of points does not change in the scatter plot, but the plotted point sizes within the visualization increase based on the metric value.



6. Employ the **Orthographic Camera**. This option lets you identify the plotted points in relation to their true perspective based on the radius metric to avoid three-dimensional distortion.

When the 3D Scatter Plot first appears, it displays in a three-dimensional rotating projection, which causes some distortion for points plotted nearer to the perspective, or virtual "camera." (The plots nearer to the camera show up much larger than the points rotating further away from the camera.)

To avoid this perspective distortion, you can select the **Orthographic Camera** option by right-clicking on the title and selecting from the menu. This allows you to represent the three-dimensional objects in two-dimensions. This renders the plotted points as flat and displays the points as relative to the radius metric, lessening the 3-dimensional offsets.

7. Select points from the scatter plot.

- **To remove a point or group of points:** Click the point.
- **To add another point or group of points to your selection:** **Ctrl + click** a point or **Ctrl + drag** across multiple points.
- **To remove a point or group of points from your selection:** **Shift + click** a point or **Shift + drag** across several points.

Analytics Reports

Analytic reports are provided for users of the Adobe SC profile. These data workbench reports—Page Views, Traffic, Unique Visitors, Referring Domain and other significant report types—are standard reports in Adobe Analytics.

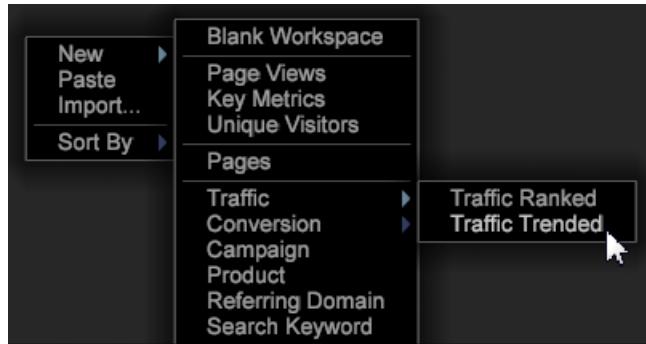
For Adobe Analytics' users, these templates allow users employing the Adobe SC profile (which uses the SC/Insight data feed) to view reports generated from Data Workbench using templates similar to the Adobe *Marketing Reports & Analytics* capability. Users can access these pre-configured reports using data workbench templates.



Note: This menu will only appear for those users who have implemented the Adobe SC profile.

To Open Reports

On the worktop, right-click and select **New** to view and open report templates.



The following are a list of Analytic reports:

Reports	Description
Page Views	A trended report that displays the number of times your website pages were viewed for the selected time period (hour, day, week, month, quarter, or year). This report allows you to track page views for each page on your site, as well as an aggregate of page views for your entire site.
Key Metrics	The Key Metrics Report lets you compare metrics to see whether they trend together. For example, as your page views increase, does your visitor count increase?
Unique Visitors	Shows you the number of unique visitors who accessed your site. Each visitor is counted once, regardless of how many times the person visits your website.
Pages	Ranks the pages on your site based on the pages that receive the most traffic. If your business question deals with quantitative data for pages, you can use this report to answer that question by adding the right metrics.
Traffic	Traffic reports give you in-depth insight into how visitors interact with your website.
Conversion	Provide comprehensive, accurate, and detailed analysis of customer activity. Metrics such as campaign management, sales cycle, customer fallout, and customer conversion let you measure e-commerce transactions, sources of sales, advertising effectiveness, customer loyalty, and more.
Campaign	Displays information about the effectiveness of your advertising efforts. You can see which types advertising efforts give you the most traffic and which of your employees is responsible for driving those efforts.
Product	Identifies how individual products and groups of products (categories) contribute to your various conversion metrics, such as Revenue or Checkouts.
Referring Domain	Shows the domains that referred the customers that most impacted your site's success metrics. Referrers fall into two main categories: Domains and URLs. Domains refer to the domain name, and appear as the base domain without the query string or subdirectories attached. URLs include the base domain name, as well as any query strings or subdirectories.
Search Keyword	Reports that display a breakdown of search keywords.

Data Workbench 6.11 Update

Data Workbench 6.1.1 updates OpenSSL in your system necessary to fix the "Heartbleed" vulnerability found in previous versions of OpenSSL. Certain Data Workbench versions (5.5.x, 6.0.x and 6.1.x) used a version of OpenSSL deemed vulnerable.

Managed users of Data Workbench have been updated to version 6.1.1, but if you have an on-premise implementation you will need to update.



Important: Because Data Workbench servers require authentication, it is important to understand that external attackers cannot exploit this vulnerability because they can't get past the initial handshake.

See [Data Workbench 6.0 to 6.1 Upgrade](#) on page 140 for detailed information about upgrading from previous versions to Data Workbench 6.1.1.

Updates for April MR 2014

Navigation title:April MR Release Notes

Adobe Analytics will update and enhance its data feed powering Data Workbench to take advantage of the Master Marketing Profile within the Adobe Marketing Cloud. All Data Workbench users are required to be prepared to transition to this new data feed format by April 21, 2014.

The Master Marketing Profile provides a complete and fully actionable view of customers for use across the Adobe Marketing Cloud and third-party services. To start establishing the foundation for new features within Analytics, the new marketing cloud visitor identifier will be added to the data feed in this upcoming April MR release along with other enhancements and improvements. This new service is available to any customer using the Adobe Marketing Cloud.

Additional enhancements and bug fixes provided with this new data feed are listed below in this topic, as well as instructions for preparing your Data Workbench configuration.

Enhancements and Bug Fixes for the New Data Feed

- Addition of the Marketing Cloud Visitor ID.
- Addition of the Target Action field to support Target reporting and analysis within Analytics.
- Improved alignment with Report & Analytics by incorporating the `bot_id` value on the product row types.
- Reduction in size of merchandising eVars without the loss of integrity on the data values.
- Resolved an issue where some times hashed values would show up in merchandising eVars.
- Internal improvements within the process to continue to deliver better stability and performance.

Updating the Data Workbench Configuration

Changes to Data Workbench will ensure uninterrupted integration of the data feed after the transition. During your next scheduled reprocess update, the current Decoder Group within the Log Processing phase of the configuration. The team responsible for managing the Data Workbench architecture will be able to perform additional tasks to help expose and take advantage of the new marketing cloud visitor identifier.

Tasks to update Data Workbench

1. Locate the current Decoder Group. Typically, located in the `Dataset\Log Processing\Decoding Instructions.cfg` found within a Profile Manager.
2. Within the configuration file, right-click on the current decoder format and select Copy.
3. Right-click and paste a new entry below the current decoder format.
4. Expand the new decoder format and add two new field positions to the end of the list. These positions can be left empty if you aren't ready to use the new fields in the dataset, or define them as follows.
 - `x-mcvisid`
 - `x-tnt-action`

Decoding Instructions

```
Log Processing Include = LogProcessingInclude:  
Decoder Groups = vector: 1 items  
  0 = TextFileDecoderGroup:  
    Decoders = vector: 2 items
```

```
0 = DelimitedDecoder:  
    Delimiter = string: \t  
    Fields = vector: 389 items  
        0 = string: x-insight-row_type  
        1 = string: x-exclude_hit  
        10 = string: x-visit_num  
        100 = string: x-prop33  
        101 = string: x-prop34  
        102 = string: x-prop35  
        103 = string: x-prop36  
        104 = string: x-prop37  
        105 = string: x-prop38  
        106 = string: x-prop39  
        107 = string: x-prop40  
        108 = string: x-prop41  
        109 = string: x-prop42  
        11 = string: x-visit_page_num  
        110 = string: x-prop43  
        111 = string: x-prop44  
        112 = string: x-prop45  
        113 = string: x-prop46  
        114 = string: x-prop47  
        115 = string: x-prop48  
        116 = string: x-prop49  
        117 = string: x-prop50  
        118 = string: x-prop51  
        119 = string: x-prop52  
        12 = string: x-hitid_high  
        120 = string: x-prop53  
        121 = string: x-prop54  
        122 = string: x-prop55  
        123 = string: x-prop56  
        124 = string: x-prop57  
        125 = string: x-prop58  
        126 = string: x-prop59  
        127 = string: x-prop60  
        128 = string: x-prop61  
        129 = string: x-prop62  
        13 = string: x-hitid_low  
        130 = string: x-prop63  
        131 = string: x-prop64  
        132 = string: x-prop65  
        133 = string: x-prop66  
        134 = string: x-prop67  
        135 = string: x-prop68  
        136 = string: x-prop69  
        137 = string: x-prop70  
        138 = string: x-prop71  
        139 = string: x-prop72  
        14 = string: x-accept_language  
        140 = string: x-prop73  
        141 = string: x-prop74  
        142 = string: x-prop75  
        143 = string: cs(referrer)  
        144 = string: x-ref_domain  
        145 = string: x-ref_type  
        146 = string: x-resolution  
        147 = string: x-s_resolution  
        148 = string: x-search_engine  
        149 = string: x-search_page_num  
        15 = string: x-bot_type
```

```
150 = string: x-state
151 = string: x-transactionid
152 = string: x-truncated_hit
153 = string: x-ua_color
154 = string: x-ua_os
155 = string: x-ua_pixels
156 = string: x-uniques_exceeded
157 = string: cs(user-agent)
158 = string: x-user_server
159 = string: x-va_finder_id
16 = string: x-bot_id
160 = string: x-va_finder_detail
161 = string: x-va_closer_id
162 = string: x-va_closer_detail
163 = string: x-va_instance_event
164 = string: x-va_new_engagement
165 = string: x-zip
166 = string: x-last_hit_time_gmt
167 = string: x-first_hit_time_gmt
168 = string: x-visit_start_time_gmt
169 = string: x-last_purchase_time_gmt
17 = string: x-browser
170 = string: x-last_purchase_num
171 = string: x-first_hit_page_url
172 = string: x-first_hit_pagename
173 = string: x-visit_start_page_url
174 = string: x-visit_start_pagename
175 = string: x-first_hit_referrer
176 = string: x-visit_referrer
177 = string: x-visit_search_engine
178 = string: x-visit_keywords
179 = string: x-daily_visitor
18 = string: x-browser_height
180 = string: x-hourly_visitor
181 = string: x-monthly_visitor
182 = string: x-yearly_visitor
183 = string: x-weekly_visitor
184 = string: x-quarterly_visitor
185 = string: x-preloaded
186 = string: x-tnt
187 = string: x-survey
188 = string: x-mvvar1
189 = string: x-mvvar2
19 = string: x-browser_width
190 = string: x-mvvar3
191 = string: x-media
192 = string: x-page_event_media
193 = string: x-page_event_var3
194 = string: x-tnt_instances
195 = string: x-survey_instances
196 = string: x-mvvar1_instances
197 = string: x-mvvar2_instances
198 = string: x-mvvar3_instances
199 = string: x-campaign
2 = string: x-userid
20 = string: x-channel
200 = string: x-purchaseid
201 = string: x-product-num
202 = string: x-category
203 = string: x-product
204 = string: x-units
```

```
205 = string: x-revenue
206 = string: x-order
207 = string: x-cart_open
208 = string: x-cart_view
209 = string: x-checkout
21 = string: x-click_action
210 = string: x-cart_add
211 = string: x-cart_remove
212 = string: x-product_view
213 = string: x-evar1
214 = string: x-evar2
215 = string: x-evar3
216 = string: x-evar4
217 = string: x-evar5
218 = string: x-evar6
219 = string: x-evar7
22 = string: x-click_action_type
220 = string: x-evar8
221 = string: x-evar9
222 = string: x-evar10
223 = string: x-evar11
224 = string: x-evar12
225 = string: x-evar13
226 = string: x-evar14
227 = string: x-evar15
228 = string: x-evar16
229 = string: x-evar17
23 = string: x-click_context
230 = string: x-evar18
231 = string: x-evar19
232 = string: x-evar20
233 = string: x-evar21
234 = string: x-evar22
235 = string: x-evar23
236 = string: x-evar24
237 = string: x-evar25
238 = string: x-evar26
239 = string: x-evar27
24 = string: x-click_context_type
240 = string: x-evar28
241 = string: x-evar29
242 = string: x-evar30
243 = string: x-evar31
244 = string: x-evar32
245 = string: x-evar33
246 = string: x-evar34
247 = string: x-evar35
248 = string: x-evar36
249 = string: x-evar37
25 = string: x-click_source_id
250 = string: x-evar38
251 = string: x-evar39
252 = string: x-evar40
253 = string: x-evar41
254 = string: x-evar42
255 = string: x-evar43
256 = string: x-evar44
257 = string: x-evar45
258 = string: x-evar46
259 = string: x-evar47
26 = string: x-click_tag
```

```
260 = string: x-evar48
261 = string: x-evar49
262 = string: x-evar50
263 = string: x-evar51
264 = string: x-evar52
265 = string: x-evar53
266 = string: x-evar54
267 = string: x-evar55
268 = string: x-evar56
269 = string: x-evar57
27 = string: x-code_ver
270 = string: x-evar58
271 = string: x-evar59
272 = string: x-evar60
273 = string: x-evar61
274 = string: x-evar62
275 = string: x-evar63
276 = string: x-evar64
277 = string: x-evar65
278 = string: x-evar66
279 = string: x-evar67
28 = string: x-c_color
280 = string: x-evar68
281 = string: x-evar69
282 = string: x-evar70
283 = string: x-evar71
284 = string: x-evar72
285 = string: x-evar73
286 = string: x-evar74
287 = string: x-evar75
288 = string: x-cust1
289 = string: x-cust2
29 = string: x-color
290 = string: x-cust3
291 = string: x-cust4
292 = string: x-cust5
293 = string: x-cust6
294 = string: x-cust7
295 = string: x-cust8
296 = string: x-cust9
297 = string: x-cust10
298 = string: x-cust11
299 = string: x-cust12
3 = string: x-service
30 = string: x-cookies
300 = string: x-cust13
301 = string: x-cust14
302 = string: x-cust15
303 = string: x-cust16
304 = string: x-cust17
305 = string: x-cust18
306 = string: x-cust19
307 = string: x-cust20
308 = string: x-cust21
309 = string: x-cust22
31 = string: x-ct_connect_type
310 = string: x-cust23
311 = string: x-cust24
312 = string: x-cust25
313 = string: x-cust26
314 = string: x-cust27
```

```
315 = string: x-cust28
316 = string: x-cust29
317 = string: x-cust30
318 = string: x-cust31
319 = string: x-cust32
32 = string: x-connection_type
320 = string: x-cust33
321 = string: x-cust34
322 = string: x-cust35
323 = string: x-cust36
324 = string: x-cust37
325 = string: x-cust38
326 = string: x-cust39
327 = string: x-cust40
328 = string: x-cust41
329 = string: x-cust42
33 = string: x-country
330 = string: x-cust43
331 = string: x-cust44
332 = string: x-cust45
333 = string: x-cust46
334 = string: x-cust47
335 = string: x-cust48
336 = string: x-cust49
337 = string: x-cust50
338 = string: x-cust51
339 = string: x-cust52
34 = string: x-currency
340 = string: x-cust53
341 = string: x-cust54
342 = string: x-cust55
343 = string: x-cust56
344 = string: x-cust57
345 = string: x-cust58
346 = string: x-cust59
347 = string: x-cust60
348 = string: x-cust61
349 = string: x-cust62
35 = string: x-curr_rate
350 = string: x-cust63
351 = string: x-cust64
352 = string: x-cust65
353 = string: x-cust66
354 = string: x-cust67
355 = string: x-cust68
356 = string: x-cust69
357 = string: x-cust70
358 = string: x-cust71
359 = string: x-cust72
36 = string: x-curr_factor
360 = string: x-cust73
361 = string: x-cust74
362 = string: x-cust75
363 = string: x-cust76
364 = string: x-cust77
365 = string: x-cust78
366 = string: x-cust79
367 = string: x-cust80
368 = string: x-cust81
369 = string: x-cust82
37 = string: x-domain
```

```
370 = string: x-cust83
371 = string: x-cust84
372 = string: x-cust85
373 = string: x-cust86
374 = string: x-cust87
375 = string: x-cust88
376 = string: x-cust89
377 = string: x-cust90
378 = string: x-cust91
379 = string: x-cust92
38 = string: x-geo_city
380 = string: x-cust93
381 = string: x-cust94
382 = string: x-cust95
383 = string: x-cust96
384 = string: x-cust97
385 = string: x-cust98
386 = string: x-cust99
387 = string: x-cust100
388 = string: x-ecom_additional_data
39 = string: x-geo_country
4 = string: x-page_event
40 = string: x-geo_dma
41 = string: x-geo_region
42 = string: x-geo_zip
43 = string: x-hier1
44 = string: x-hier2
45 = string: x-hier3
46 = string: x-hier4
47 = string: x-hier5
48 = string: x-homepage
49 = string: c-ip
5 = string: x-hit_source
50 = string: x-j_jscript
51 = string: x-javascript
52 = string: x-java_enabled
53 = string: x-keywords
54 = string: x-language
55 = string: x-mobile_id
56 = string: x-new_visit
57 = string: x-os
58 = string: x-p_plugins
59 = string: x-plugins
6 = string: x-hit_time_gmt
60 = string: x-page_event_var1
61 = string: x-page_event_var2
62 = string: x-page_type
63 = string: cs-uri
64 = string: x-pagename
65 = string: x-paid_search
66 = string: x-partner_plugins
67 = string: x-persistent_cookie
68 = string: x-prop1
69 = string: x-prop2
7 = string: x-date_time
70 = string: x-prop3
71 = string: x-prop4
72 = string: x-prop5
73 = string: x-prop6
74 = string: x-prop7
75 = string: x-prop8
```

```
76 = string: x-prop9
77 = string: x-prop10
78 = string: x-prop11
79 = string: x-prop12
8 = string: x-visid_high
80 = string: x-prop13
81 = string: x-prop14
82 = string: x-prop15
83 = string: x-prop16
84 = string: x-prop17
85 = string: x-prop18
86 = string: x-prop19
87 = string: x-prop20
88 = string: x-prop21
89 = string: x-prop22
9 = string: x-visid_low
90 = string: x-prop23
91 = string: x-prop24
92 = string: x-prop25
93 = string: x-prop26
94 = string: x-prop27
95 = string: x-prop28
96 = string: x-prop29
97 = string: x-prop30
98 = string: x-prop31
99 = string: x-prop32
1 = DelimitedDecoder:
Delimiter = string: \t
Fields = vector: 391 items
0 = string: x-insight-row_type
1 = string: x-exclude_hit
10 = string: x-visit_num
100 = string: x-prop33
101 = string: x-prop34
102 = string: x-prop35
103 = string: x-prop36
104 = string: x-prop37
105 = string: x-prop38
106 = string: x-prop39
107 = string: x-prop40
108 = string: x-prop41
109 = string: x-prop42
11 = string: x-visit_page_num
110 = string: x-prop43
111 = string: x-prop44
112 = string: x-prop45
113 = string: x-prop46
114 = string: x-prop47
115 = string: x-prop48
116 = string: x-prop49
117 = string: x-prop50
118 = string: x-prop51
119 = string: x-prop52
12 = string: x-hitid_high
120 = string: x-prop53
121 = string: x-prop54
122 = string: x-prop55
123 = string: x-prop56
124 = string: x-prop57
125 = string: x-prop58
126 = string: x-prop59
```

```
127 = string: x-prop60
128 = string: x-prop61
129 = string: x-prop62
13 = string: x-hitid_low
130 = string: x-prop63
131 = string: x-prop64
132 = string: x-prop65
133 = string: x-prop66
134 = string: x-prop67
135 = string: x-prop68
136 = string: x-prop69
137 = string: x-prop70
138 = string: x-prop71
139 = string: x-prop72
14 = string: x-accept_language
140 = string: x-prop73
141 = string: x-prop74
142 = string: x-prop75
143 = string: cs(referrer)
144 = string: x-ref_domain
145 = string: x-ref_type
146 = string: x-resolution
147 = string: x-s_resolution
148 = string: x-search_engine
149 = string: x-search_page_num
15 = string: x-bot_type
150 = string: x-state
151 = string: x-transactionid
152 = string: x-truncated_hit
153 = string: x-ua_color
154 = string: x-ua_os
155 = string: x-ua_pixels
156 = string: x-uniques_exceeded
157 = string: cs(user-agent)
158 = string: x-user_server
159 = string: x-va_finder_id
16 = string: x-bot_id
160 = string: x-va_finder_detail
161 = string: x-va_closer_id
162 = string: x-va_closer_detail
163 = string: x-va_instance_event
164 = string: x-va_new_engagement
165 = string: x-zip
166 = string: x-last_hit_time_gmt
167 = string: x-first_hit_time_gmt
168 = string: x-visit_start_time_gmt
169 = string: x-last_purchase_time_gmt
17 = string: x-browser
170 = string: x-last_purchase_num
171 = string: x-first_hit_page_url
172 = string: x-first_hit_pagename
173 = string: x-visit_start_page_url
174 = string: x-visit_start_pagename
175 = string: x-first_hit_referrer
176 = string: x-visit_referrer
177 = string: x-visit_search_engine
178 = string: x-visit_keywords
179 = string: x-daily_visitor
18 = string: x-browser_height
180 = string: x-hourly_visitor
181 = string: x-monthly_visitor
```

```
182 = string: x-yearly_visitor
183 = string: x-weekly_visitor
184 = string: x-quarterly_visitor
185 = string: x-preloaded
186 = string: x-tnt
187 = string: x-survey
188 = string: x-mvvar1
189 = string: x-mvvar2
19 = string: x-browser_width
190 = string: x-mvvar3
191 = string: x-media
192 = string: x-page_event_media
193 = string: x-page_event_var3
194 = string: x-tnt_instances
195 = string: x-survey_instances
196 = string: x-mvvar1_instances
197 = string: x-mvvar2_instances
198 = string: x-mvvar3_instances
199 = string: x-campaign
2 = string: x-userid
20 = string: x-channel
200 = string: x-purchaseid
201 = string: x-product-num
202 = string: x-category
203 = string: x-product
204 = string: x-units
205 = string: x-revenue
206 = string: x-order
207 = string: x-cart_open
208 = string: x-cart_view
209 = string: x-checkout
21 = string: x-click_action
210 = string: x-cart_add
211 = string: x-cart_remove
212 = string: x-product_view
213 = string: x-evar1
214 = string: x-evar2
215 = string: x-evar3
216 = string: x-evar4
217 = string: x-evar5
218 = string: x-evar6
219 = string: x-evar7
22 = string: x-click_action_type
220 = string: x-evar8
221 = string: x-evar9
222 = string: x-evar10
223 = string: x-evar11
224 = string: x-evar12
225 = string: x-evar13
226 = string: x-evar14
227 = string: x-evar15
228 = string: x-evar16
229 = string: x-evar17
23 = string: x-click_context
230 = string: x-evar18
231 = string: x-evar19
232 = string: x-evar20
233 = string: x-evar21
234 = string: x-evar22
235 = string: x-evar23
236 = string: x-evar24
```

```
237 = string: x-evar25
238 = string: x-evar26
239 = string: x-evar27
24 = string: x-click_context_type
240 = string: x-evar28
241 = string: x-evar29
242 = string: x-evar30
243 = string: x-evar31
244 = string: x-evar32
245 = string: x-evar33
246 = string: x-evar34
247 = string: x-evar35
248 = string: x-evar36
249 = string: x-evar37
25 = string: x-click_source_id
250 = string: x-evar38
251 = string: x-evar39
252 = string: x-evar40
253 = string: x-evar41
254 = string: x-evar42
255 = string: x-evar43
256 = string: x-evar44
257 = string: x-evar45
258 = string: x-evar46
259 = string: x-evar47
26 = string: x-click_tag
260 = string: x-evar48
261 = string: x-evar49
262 = string: x-evar50
263 = string: x-evar51
264 = string: x-evar52
265 = string: x-evar53
266 = string: x-evar54
267 = string: x-evar55
268 = string: x-evar56
269 = string: x-evar57
27 = string: x-code_ver
270 = string: x-evar58
271 = string: x-evar59
272 = string: x-evar60
273 = string: x-evar61
274 = string: x-evar62
275 = string: x-evar63
276 = string: x-evar64
277 = string: x-evar65
278 = string: x-evar66
279 = string: x-evar67
28 = string: x-c_color
280 = string: x-evar68
281 = string: x-evar69
282 = string: x-evar70
283 = string: x-evar71
284 = string: x-evar72
285 = string: x-evar73
286 = string: x-evar74
287 = string: x-evar75
288 = string: x-cust1
289 = string: x-cust2
29 = string: x-color
290 = string: x-cust3
291 = string: x-cust4
```

```
292 = string: x-cust5
293 = string: x-cust6
294 = string: x-cust7
295 = string: x-cust8
296 = string: x-cust9
297 = string: x-cust10
298 = string: x-cust11
299 = string: x-cust12
3 = string: x-service
30 = string: x-cookies
300 = string: x-cust13
301 = string: x-cust14
302 = string: x-cust15
303 = string: x-cust16
304 = string: x-cust17
305 = string: x-cust18
306 = string: x-cust19
307 = string: x-cust20
308 = string: x-cust21
309 = string: x-cust22
31 = string: x-ct_connect_type
310 = string: x-cust23
311 = string: x-cust24
312 = string: x-cust25
313 = string: x-cust26
314 = string: x-cust27
315 = string: x-cust28
316 = string: x-cust29
317 = string: x-cust30
318 = string: x-cust31
319 = string: x-cust32
32 = string: x-connection_type
320 = string: x-cust33
321 = string: x-cust34
322 = string: x-cust35
323 = string: x-cust36
324 = string: x-cust37
325 = string: x-cust38
326 = string: x-cust39
327 = string: x-cust40
328 = string: x-cust41
329 = string: x-cust42
33 = string: x-country
330 = string: x-cust43
331 = string: x-cust44
332 = string: x-cust45
333 = string: x-cust46
334 = string: x-cust47
335 = string: x-cust48
336 = string: x-cust49
337 = string: x-cust50
338 = string: x-cust51
339 = string: x-cust52
34 = string: x-currency
340 = string: x-cust53
341 = string: x-cust54
342 = string: x-cust55
343 = string: x-cust56
344 = string: x-cust57
345 = string: x-cust58
346 = string: x-cust59
```

```
347 = string: x-cust60
348 = string: x-cust61
349 = string: x-cust62
35 = string: x-curr_rate
350 = string: x-cust63
351 = string: x-cust64
352 = string: x-cust65
353 = string: x-cust66
354 = string: x-cust67
355 = string: x-cust68
356 = string: x-cust69
357 = string: x-cust70
358 = string: x-cust71
359 = string: x-cust72
36 = string: x-curr_factor
360 = string: x-cust73
361 = string: x-cust74
362 = string: x-cust75
363 = string: x-cust76
364 = string: x-cust77
365 = string: x-cust78
366 = string: x-cust79
367 = string: x-cust80
368 = string: x-cust81
369 = string: x-cust82
37 = string: x-domain
370 = string: x-cust83
371 = string: x-cust84
372 = string: x-cust85
373 = string: x-cust86
374 = string: x-cust87
375 = string: x-cust88
376 = string: x-cust89
377 = string: x-cust90
378 = string: x-cust91
379 = string: x-cust92
38 = string: x-geo_city
380 = string: x-cust93
381 = string: x-cust94
382 = string: x-cust95
383 = string: x-cust96
384 = string: x-cust97
385 = string: x-cust98
386 = string: x-cust99
387 = string: x-cust100
388 = string: x-ecom_additional_data
389 = string: x-mcvisid
39 = string: x-geo_country
390 = string: x-tnt-action
4 = string: x-page_event
40 = string: x-geo_dma
41 = string: x-geo_region
42 = string: x-geo_zip
43 = string: x-hier1
44 = string: x-hier2
45 = string: x-hier3
46 = string: x-hier4
47 = string: x-hier5
48 = string: x-homepage
49 = string: c-ip
5 = string: x-hit_source
```

```
50 = string: x-javascript
51 = string: x-javascript
52 = string: x-java_enabled
53 = string: x-keywords
54 = string: x-language
55 = string: x-mobile_id
56 = string: x-new_visit
57 = string: x-os
58 = string: x-p_plugins
59 = string: x-plugins
6 = string: x-hit_time_gmt
60 = string: x-page_event_var1
61 = string: x-page_event_var2
62 = string: x-page_type
63 = string: cs-uri
64 = string: x-pagename
65 = string: x-paid_search
66 = string: x-partner_plugins
67 = string: x-persistent_cookie
68 = string: x-prop1
69 = string: x-prop2
7 = string: x-date_time
70 = string: x-prop3
71 = string: x-prop4
72 = string: x-prop5
73 = string: x-prop6
74 = string: x-prop7
75 = string: x-prop8
76 = string: x-prop9
77 = string: x-prop10
78 = string: x-prop11
79 = string: x-prop12
8 = string: x-visid_high
80 = string: x-prop13
81 = string: x-prop14
82 = string: x-prop15
83 = string: x-prop16
84 = string: x-prop17
85 = string: x-prop18
86 = string: x-prop19
87 = string: x-prop20
88 = string: x-prop21
89 = string: x-prop22
9 = string: x-visid_low
90 = string: x-prop23
91 = string: x-prop24
92 = string: x-prop25
93 = string: x-prop26
94 = string: x-prop27
95 = string: x-prop28
96 = string: x-prop29
97 = string: x-prop30
98 = string: x-prop31
99 = string: x-prop32
Name = string: Adobe SC decoder
Fields = vector: 0 items
Log Entry Condition = AndCondition: 0 items
Parameters = vector: 0 items
Stage = string: Default
Transformations = vector: 0 items
```

Data Workbench 6.1 Release Notes

Data Workbench 6.1 release notes include new features, upgrade requirements, bug fixes, and known issues.

To view previous features and fixes based for each past release, see the [release note archives](#).

New Features

Data Workbench 6.1 includes these new features:

Features	Description
64-bit Windows upgrade	The data workbench server, report server, and client components are upgraded to run only on 64-bit Windows operating systems.
Propensity Scoring	<p>Scoring your audience lets you identify customer loyalty and statistically perceive who is likely to convert a sale or interact with a story or campaign. Propensity scoring now includes these visualizations to view models and show the changing correlation of selected metrics.</p> <ul style="list-style-type: none">• The Model Viewer examines a logistic regression model generated with Propensity Scoring, displaying the coefficient weights of each input variable (including the constant term) and their statistical error range.• Lift and Gain charts are used to evaluate the potential increase of a scored data model.• The Confusion Matrix gives four counts by the combination of Actual Positive (AP), Actual Negative (AN), Predicted Positive (PP), and Predicted Negative (PN).• Starting with v6.1, you now have a Save option to save propensity scores based on two types: dimensions, or dimensions and metrics.• You can now click Ctrl-Alt and drag and drop to add elements in Propensity Scoring and the Cluster Builder. Previously to add table elements, you had to drag from the table to the Elements box.
Simplified Chinese localization on page 133	Data workbench now supports Simplified Chinese for the client application. Data workbench also supports the Installing the Input Method Editor on page 133 as a secondary text entry process for international languages.
Math Functions	You can now add Mathematical functions to metrics, math transformations, and worksheet cells to further calculate datasets.
Statistical Callouts	Tables now offer a statistics summary call-out for metric columns. The call-out can display the mean, standard deviation, minimum and maximum values, variance, and total count for the column. It can be factored in to any selection and evaluation.
Correlation Matrix filter	The Correlation Matrix has been updated with a Binary Filter to let you constrain values for one or both of the correlated metrics, allowing you to better focus your comparison. Also, you can now add Dimension elements from a Dimension table by clicking Ctrl + Alt and dragging elements to the matrix column or row to be evaluated.
Hide Fallout Labels in Funnel on page 136	Toggle between displaying and hiding fallout labels in a Funnel visualization by right-clicking the title and selecting Hide Fallout .
Sorting Table Columns on page 136 alphabetically or by ordinals	From a Dimension table, you can sort elements alphabetically or by ordinals using the new arrow in the title of the column. The # character will display when a column is sorted by ordinals.
New Keyboard Shortcuts on page 135	Additional shortcut keys have been added for adding and editing workspaces and configuration files.

Known Issues

- When importing a workspace, an error message is displayed even though the import was successful.

Workaround: Click OK to ignore the error. The workspace is imported successfully.

Simplified Chinese Localization Issues

- The dialog title and message displayed after clicking "Submit" when setting the target in the Scoring visualization are unreadable.

Workaround: None.

- When using word wrap in the Worksheet visualization, localized words are not being wrapped correctly. Extra junk characters are being added to the string.

Workaround: None

- Unable to launch `Insight.exe` if the installation directory is named with non-English characters.

Workaround: Keep default names or rename using only English characters in the folder path to launch executables.

Data Workbench 6.1 features

Data Workbench 6.1 includes the following features.

Propensity Scoring

Propensity scoring lets you define customers based on their possibility of a successful conversion or completion of a specified event. It allows you to maximize the potential impact of efforts before executing a process or directing a campaign.

The Value of Propensity Scoring

Propensity scoring lets you perform data discovery to identify hidden behaviors or patterns that exists across your data. Specifically, propensity scoring helps you identify clusters of similar customers using more focused and objective means rather than simple segmentation or filtering. In addition, propensity scoring lets you set up predictive capabilities to identify behavior for your company's high-value customer.

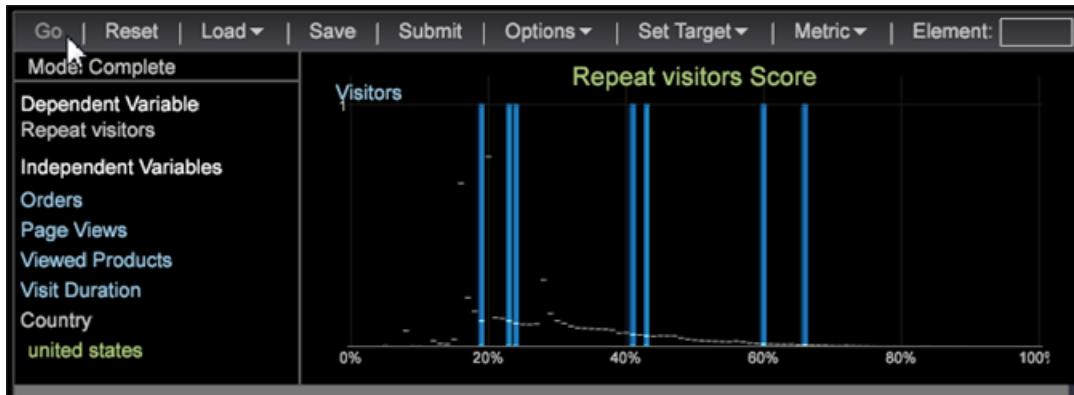
Once you have identified the high-value audience, you can then engage them for the greatest effect. For example, if you are Business to Business company, you may have sales call leads that allow you to then score the leads and identify their likelihood to convert offline. Because every lead increases costs, creating an incentive to identify prospective customers with the highest likelihood of converting a sale is the most effective and the least expensive way to focus your resources.

Propensity scoring provides the ability to identify those factors that are most predictive of a particular score or to increase the likelihood of an event taking place, but it can also be applied to answer specific questions: Will the customer convert? Will the customer respond to an email? Will the customer repurchase? Propensity scoring lets you answer these questions and identify visitors with an inclination for action that can then be set up and scored.

In addition, you can use filters to define a subset of visitors to be scored using the optional **Training Filter** feature. If no filter is applied, then all visitors are targeted for scoring.

Features of the Propensity Scoring Visualization

To open the Propensity Scoring Visualization, click **Add > Visualization > Predictive Analytics > Scoring > Propensity Score**.



The Propensity Scoring Visualization includes these features accessible from its toolbar:

Toolbar Feature	Description
Go	Click to run the scoring process after setting up parameters.
Reset	Clear all settings in the visualization.
Load	Loads a previously created ScoreDim that allows you to change and/or rebuild the scoring model.
Save	Save the Propensity scoring visualization as a Dim file to be accessed and opened as needed.
Submit	Submit scoring task for server-side processing.
Options	Set the Training Filter to limit the subset of visitors. The default filter is Train on Everyone , but you can change it by making workspace selections or building a filter using the Filter Editor .
Set Target	Set the Dependent Variable.
Metric	Add Metrics as Independent Variables.
Elements	Drag Dimension elements using the <Ctrl> + <Alt> keys from Dimension tables.

See also:

1. The [Propensity Gain and Lift Charts](#) on page 130. These views can be opened from a complete scoring model or from Add Visualization> Predictive Analytics > Scoring.
2. The [Model Viewer](#). These views can be opened from a complete scoring model or from Add Visualization> Predictive Analytics > Scoring.
3. The [Complex Filter Description](#) on page 131 feature.

Using the Propensity Scoring Visualization

- **Define one or more filters to define the visitor population for scoring.** This optional **Training Filter** lets you target visitors based on selected criteria. If no training filter is applied, then all visitors are targeted for scoring. If the Training Filter is set, then the scoring result is meaningful to the defined visitor population, although each visitor will still be given a score.
- **Identify the positive visitors.** To define the dependent variable to specify a target filter identifying the positive visitors that match the desired outcome. This can be as simple as Revenue > \$10, or a much more complex filter.
- **The Target filter is not allowed to be the same as the Training filter.** Logically, the Target Filter should be an addition to the Training Filter, resulting in a positive subset of the visitor population to be scored.
- **Select variables of interest (independent variables) as inputs to the Propensity Scoring algorithm.** These can be Metrics or individual elements of a Dimension. Propensity Scoring will start preprocessing just as in

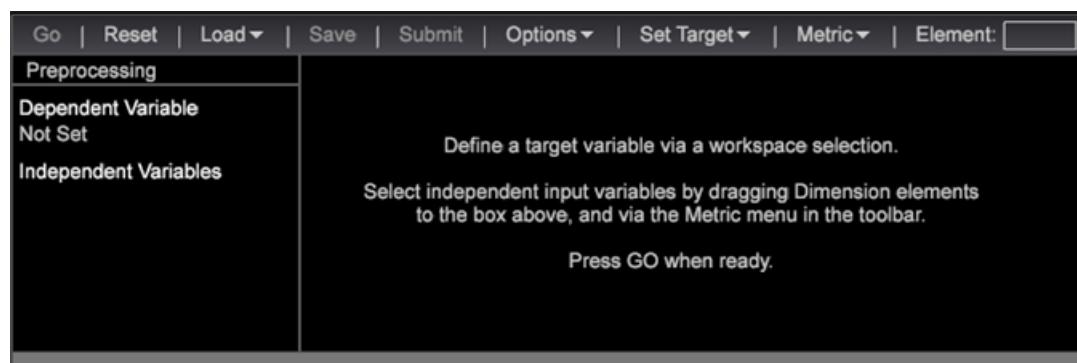
[Visitor Clustering](#) on page 155. The system begins capturing a certain amount of samples that match the definition of the previously set training filter (if any). Currently, the sample size is set as 10% of scoring population (defined by training filter), with a minimum of 20,000 and maximum of 100,000, and is bound to the scoring population size.

- A Score Dimension has elements ranging from 0% to 100% that determines the likelihood of the visitors matching the Target variable.

Setting up Propensity Scoring

Follow these steps to use the Propensity Scoring visualization.

1. Open a new workspace and click **Add > Visualization > Predictive Analytics > Scoring > Propensity Score**.



2. Set the **Target** (the dependent variable).

Set the dependent variable by selecting:

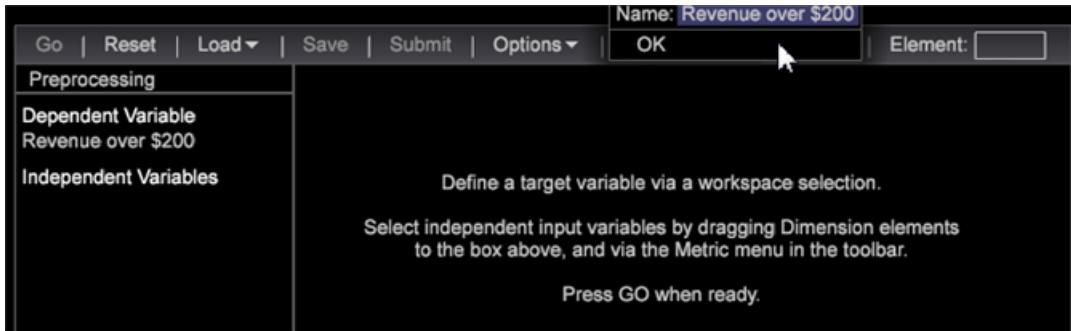
- **Dimension elements:** Right-click in the workspace and select **Table**. Then select a Dimension elements as your dependent variable.

OR

- **Filter Editor.** Click **Add > Visualization > Filter Editor** to open the Filter Editor visualization.



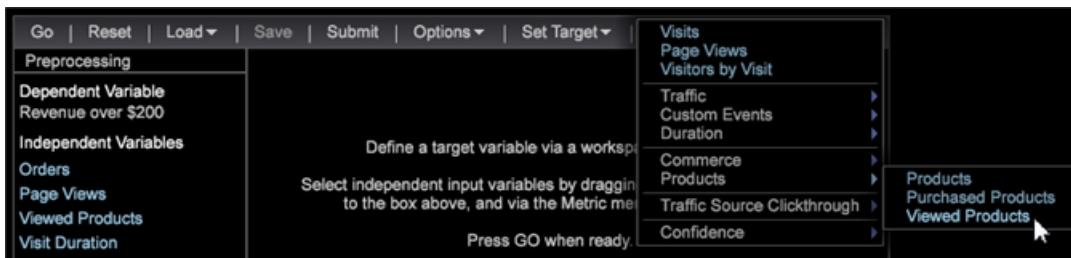
After selecting a Dimension element or Filter as the dependent variable, click **Set Target**, enter a name to describe the dependent variable. Then click **OK** (and make sure the filter box is highlighted) to set the Target.



The name you give the target is the dependent variable that will appear in the left pane.

3. Add independent variables.

Add the independent variables using Metrics or Dimension Elements.



- **Metrics.** From the Propensity Scoring toolbar, select a metric from the **Metrics** menu.
- **Dimension elements:** Right-click in the workspace and select **Table**. Select one or more Dimension elements and drag to the left column under **Independent Variables** or to the **Element** box using the <Ctrl> + <Alt> keys.

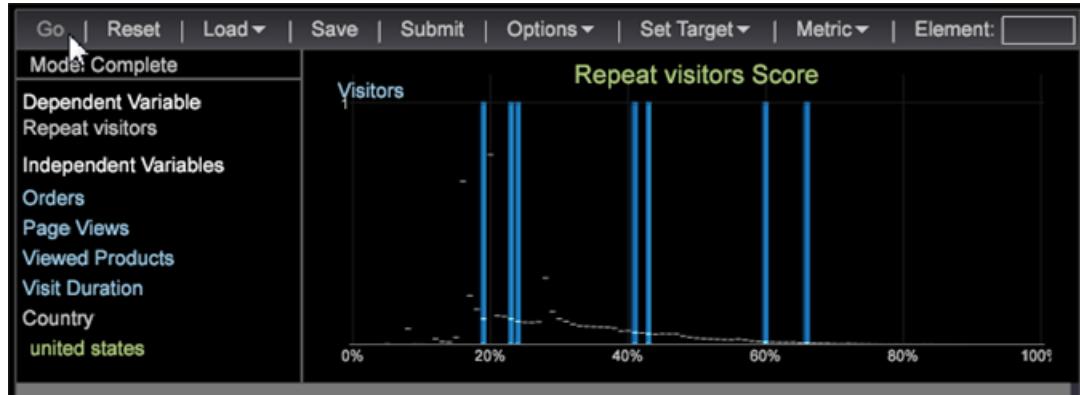
4. Set **Training Filter**.

You can define the set of visitors that you want to score by clicking **Options > Set Training Filter** from the Propensity Scoring toolbar. This will provide a subset of data built using only the visitors that you want to score. For example, who visited in the last month, visitors who reside in Australia, or visitors who viewed specific products.

The default filter is **Train on Everyone**, but you can change it by activating **Dimension Elements** in a table or building a filter using the **Filter Editor**.

After selecting a Dimension element or building a filter and while activated, click **Options > Set Training Filter**, enter a name to describe the filter, and then click **OK**.

5. Once you have identified all your inputs, press **Go**.



The scoring process will begin by passing over the data multiple times. It will then display the results as bar charts over a percentage line.

6. Save Propensity Score.

Starting with 6.1, you now have an option when using the Save Propensity Score:

- Dimension
- Dimension and Metric

You can end up with two saved files, both a dimension and a defined metric.



Note: If you submit the Propensity Score for processing you will get a dimension only.

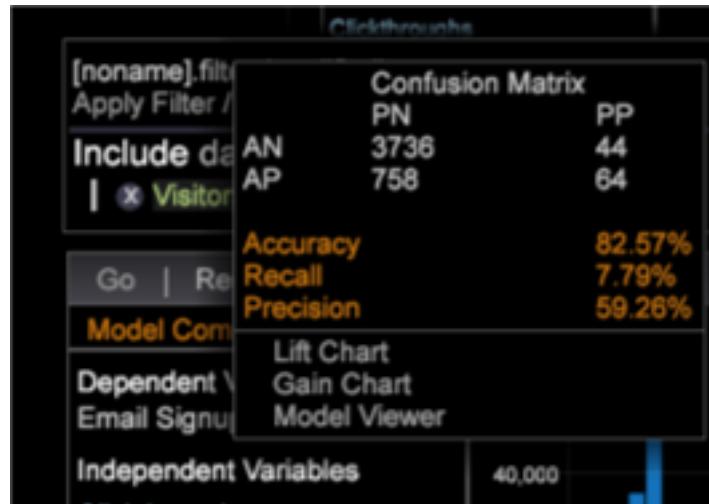
The derived metric is the associated average score metric.

7. Check for accuracy.

The system will display **Model Complete** and generate a scoring model when the process is complete.

Right-clicking on **Model Complete** will identify the accuracy of the scoring model as defined by the system. Values ranging from 0 percent to 100 percent will identify the likelihood of the visitors matching the **Target** variable.

The Confusion Matrix gives four counts by the combination of Actual Positive (AP), Actual Negative (AN), Predicted Positive (PP), and Predicted Negative (PN). These numbers are obtained by applying the resulted scoring model to the 20% withheld testing data of which we know the true answer. If the score is greater than 50%, it is predicted as a positive case (matching the defined event).



Accuracy	Indicates how accurate the model is by identifying the correct predictions over all predictions. $(TP + TN)/(TP + FP + TN + FN)$
Recall	Identifies the ability to re-identify the scoring model. $TP / (TP + FN)$
Precision	Identifies the level of discrepancy. $TP / (TP + FP)$

8. Open a *Propensity Gain and Lift Charts* on page 130, or the *Model Viewer* on page 132.

Right-click on the **Model Complete** visualization and select **Lift Chart**, **Gain Chart**, or **Model Viewer**.

Calculating Propensity Scoring

The statistical calculations for Propensity Scoring are defined.

Conceptually, the score calculated for each visitor is an estimated probability that the specified event (defined by the target filter) might happen, resulting in a score value range from 0 to 100 percent. The scoring procedure uses existing samples as training data to find the relationship between the event probability and the selected independent variables of interest.

Mathematically, such relationships are reflected in each quantitative value associated for each independent variable. Those values are called model coefficients. ScoreDim currently uses the Iteratively Reweighted Least Squares (IRLS) algorithm to estimate the model coefficients. IRLS goes through the samples multiple times until the difference of coefficients between current pass and the previous pass is less than 1.0e-6, at which point it is called **converged**. However, depending on the data, IRLS may not be able to reach convergence.

In such case, the model training iteration will terminate when

- the coefficient difference gets larger instead of smaller,
- 1,000 passes have been reached, or
- a mathematical error prevents continuing iteration.

If IRLS doesn't converge, a backup algorithm called Stochastic Gradient Decent (SGD) will be used. SGD will also go through the training samples multiple times. But unlike IRLS, the SGD model coefficients are controlled so that the difference between iteration will always decrease in an exponential manner. Similarly, SGD will terminate when the coefficient difference falls below 1.0e-6 or 100,000 passes have been reached. The failure of IRLS and engagement of SGD will be recorded in trace log.

For both algorithms, not all samples go into model training. Eighty percent are currently used to train the model. After the model is trained, the remaining twenty percent samples will be used to assess the model strength in terms of Accuracy, Recall, and Precision that is calculated from the confusion matrix. The closer to 100%, the better the scoring model.

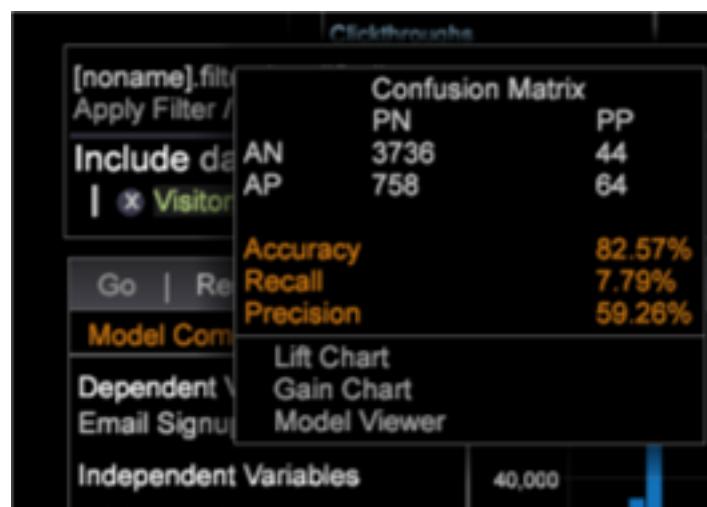
Propensity Gain and Lift Charts

The Lift and Gain charts offer visualizations for evaluating the potential performance of a scored model to evaluate performance over defined portions of the audience.

Gain and lift charts are visualizations built to evaluate the potential performance of the scored model. These charts evaluate performance over each portion of the population.

To Open a Lift or Gain chart

1. Select Add Visualization > Predictive Analytics > Scoring.
2. Hover over **Model Complete** of a saved score.



About Lift and Gain Charts

The Lift and Gain Charts are useful visual tools for measuring the value of a predictive model. Both charts consist of a lift curve (green) and a baseline (pink). For the **Gain Chart**, the distance between the lift curve and the baseline represents how much you can improve performance in responses (or the "gain") from using the predictive mode. The gain is realized by prioritizing and targeting the prospects (customers/visitors) who are most likely to convert, rather than marketing to customers/visitors at random. In this way, you can quantify the expected value of using the predictive model to choose which prospects to contact.

Similar to the Gain Chart, the **Lift Chart** shows how much more likely you are to receive positive responses than if you contacted prospects at random. You want the distance between the lift curve and the baseline to be as large as possible, representing larger expected gains from using the predictive model to contact customers. Mathematically, the gain and lift charts are defined as follows:

- **Gain** = (Expected Response using Predictive Model to Contact Prospects) / (Expected Response from Randomly Contacting Prospects)
- **Lift** = (Expected Response among a Specific Group Size of Prospects identified using the Predictive Model) / (Expected Response among the same Specific Group Size of Prospects identified Randomly)

Example of Lift and Gain Charts

For example, consider the example of a retailer who wants to launch an email re-marketing campaign to sell yoga pants. Historically, the analyst expects an average response rate of 20 percent based upon past email

re-marketing campaigns similar to this one. While the analyst has nearly 5 million customers in its email database, the business only wants to market to those customers that are most likely to respond to the email and purchase. In this way, the company will maximize the ROI of the campaign while ensuring that they don't unnecessarily send emails to uninterested customers. Given an expected response rate of 20 percent, the marketer and analyst expect that approximately 1 million customers are likely to respond and purchase. Rather than randomly guessing which of those customers will be among the 20 percent responses, the analyst wants to be smart about predicting which of the one million prospects (among the database of 5 million customers) are most likely to respond.

Using Adobe's Audience Scoring capability, the analyst defines success as a prospect clicks on an email and purchases yoga pants (the dependent variable). After selecting the independent variables (based upon experience and knowledge gained from analyzing data correlations and audience clustering among other analyses), each prospect is scored on their likelihood of positively responding the email re-marketing campaign (clicking on the email and purchasing yoga pants). The analyst opens the resulting Gain and Lift charts based upon the predictive model.

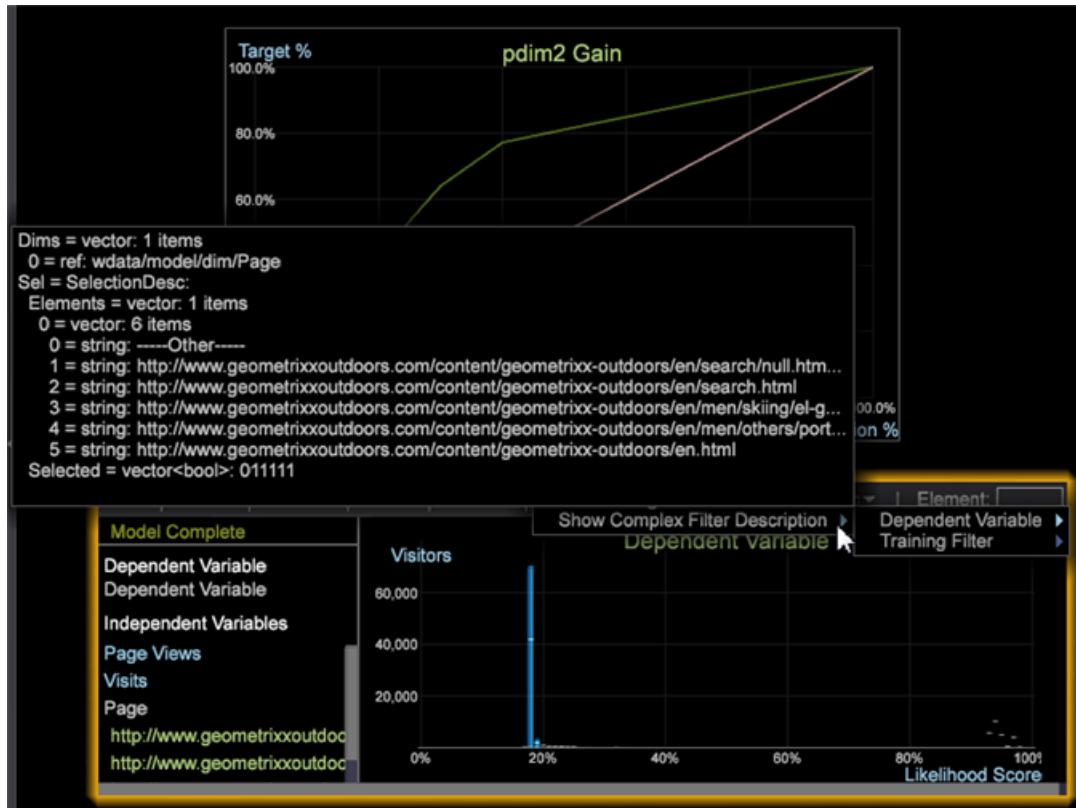
The y-axis shows the percentage of the cumulative expected positive responses. In our example, we expect a total of 1 million positive responses. A value of 20% on the y-axis corresponds to 20% of the 1 million expected positive responses, or 200,000 positive responses. The x-axis shows the percentage of prospective customers contacted. In our example, the x-axis represents a fraction of the 5 million customers in the email database. The baseline (Pink) is the overall response rate - if you contact X% of prospects then you will receive X% of the total positive responses. Using the predictive model, the lift curve (green) shows the percentage of positive responses obtained (y-axis) by contacting a given percentage of prospects (x-axis).

The Lift chart plots the expected lift as a result of using the predictive model to determine the top one million prospects most likely to purchase yoga pants after receiving and clicking on the email. For contacting 20 percent of randomly selected prospects using no predictive model, you should expect to get 20 percent of responders. However, using the predictive model to identify the top 20 percent of prospects most likely to respond, you expect to get 50% of responders. The y-value of the lift curve at 20 percent is $50/20 = 2.5$. The lift chart shows how much more likely you are to receive respondents than if you contact a random sample of prospects. For example, by contacting only 20 percent of prospects based on the predictive model you will reach 2.5 times as many respondents as compared to having not used any predictive model.

Complex Filter Description

You can add a Show Complex Filter Description for Dependent Variables and for the Training Filter.

To **Show the Complex Filter Description**, from the scoring dialog select Options > Show Complex Filter Description.



Model Viewer

A model viewer lets you generate a logistic regression model using the Propensity Scoring feature.

The Model Viewer displays the coefficient weights of each input variable (including the constant term) and their statistical error range. Input variables showing a high absolute coefficient and small margin of error are the most significant predictors in the model.

To Open a Model Viewer chart

1. Select Add Visualization > Predictive Analytics > Scoring.
2. Hover over Model Complete of a saved score.



The input variables with a coefficient ≥ 1 are positive influences on the propensity model. The coefficients that are < 1 are negative influences on the propensity model. The range defined within the parentheses is the error, and indicates the consistency of the input variable across the successful population.

Simplified Chinese localization

The data workbench client application now supports Simplified Chinese.

To install Simplified Chinese:

Before configuring `Insight.exe` and supporting files, you must exit the client application.

1. Create a shortcut that passes in the command-line setting to the `insight.exe` file.

```
Insight.exe -zh-cn
```

2. Configure `Insight.cfg` to support single and double-byte font characters.

Data workbench currently supports both English and Simplified Chinese. You can select fonts to support both of these languages:

```
FONTS = vector: 2 items
0 = string: SimSun
1 = string: Arial
```

3. Restart `Insight.exe`.

Installing the Input Method Editor

Data workbench now supports the Input Method Editor (IME) as a secondary text entry process for international languages.

IMEs allow you to enter international characters using a variety of methods suited for your local language. Data workbench provides an input dialog box that allows you to open and use your desired IME for text fields.



Note: For the data workbench 6.1 release, only the virtual Simplified Chinese keyboard will be supported. Inputting other languages through the IME could result in unexpected behavior.

Using an IME

To use the floating IME text input feature:

1. Click **Alt + Space** for any text input area.
2. Enter values using your system's IME.
3. Close the input dialog by selecting the **Enter** key or clicking the **OK** button.

The dialog will disappear and the characters will then appear in the selected field.

Updating the `Insight.cfg` file

To employ the IME, you must update the `Insight.cfg` file with this setting:

```
Localized IME = bool: true
```

If this setting does not exist in the configuration file, then pressing **Alt + Space** will not engage the IME feature.

Starting Insight in another language: To better support localized assets like a splash screen and to support multiple languages in the future, data workbench requires command-line arguments identifying the language to load. The default language is English.

Starting data workbench in Chinese requires you to invoke `Insight.exe` with the "`-zh-cn`" argument:

```
Insight.exe -zh-cn
```

(These command line arguments are not case sensitive.)

Syntax for Math Functions

Data workbench now allows you to employ additional mathematical calculations.

Statistical calculations can be used evaluate and display in Metric editor visualizations, worksheet data, and for statistical callouts.

Calculation	Function	Input
absolute	<code>abs()</code>	Math transformation, Metric editor and Worksheet cell
arccosine	<code>acos()</code>	Math transformation, Metric editor and Worksheet cell
arcsine	<code>asin()</code>	Math transformation, Metric editor and Worksheet cell
arctangent	<code>atan()</code>	Math transformation, Metric editor and Worksheet cell
cosine	<code>cos()</code>	Math transformation, Metric editor and Worksheet cell
exponential	<code>exp()</code>	Math transformation, Metric editor and Worksheet cell
logarithm	<code>log()</code>	Math transformation, Metric editor and Worksheet cell
maximum	<code>max()</code>	Math transformation and Statistics callout
minimum	<code>min()</code>	Math transformation and Statistics callout
In (natural logarithm)	<code>ln()</code>	Math transformation, Metric editor and Worksheet cell
sine	<code>sin()</code>	Math transformation, Metric editor and Worksheet cell
tangent	<code>tan()</code>	Math transformation, Metric editor and Worksheet cell

Statistical Callouts

Statistical callouts measure meaningful relationships to identify hidden opportunities and variables of interest for more advanced data mining capabilities in audience clustering and visitor response scoring.

Statistical callouts expand the algorithms so that more types of data can be correlated, such as binomial variables (yes/no, 0/1, or purchaser/non-purchaser) correlated with countable metrics (visits, orders, or downloads).

To add statistical callouts:

1. In a table, right-click the metric header.
2. Select **Statistics** and then select or clear the checkmarks for each required setting. All in the Display Callout are selected as the default setting.



The callout can return statistical values factored into the dataset columns.

Calculation	Description
Count	Returns the number of rows in a dataset.
Maximum	Identifies the maximum Metric value across all elements of the dimension.
Minimum	Identifies the minimum Metric value across all elements of the dimension.
Mean	The mean is the arithmetic average of the Metric values of elements in the Dimension, calculated by the total sum divided by the count (sum/count).
Standard Deviation	The standard deviation shows how much variation exists from the expected mean. A lower standard deviation shows the data points close to the mean. A higher standard deviation shows that the data points are spread across a large range of values.
Total	Returns the total sum of the Metric values.
Variance	A measure of the variance of the Metric values from the Metric mean for that dimension. It is equal to the square of the standard deviation.

New Keyboard Shortcuts

New shortcut keys are available for workspaces and configuration files.

Keyboard shortcuts from the worktop thumbnail view

Keys	Action
Ctrl + N	Opens a new workspace. If you are limited to creating a single type of blank workspace, then the workspace will open immediately. If you have rights to open multiple types of workspaces based on your implementation, then Ctrl + N displays a menu allowing you to choose the type of workspace to be created.
Ctrl + O	Imports a workspace.

Keyboard shortcuts within a workspace

Keys	Action
Ctrl + P	Prints a workspace.
Ctrl + W	Closes a workspace without saving.
Ctrl + S	Saves a workspace.

Sorting Table Columns

Sort table columns alphabetically or by ordinals.

To better select elements in a Dimension table, you can order the first column alphabetically or by ordinals by selecting the **Sort** menu option.

The # character will display when a column is sorted by ordinals (the default).

Select Sort Option

To change sorting options between ordinal and alphabet, right-click and select **Sort**. Click the arrow to reverse the order.

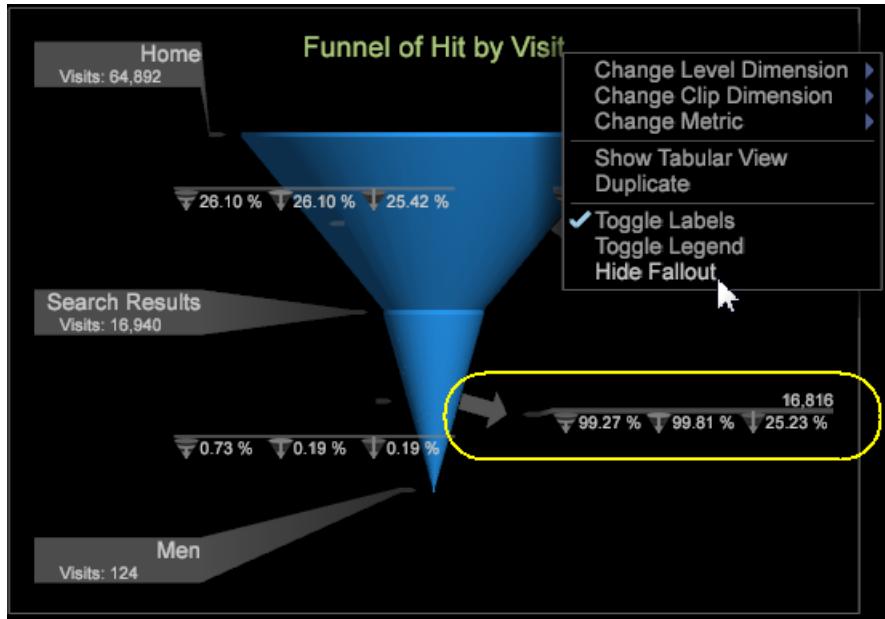


Note: You can sort other columns by ordinal by clicking the name of the column.

Hide Fallout Labels in Funnel

Toggle to open fallout labels in a Funnel visualization.

The Funnel visualization identifies where a customer abandons a marketing campaign or diverts from a defined conversion path while interacting with your website or cross-channel campaign. The left side of the Funnel visualization displays the results of a visit or visitors, while the right side displays the "Fallout" of those who abandon a specified path.



When in a **Funnel** visualization, you can right-click the title and select **Hide Fallout** from the menu to hide the fallout labels.

Data Workbench 5.5 to 6.1 Upgrade

Follow these steps to update to data workbench v6.1 from your Insight v5.5x installation.

Step 1: [Server Upgrade](#) on page 137

Step 2: [Report Server Upgrade](#) on page 139

Step 3: [Client Upgrade](#) on page 139



Important: The server, report server, and client components are upgraded to run on 64-bit Windows operating systems.

Server Upgrade

Follow these steps to update the **Server v6.1** components:

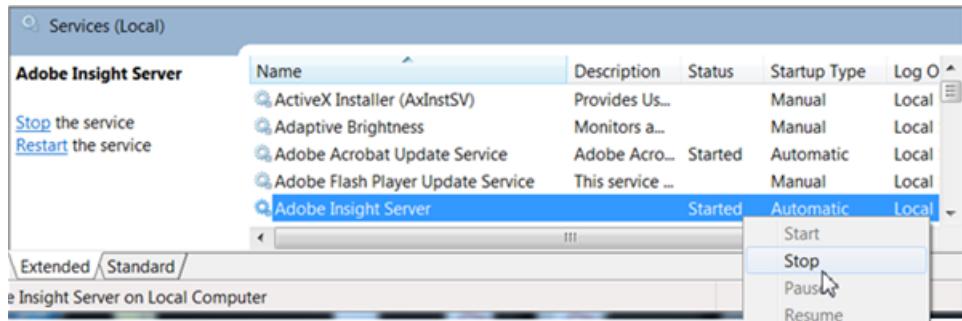
1. Using the **Software and Docs** profile, open the **Start Here** workspace and download all needed server packages to a local folder.

- Download **Server Packages \ v6.1** zip folders and extract all files.

The **Server** package includes **Lookup** and **Profile** folders with the **Base** and **Transform** lookup files to add and replace to update the server.

- Download new **Profiles** folders.
- Download updated **Lookup** folders.
- Download the **Report Server \ v6.1** package.
- Download additional **Sensor**, **Documentation**, and **Dashboard** files as needed for your system.

2. Stop the **Adobe Insight Server** service.



3. From the downloaded **Server** package:
 - a. Replace the **Server\Bin** folder to update the **InsightServer64.exe** and supporting files.
 - b. Replace the **Server\Profiles** folder. You can overwrite all files.
 - c. Update the **Server\Lookups** folder. You will want to add the newly downloaded files to the custom files already located in the folder.
 - d. Replace the **Server\Software** folder to update **Insight.exe** and **ReportServer.exe**
 - e. Update the **Server\Scripts** folder to update **TnTSend.exe**.
4. If you employ **DeviceAtlas**, then you will need to *update the bundle* located in the **Server\Lookups** folder.
5. Set Directories in the **Profile.cfg** file to ensure that the vector is updated to reflect the number of items for each profile.

For example, to enable the **Predictive Analytics** profile you will need to update this setting.

```
Directories = vector: 5 items
  0 = string: Base\\
  1 = string: Geography\\
  2 = string: Predictive Analytics\\
  3 = string: Adobe SC\\
  4 = string: Profile Name\\
```

6. Configure and save the **PAServer.cfg** file to upgrade the Predictive Analytics feature.

If you want to submit Predictive Analytics jobs to the servers, then you will need to configure the **Server > Predictive Analytics > Dataset > PAServer.cfg** file to manage server-side clustering submissions.

The custom profile should inherit the settings from the Predictive Analytics configuration profile, allowing you to configure and save the **PAServer.cfg** based on your site's implementation.

7. Define the **Log Source ID**.

The **Recording of Rows per Log Source** was added in **v6.04** and defined in the custom profile's **Log Processing.cfg** file by adding a uniquely named **Log Source ID**.

```
Log Processing.cfg
  Log Source ID = string: <Name your ID Here>
```

If you do not have the Log Source ID defined, then you will get the following error:

```
Missing Log Source ID in log processing.cfg.
Log Source ID must be defined for all log sources.
```

8. Because the **EventMessages.dll** has been updated, it is required that you unregister and then register the **Adobe Insight Server** across the cluster.

-
- InsightServer64.exe /unregserver
 - InsightServer64.exe /regserver
9. Start the **Adobe Insight Server** service across the cluster.

The server installation is now complete.

Report Server Upgrade

 **Important:** Before upgrading to **Report Server v6.1**, you must first upgrade to **Server v6.1**.

1. Using the **Software and Docs** profile, download **v6.1** from the **Report Server** package to a local folder.
2. Copy **Report Server 6.1** from the downloaded package and replace the profile packages.

 **Note:** The `Insight.zbin` file in the `install` folder is a backup file used for localization, and must be present in the `install` directory. This file or other `.zbin` files will be used depending on the command-line settings passed when starting up.

3. (optional) Modify the report server configuration file to support double-byte characters.

Data workbench currently supports English (-en-us) and Chinese (-zh-cn). You need to set a font to support single and double-byte characters:

```
Report Server.cfg - Add Fonts
  Fonts = vector: 2 items
    0 = string: SimSun
    1 = string: Arial
```

The Windows operating system must also have the listed fonts installed.

4. Configure Report Server v6.1.
 - a. Stop the **Adobe Insight Report Server** service.
 - b. Launch a command prompt as an "Administrator".
 - c. Navigate to the Report Server install folder.
 - d. Delete the Report Server service using the following command:

```
ReportServer.exe /unregserver
```

5. Start the service based on the language settings:

```
ReportServer.exe -RegServer -Locale -en-us (English)
ReportServer.exe -RegServer -Locale -zh-cn (Simplified Chinese)
```

6. To verify that Report Server is running with the correct settings, open up **Windows Service Manager** and right-click **Adobe Insight Report Server - Properties**. The path to the executable will display the updated command-line settings.

The report server installation is now complete.

Client Upgrade

 **Important:** Before upgrading to **Client v6.1**, the administrator must first upgrade to **Server v6.1**.

1. Launch `Insight.exe` but DO NOT connect to any profiles.

-
2. Edit the `Insight.cfg` file to not update software automatically.

```
Update Software = bool: false
```

3. Connect to **Software and Docs** profile (softdocs).
4. Download `Software\Insight Client\v6.10`.
5. (optional) Modify `insight.cfg` to support double-byte characters.

Data workbench currently supports both English and Simplified Chinese. Select fonts to support both of these languages:

```
Fonts = vector: 2 items
0 = string: SimSun
1 = string: Arial
```

6. Exit out of the client.
7. Copy the files in the downloaded **v6.1** client package to the `Install` folder.



Note: The `Insight.zbin` file in the install folder is a backup file used for localization, and must be present in the install directory. This file or other `.zbin` files will be used depending on the command-line settings passed when starting up.

For example, to launch Simplified Chinese, create a shortcut that passes in the command-line setting.

```
Insight.exe -zh-cn
```

If you want to launch in English (default), then no command-line change is necessary.

8. Launch `Insight.exe` for English or the shortcut that you created for another language.
9. Connect to your profile and allow the client to synchronize with the server.
10. (optional) To employ the IME, make these changes to the `Insight.cfg` file:

```
Localized IME = bool: true
```

The Input Method Editor (IME) allows you to input international characters.

11. (optional) Edit the `Insight.cfg` file to automatically update software:

```
Update Software = bool: true
```

See instructions for implementing the IME.

12. Restart again after the profile synchronization to employ the most recent `.zbin` file.

The client installation is now complete.

Data Workbench 6.0 to 6.1 Upgrade

Follow these steps to update to data workbench v6.1 from your data workbench v6.0x installation.

Step 1: [Server Upgrade](#) on page 140

Step 2: [Report Server Upgrade](#) on page 142

Step 3: [Client Upgrade](#) on page 143



Important: The server, report server, and client components are upgraded to run on 64-bit Windows operating systems.

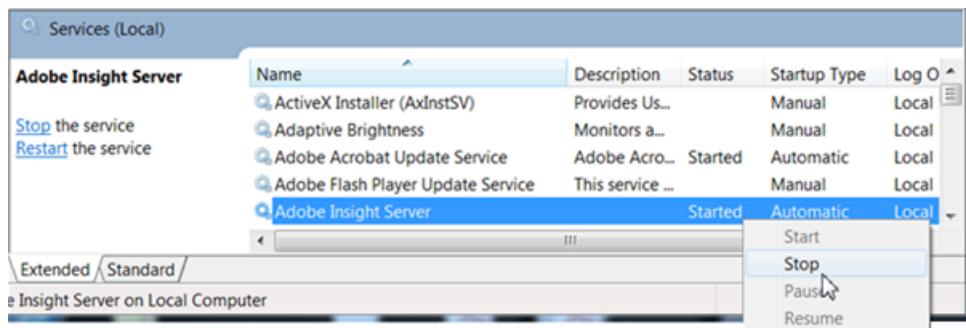
Server Upgrade

Follow these steps to update the **Server v6.1** components:

- Using the **Software and Docs** profile, open the **Start Here** workspace and download all needed server packages to a local folder.
 - Download **Server Packages \ v6.1** zip folders and extract all files.

The Server package includes **Lookup** and **Profile** folders with **Base** and **Transform** profiles to update the server.

 - Download the **Profiles** folders.
 - Download the **Lookup** folders.
 - Download the **Report Server \ v6.1** package.
 - Download additional **Sensor**, **Documentation**, and **Dashboard** files as needed for your system.
- Stop the **Adobe Insight Server** service.



- From the downloaded **Server** package:
 - Replace the **Server\Bin** folder to update the **InsightServer64.exe** and supporting files.
 - Replace the **Server\Profiles** folder. You can overwrite all files.
 - Update the **Server\Lookups** folder. You will want to add the newly downloaded files to the custom files already located in the folder.
 - Replace the **Server\Software** folder to update **Insight.exe** and **ReportServer.exe**
 - Update the **Server\Scripts** folder to update **TnTSend.exe**.
- If you employ **DeviceAtlas**, then you will need to *update the bundle* located in the **Server\Lookups** folder.
- Configure the **Profile.cfg** file to ensure that the vector is updated to reflect the number of items for each profile.

For example, to enable the **Predictive Analytics** profile you will need to update this setting.

```
Directories = vector: 5 items
  0 = string: Base\\
  1 = string: Geography\\
  2 = string: Predictive Analytics\\
  3 = string: Adobe SC\\
  4 = string: Profile Name\\
```

- Configure and save the **PAServer.cfg** file for the Predictive Analytics feature.

If you want to submit Predictive Analytics jobs to the servers, then you will need to configure the **Server > Predictive Analytics > Dataset > PAServer.cfg** file to manage server-side clustering submissions.

The custom profile should inherit the settings from the Predictive Analytics configuration profile, allowing you to configure and save the `PAServer.cfg` file based on your site's implementation.

7. Define the **Log Source ID**.

The **Recording of Rows per Log Source** was added in **v6.04** and defined in the custom profile's `Log Processing.cfg` file by adding a uniquely named **Log Source ID**.

```
Log Processing.cfg
  Log Source ID = string: <Name your ID Here>
```

If you do not have the Log Source ID defined, then you will get the following error:

```
Missing Log Source ID in log processing.cfg.
Log Source ID must be defined for all log sources.
```

8. Because the `EventMessages.dll` has been updated, it is required that you unregister and then register the **Adobe Insight Server** across the cluster.
 - `InsightServer64.exe /unregserver`
 - `InsightServer64.exe /regserver`

9. Start the **Adobe Insight Server** service across the cluster.

The server installation is now complete.

Report Server Upgrade



Important: Before upgrading to **Report Server v6.1**, you must first upgrade to **Server v6.1**.

1. Using the **Software and Docs** profile, download **v6.1** from the **Report Server** package to a local folder.
2. Copy **Report Server 6.1** from the downloaded package and replace the profile packages.



Note: The `Insight.zbin` file in the `install` folder is a backup file used for localization, and must be present in the `install` directory. This file or other `.zbin` files will be used depending on the command-line settings passed when starting up.

3. (optional) Data workbench currently supports English (`-en-us`) and Chinese (`-zh-cn`). You need to set a font to support single and double-byte characters:

```
Report Server.cfg - Add Fonts
  Fonts = vector: 2 items
  0 = string: SimSun
  1 = string: Arial
```

The Windows operating system must also have the listed fonts installed.

4. Configure Report Server v6.1 for localization.

- a. Stop the **Adobe Insight Report Server** service.
- b. Launch a command prompt as an "Administrator".
- c. Navigate to the Report Server `install` folder.
- d. Delete the Report Server service using the following command:

```
ReportServer.exe /unregserver
```

- e. Start the service based on language settings:

```
ReportServer.exe -RegServer -Locale -en-us (English)
ReportServer.exe -RegServer -Locale -zh-cn (Simplified Chinese)
```

-
5. To verify that Report Server is running with the correct settings, open up **Windows Service Manager** and right-click **Adobe Insight Report Server - Properties**. The path to the executable will display the updated command-line settings.

The report server installation is now complete.

Client Upgrade



Important: Before upgrading to **Client v6.1**, the administrator must first upgrade to **Insight Server v6.1**.

1. Launch `Insight.exe` but do not connect to any profiles.
2. Edit the `Insight.cfg` file.

```
Update Software = bool: true
```

3. Connect to your profile.

Allow the client to synchronize with the server and your client will be upgraded with the latest v6.1client profiles, executables, and configuration files.



Note: The `Insight.zbin` file in the `install` folder is a backup file used for localization and must be present. This file or other `.zbin` files will be used depending on the command-line settings passed when starting up.

See [Installing the Input Method Editor](#) on page 133 to add an `insight.zbin` file required for localized settings.

Additional Client Settings

Before configuring `Insight.exe` and supporting files, you must exit the client application.

To install Simplified Chinese:

1. Create a shortcut that passes in the command-line setting to the `Insight.exe` file.

```
Insight.exe -zh-cn
```

2. Configure `Insight.cfg` to support single and double-byte font characters.

Data workbench currently supports both English and Simplified Chinese. You can select fonts to support both of these languages:

```
Fonts = vector: 2 items
0 = string: SimSun
1 = string: Arial
```

The Windows operating system must also have the requested fonts installed.

3. Launch the shortcut that you created to synchronize profiles and the updated `.zbin` file.

To employ the Input Method Editor (IME).

IME allows you to input international characters.

1. Update the `Insight.cfg` file with these settings:

```
Localized IME = bool: true
```

2. Launch the shortcut that you created to synchronize profiles and the updated `.zbin` file.

The client installation is now complete.

Data Workbench 6.0 Release Notes

View the previous Data Workbench release notes.

See also the [previous Insight release notes](#).

Data Workbench 6.0 Release Notes

New features introduced in Data Workbench 6.0.4, including bug fixes and known issues.

New Features

Data workbench (Insight 6.0) includes these new features and visualizations for added reporting capabilities and predictive analysis tools.

To view previous features and fixes based for each past release, see the [release note archives](#).

Data Workbench Features	Description
Funnel Visualization on page 148	The Funnel visualization lets you define the sequential process flow of your customers and provides visibility into the fallout of visitors at each step in the process.
Visitor Clustering on page 155	Clustering lets you leverage customer characteristics to dynamically categorize visitors and generate cluster sets based on selected data inputs for customer analysis and targeting.
Correlation Matrix on page 164	Correlation Analysis lets you quickly identify relevant data relationships to extend and enhance your analysis.
DeviceAtlas Distribution on page 555	The DeviceAtlas JSON file will now be distributed in a .bundle file (a renamed .tar.gz) along with DeviceAtlas.dll and DeviceAtlas64.dll.

Client Upgrade Requirements

Complete these upgrade tasks for data workbench (Insight 6.0) client features:

Updating the .zbin file for the client

Data workbench now supports an Input Method Editor (IME) as a secondary text entry process that allows you to enter international characters from your keyboard using a floating text box. Data workbench will support English by default but also allows you to load other files to support international languages, such as a virtual Chinese keyboard (Pinyin IME).

A new dictionary file (a .zbin file) is required by the client application before updating to version 6.0. You can obtain the needed .zbin file from the Software and Docs profile (Softdocs).

Prerequisites:

- Before upgrading to the Insight 6.0 client and Report Server 6.0, the Insight administrator must first upgrade to Insight Server 6.0.
- The Insight administrator will need to choose a zbin file based on language (en-us.zbin, zh-cn.zbin), copy the language file, then rename it to insight.zbin, and place the renamed file in the root directory of the Report Server where the executable is located. Then restart the Insight Report Server.

See the [Server Upgrade Requirements](#) for additional server-side upgrade information.

To upgrade the zbin file for the client (from version 5.x to 6.0):

1. To make sure the client does not get updated from the Insight Server during this upgrade, set your Insight.cfg argument to false.

```
Update Software = bool: false
```

2. Restart the Insight client.

-
3. Navigate to the Software and Docs profile (SoftDocs profile) and download the required **Insight.zbin** file:
Software\Insight Client\v6.00\Insight_6.00.zip
 4. Copy the Insight.zbin file to the same folder as the Insight.exe file.
 5. To make sure the Insight client now gets updated from the Insight Server, change the Insight.cfg file argument to true:

```
Update Software = bool: true
```

6. Restart the client.

Your client will synchronize with the server and you will see a message stating that your client is downloading. At the conclusion of the download, you will get a message asking if you want to restart your Insight client.

7. Click **OK** to restart the client.

The client will start and upgrade to version 6.0.

8. Restart the client again for the Insight.zbin client synchronization to take effect.

If you get the following message, then it means the zbin was not placed in the correct folder location alongside the Insight.exe file.

```
Insight Terminated: The backup dictionary file insight.zbin  
is missing.
```

To correct the issue, delete Insight.exe and rename the latest version of Insight.exe.old to Insight.exe, and then start again with Step 1 above.

Server Upgrade Requirements

Complete these upgrade tasks for Insight 6.0 server features:

Update all Insight Server 6.0 packages. Insight 6.0 includes server packages that need to be updated, including the new Predictive Analytics profile.

-  **Important:** It is recommended that users upgrade their server clusters with fresh installations of Insight Server 6.0 when updating.

It is also recommended that client upgrade their server clusters with fresh installation of Insight Server 6.0.

Upgrade Server cluster

Prepare the language file (.zbin file). The Insight administrator selects the <language>.zbin file for the required language (for example: en-us.zbin , zh-cn.zbin) located in the base/localization/<language>.zbin folder. The administrator then copies the language file and renames it to "insight.zbin".

After preparing the language file (.zbin), both the Insight Client and Report Server need to be updated. The Insight Client is updated during the [client upgrade process](#), but in most cases the Insight administrator will update the Report Server.

Update Report Server with a language file (.zbin file).

For all languages, Report Server 6.0 requires the "insight.zbin" file copied to the Report Server root folder.

Update the Report Server language files:

1. Add the renamed "insight.zbin" file to the root ReportServer directory.
2. The Report Server configuration file (reportserver.cfg) requires font settings for double-byte languages. For example, Chinese requires the addition of fonts using SimSun:

```
Report Server.cfg - Add Fonts
```

```
Fonts = vector: 2 items
```

```
0 = string: SimSun  
1 = string: Arial
```

3. A parameter for Report Server 6.0 needs to be passed in the command line for localization, for example:

```
ReportServer.exe -Locale -zh-cn  
ReportServer.exe -Locale -en-us
```



Note: If a locale is not specified, then the Report Server defaults to the language selected in the insight.zbin file.

Follow the steps to launch the ReportServer as a service with the Locale parameters:

- a. Launch a Command Prompt as an Administrator.
- b. Navigate to the ReportServer install folder.
- c. Type the following command to start the service:
 - For English: ReportServer.exe -RegServer -Locale -en-us
 - For Chinese: ReportServer.exe -RegServer -Locale -zh-cn

4. To verify if the ReportServer is running with the correct parameters:

- a. Open up Windows Service Manager.
- b. Right-click Adobe Insight Report Server - Properties.

The path to executable will contain the parameters:

```
ReportServer.exe -Service ReportServer -Locale -en-us
```

Modify Profile Configuration file for Predictive Analytics. Insight administrator will need to modify the custom profile.cfg file to include the Predictive Analytics profile to be available in Insight.

Example of the profile.cfg entry:

```
Example ("profile.cfg"):  
Profile = profileInfo:  
    Active = bool: true  
    Directories = vector: 5 items  
        0 = string: Base\  
        1 = string: Predictive Analytics\  
        2 = string: Geography\  
        3 = string: Adobe SC\  
        4 = string: Custom Profile\  
    
```

Update the PAServer.cfg file. If you want to submit Predictive Analytics clustering jobs to Insight Servers, then you will need to configure the PAServer.cfg file for handling server-side clustering submissions.

The custom profile should inherit the PAServer.cfg from the Predictive Analytics profile (Server\Profiles\Predictive Analytics\Dataset). Configure and save the PAServer.cfg per your implementation site.



Note: Once PAServer.cfg is configured and saved to custom profile, an Insight Server restart is required across the site.

Upgrade Report Server. You will need to update the fonts and the start-up parameters for Report Server.

Prerequisites:

- Before upgrading Report Server 6.0, the Insight administrator must first upgrade to Insight Server 6.0.

-
- For all languages, Report Server 6.0 requires the addition of Insight.zbin to the Report Server root folder. Make sure the base/localization/<language>.zbin is copied and renamed to "insight.zbin". Copy it to the root of the Report Server directory.

Update the Fonts and Start-up parameters:

1. Report Server requires font setting for double byte in order to output to different languages, for example:

Report Server.cfg - Add Fonts

```
Fonts = vector: 2 items
0 = string: SimSun
1 = string: Arial
```

2. Parameter for Report Server 6.0 needs to be passed in the command-line for localization purposes.

To launch the Report Server as a service with the Locale parameters:

1. Stop the Report Server Service.
2. Launch a Command Prompt as an Administrator.
3. Navigate to the Report Server install folder.
4. Type the following command to start the service:

```
ReportServer.exe -RegServer -Locale -en-us
```

To verify if the Report Server is running with the correct parameters:

1. Open up Windows Service Manager
2. Right-click **Adobe Insight Report Server - Properties**.
3. The path to executable will contain the parameters:

```
ReportServer.exe -Service ReportServer -Locale -en-us
```

Upgrade the SiteCatalyst data feed for Insight 6.0. The filename format of the SiteCatalyst data feed for Insight 6.0 has changed.

Current filename format:

```
RSID_YYYYMMDD_HH0000.tsv.gz
```

New filename format:

```
YYYYMMDD-RSID_HH0000.tsv.gz
```



Note: This change does not affect users currently deployed with the *wbench/ecom* version of the SiteCatalyst data feed.

The filename format change will allow for the full use of the Insight Start and End time declarations during log processing. This enables the process to evaluate if the contents of the file should be read, rather than filter all source files using a row by row search.

In most cases, a rename process was implemented upon receipt of the file to provide the full use of this capability. This modification provides the required naming convention by default without the need and overhead of a secondary process.

[To use the new SiteCatalyst data feed:](#)

1. Determine how the receiving process will handle the new filename format.

The standard rename/move scripts deployed during implementation moves the files with a ".gz" extension, and only performs a rename if the filename matched the filename format with the preceding RSID.

The new filename format:

```
YYYYMMDD-RSID_HH0000.tsv.gz
```

2. Evaluate the defined log source paths to confirm that all files will be read.

If you already have a rename script implemented, then you are already defining your log sources to read this new filename format.

Fixes

- Now, the key combination to leave a workspace without saving changes has been updated to <Ctrl> + <Backspace>. Previously, you voided changes and closed a workspace by pressing <Ctrl> + <Delete>.

Data Workbench 6.0 features

Data Workbench 6.0 included the following features.

Funnel Visualization

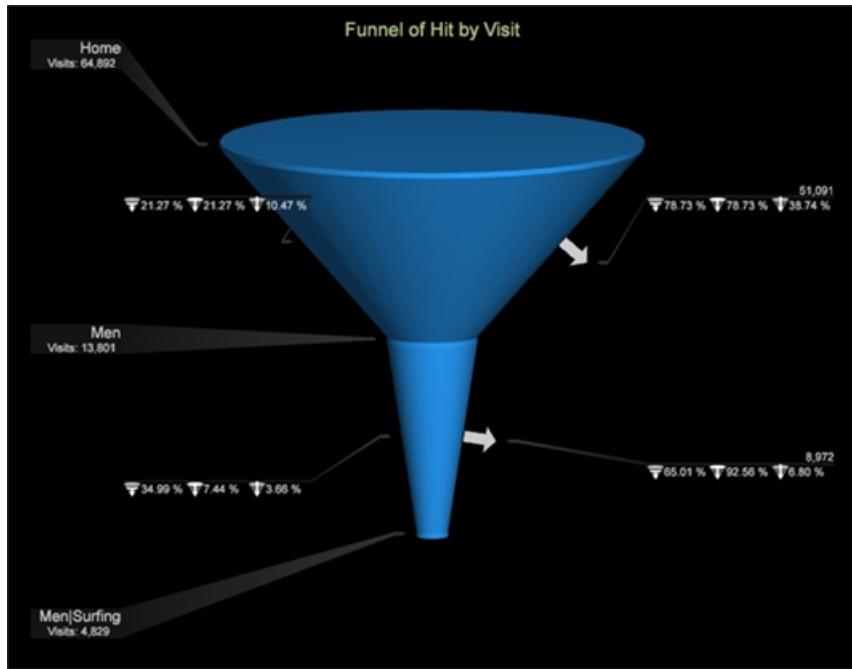
The Funnel visualization lets you identify where customers abandon a marketing campaign or divert from a defined conversion path while interacting with your website or cross-channel campaign.

The funnel visualization identifies converting page views to purchases, and lets you see where customers fall out of the process. Gaining visibility into customer decisions at each step lets you understand where they are being deterred, what path they tend to follow, and when customers leave your site and where they go.

Proper understanding of customer navigation allows architects to design and target web pages based on trending interest, and lets marketers better interpret customer actions and interactions against specific campaigns.

About the Funnel

The Funnel visualization is much like the Path Browser in building a path to analyze your visitor's sequence of hits, to identify who is falling out (who left the path) or falling through (who followed the path). It also identifies where visitors went after each step along the campaign path, and where they navigated to after falling out or falling through the defined steps.



In addition to web data, you can perform funnel analysis across all types of cross-channel data supported by the platform. Any data element from any source can be represented in the Funnel visualization.

The Funnel visualization provides various levels of data:

- **Funnel by Visit.** This funnel is based on a "per visit" interaction.
- **Funnel by Visitor.** This funnel is based on a "per visitor" interaction. This view shows the data based on the visit organized by visitor.
- The **Clip dimension and Level dimension.** Funnel dimensions can be modified by choosing Clickthrough, Hit, Product, Visit or Visitor Level dimensions.
- **Metrics.** You can change the underlying Funnel metrics from Visits used to build the funnel to any metric in your profile. The Funnel visualization allows you to drill into your data and analyze the patterns of visitors with multiple baseline metrics.

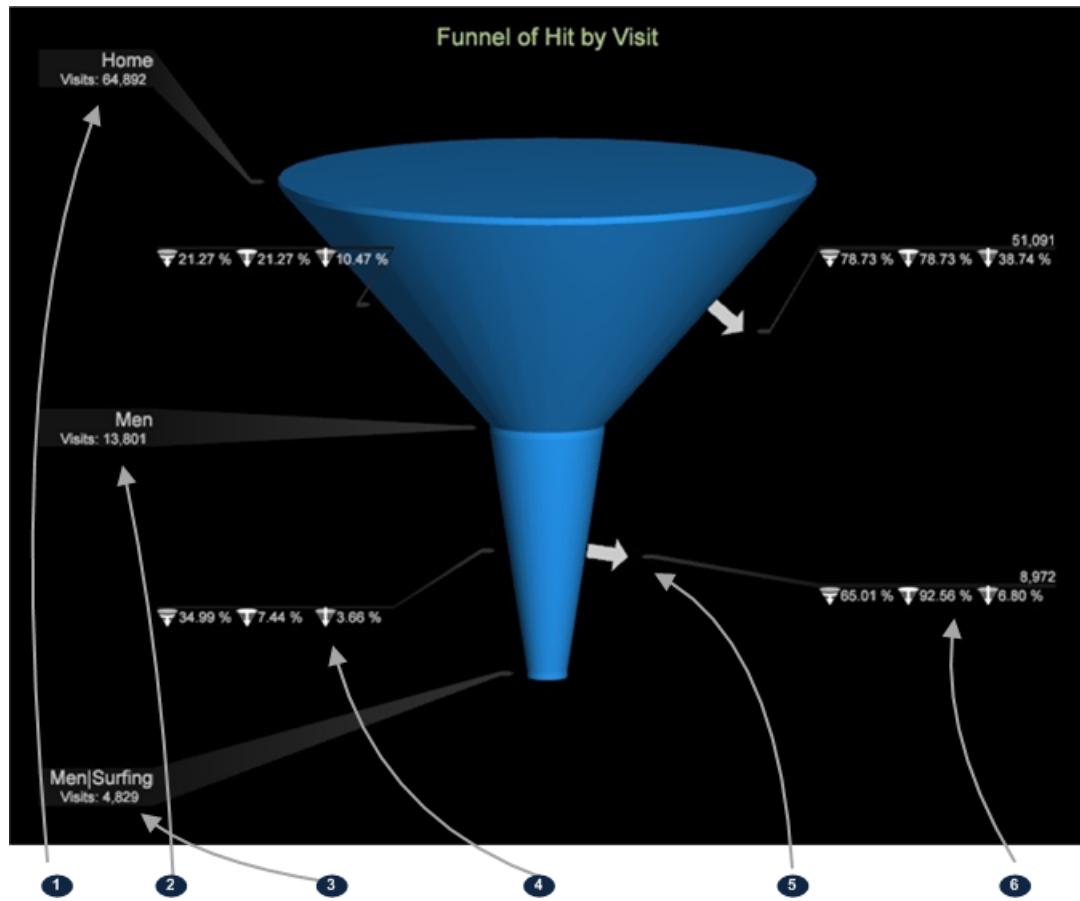


Note: In the Funnel visualization, you can apply any dimension at the countable level. This is in contrast to the Path Browser and Process Maps that restrict your choice of metrics. Analysts have multiple choices when applying a metric in the Funnel visualization.

Funnel Features

The Funnel visualization includes features to build a funnel with multiple dimensions, raw visitor numbers, percentage of visitors at each step, and separate scopes.

Here are the basic features of the funnel visualization.



1	First Element	First funnel step in the process.
2	Third Element	Third funnel step in the process. Note: The selected elements do not have to be from the same dimension.
3	Fall-through Percentage	Percentage who completed the defined path displayed in three scopes.
4	Fallout Browser	Fallout arrow. Right-click and select Add Path Browser to see what other path visitors took.
5	Percent Fallout	Percentages that describe three scopes of fallout for users that did not complete the path. Percentages are presented in three scopes: 
		 The percentage of fallout from the step previous to this step.
		 The percentage of fallout from the first step in the funnel.



The percentage of fallout based on the total number of visitors.

Funnel Steps

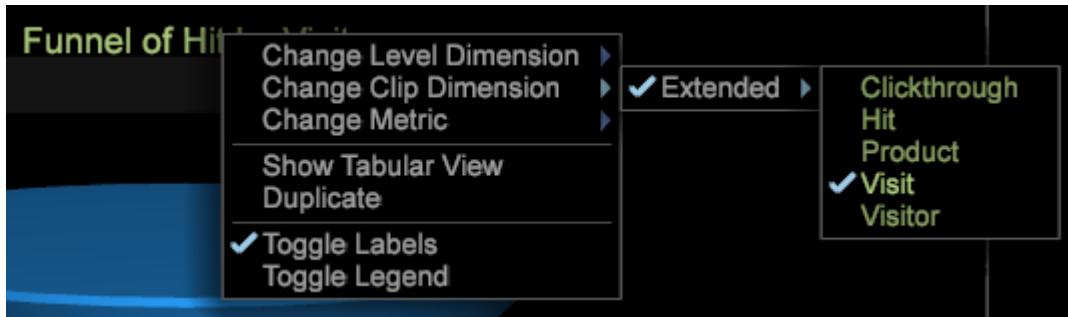
The disks in a funnel represent the steps in the navigation, the cones represent the fall-through from one step to the next, and the arrows represent the fallout. Clicking a cone will select the users who fell through at that point and include them in the current workspace filter. Clicking an arrow will select the visitors who fell out.



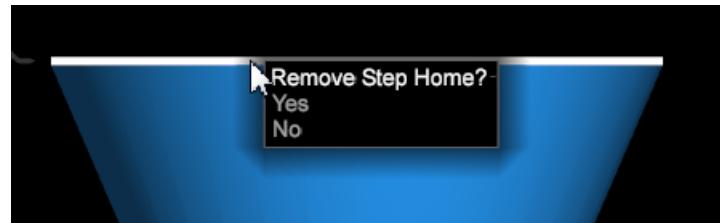
Note: The Funnel visualization has a limit of eight steps that can be applied.

Additional Funnel Features and Functionality

- **Adjust the clip and level of the funnel.** Select the Funnel option from the Visualization menu. After the funnel is created, you can right-click on the title to adjust the clip and level to any countable metric in your system.



- **Drag more elements.** Add more elements to your funnel by dragging and dropping them from the Dimension table to the funnel using the <Ctrl> + <Alt> keys. You can drag multiple steps at the same time from the Dimension table by selecting multiple items (using <Ctrl> + click) and then dragging them to the Funnel visualization using the <Ctrl> + <Alt> keys..
- **Delete a step:** Delete elements by right-clicking on the step in the visualization and clicking Yes.



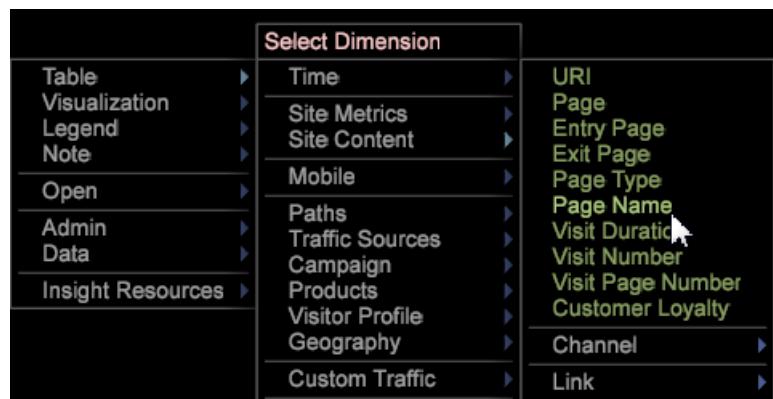
- **Rearrange the steps you have dragged to the funnel.** Simply click the step to select it and drag it to another position to rearrange the steps.
- **Open a Path Browser.** You can see more detail about where customers fall through or fall out of the process through the [Adding a Path Browser](#) on page 154 feature.
- **Add more steps.** You can add a maximum of eight steps to each funnel visualization.
- **Change the metric.** The metric can be changed so the steps are counting visits or some other metric at each step. Available options vary by dataset.

-
- **Display in a tabular view.** Right-click the title to display the Funnel Visualization menu and click **Show Tabular View**. Once in tabular view, you can select **Show Graph View** to return to graphic representation of the funnel. To open the Tabular View, right-click on the title and select Show Tabular View from the menu.
 - **Compare sequences.** An efficient way to compare two similar sequences is to display their two visualizations side-by-side. You can also display both the tabular view and the graph view side-by-side using the Duplicate feature. To open, right-click on the title and select Duplicate from the menu.

Building a Funnel

Follow these steps to build a new funnel visualization

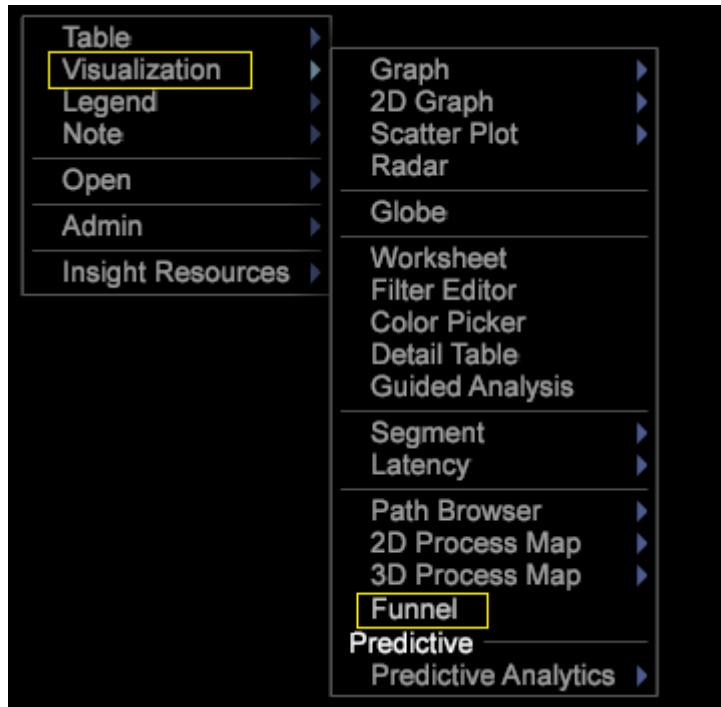
1. From a new workspace, open a dimension table.



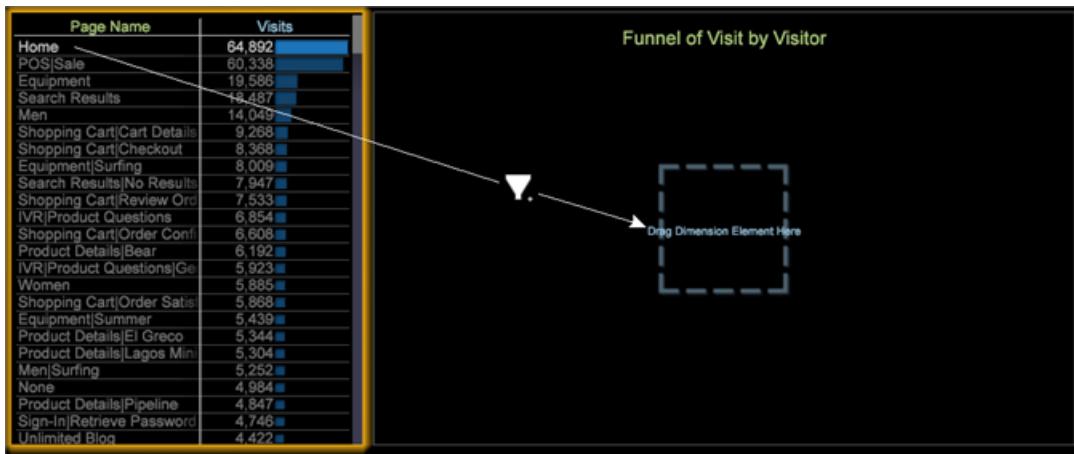
For example, right-click on the workspace and select **Table > Site Content > Page Name**

2. Open a funnel visualization.

For example, right-click on the workspace and select **Visualization > Funnel**.



3. Add a dimension by selecting the element, then drag while then holding <Ctrl> + <Alt>.
4. Drag the element into position over the target zone, **Drag Dimension Element Here**.



5. Add additional dimensions.

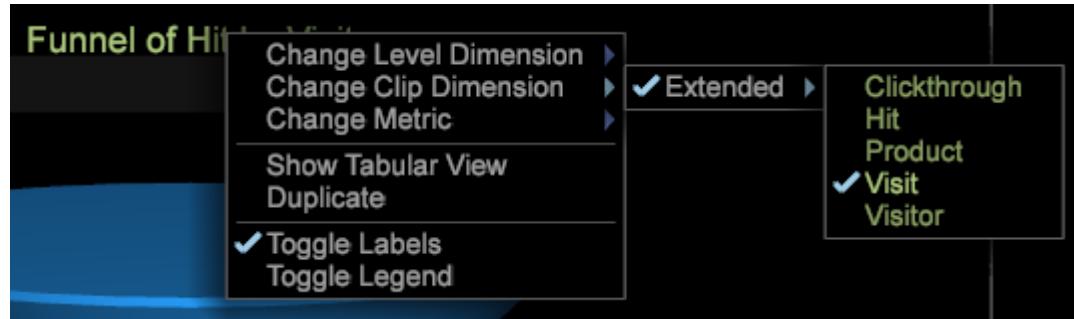
You can add up to a total of eight steps to your funnel.

6. Release the table selection you made to view your funnel based on a total audience.

Funnel visualizations follow the pattern of all other visualizations, allowing the selected element to control the segment of data that you are visualizing.

7. Modify the Funnel clip dimension, level dimension, and metric.

Right-click on the title to adjust these options based on the metrics and dimensions defined in your custom profile.



From the Funnel, you can identify your fall out and fall through rates and the percentage of customers converting sales and following campaigns or content paths. For additional information, you can open a [Adding a Path Browser](#) on page 154 to see a detailed list of the pages taken before arriving at the selected page, and the paths taken after they left.

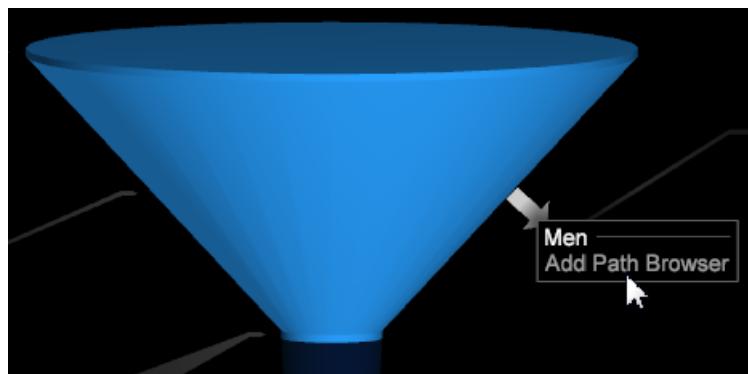
See additional [Funnel Features](#) on page 149.

Adding a Path Browser

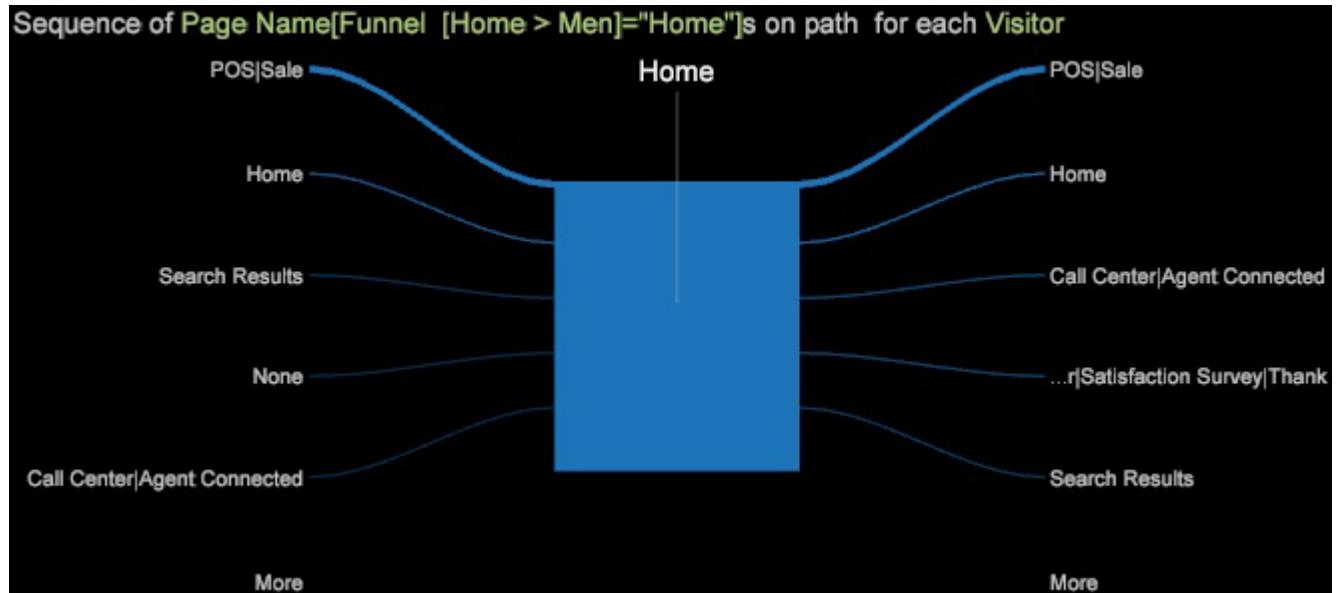
From the Funnel visualization, you can open a Path Browser to identify additional detailed visitor fallout and fall through paths.

You can identify where visitors "fallout" (who left the path) or where visitors are "falling-through" (who followed the path). You can open multiple Path Browsers from the same Funnel visualization by selecting fallout arrows or fall through cones.

1. Open the Path Browser for the fall out visitor, the visitors who left the site during a specific step. Right-click the fallout arrow and select the Add Path Browser option.



From the Path Browser, you can see where visitors went before hitting a selected element in the Funnel (on the left), and where they navigated to after leaving the page (on the right).



In the Path Browser visualization, the width of the lines going in and out of the selected dimension identifies the volume of traffic. For example, before hitting the Men's page, a few visitors came in through other routes, but most came in through the Home page. When leaving, most went to the Men Surfing page.

2. Open the Path Browser for the "fall-through" visitors. Right-click the cone in the funnel to identify the path for visitors who fell through, or moved to the next step, of the funnel.
3. Click the **More** button at the bottom of the Path Browser visualization to open a table with all visits listed in a tabular view for the previous and current visits.

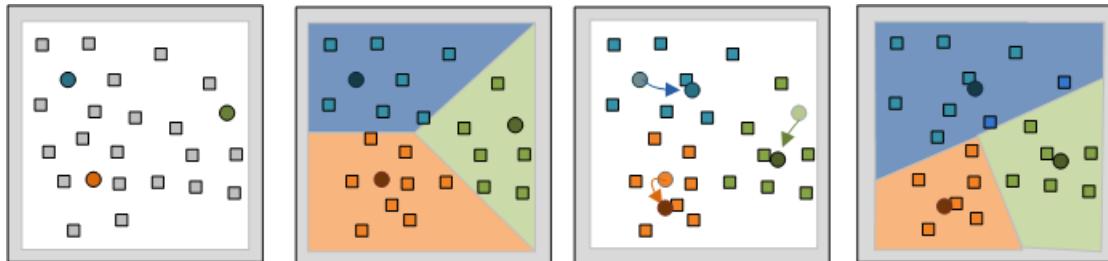
More Page Name[Funnel [Home > Men] = "Home"]s on path	Visits
POS Sale	1,132
Home	253
IVR Product Questions	30
IVR Product Questions General Questions	41
None	73
IVR Product Questions Specific Product Error: Product Not Fo	4
Search Results	74
Product Details Bear	34
Equipment Surfing	30
Product Details Lagos Mini Longboard	20
IVR Welcome	28
Call Center Agent Connected	107
Call Center Satisfaction Survey	31
Men	0
Men Winter	0
Men Winter Skiing	0
Product Details Midweight	0
Men Surfina	5

Visitor Clustering

Visitor clustering lets you leverage customer characteristics to dynamically categorize visitors and generate cluster sets based on selected data inputs, thus identifying groups that have similar interests and behaviors for customer analysis and targeting.

Clustering process

The clustering process requires you to identify metrics and dimension elements to use as inputs, and allows you to choose a specific target population to apply these elements to create specified clusters. When you run the clustering process, the system uses the metric and dimension inputs to determine appropriate initial centers for the specified number of clusters. These centers are then used as a starting point to apply the K-Means algorithm.



The initial centers are intelligently chosen via a Canopy Clustering pass.

Data clusters are created by associating every data point to the nearest center.

The mean of each of the K clusters becomes the new center.

The algorithm is repeated in steps 2 and 3 until convergence is reached. This can take multiples passes.

The **Maximum Iterations** in the **Options** menu allows the analyst to specify the maximum number of iterations to be performed by the clustering algorithm. Setting this option may result in faster completion of the clustering process based on the maximum iterations cap at the expense of exact convergence of the cluster centers.



Note: Once the clusters have been defined, the Cluster Dimension can be saved for use just like any other dimension. It can also be loaded into the Cluster Explorer to examine the separation of cluster centers.

In the **Cluster Builder**, you can select **Options > Algorithm** to select algorithms when defining clusters. Currently, there are 3 supported algorithms:

- KMeans
- Kmeans++
- Expectation Maximization

There are 2 ways to run the clustering process:

- Method 1 - Click **Go** in the cluster visualization window.
- Method 2 - Click **Submit** in the cluster visualization window, which directly sends the clustering job to the Server. You can track the progress through the "Detailed Status for Query" option.

The screenshot shows a software interface for monitoring a Predictive Analytics Server. At the top, there are two tabs: "New Layout" (selected) and "Original Layout". Below them is a search bar labeled "Search:" with a lock icon. The main area displays a hierarchical tree view and a list of server details.

- Detailed Status for Query Server** (selected)
 - + Component Status map
 - + Disk Space
 - + Execution Engine
 - + Hardware
 - + Memory Status ServerMemoryStatus
 - + Network Connections 5 items
 - + Performance
 - + Poll Latency
- Predictive Analytics Server** (highlighted with a red box)
 - [redacted] 0 Tasks
- + Processing Status
 - + Profiles 1 items
 - Location E:\server\
 - Product Insight Server
 - Query Memory Usage 20,971,548/500,000,000 bytes
 - Server Common Name [redacted]
 - Service Status running
 - Status Generated At 04/21/17 00:57:24
 - Status Generation Took (s) 0.075087826699018478
 - Version 6.61 (16101405)

The algorithm has the following restrictions:

1. If you are using Method 1, you can select any of the supported clustering algorithms.
2. If you are using Method 2, you can select kmeans or kmeans++. The Expectation Maximization option will not available.



Note: In the DPU.cfg file, the value for 'Query, Memory Limit' is set to 500 MB by default. This value must be increased while running multiple clustering jobs. For instance, if you are running 5 clustering jobs in parallel, increase this value to 1 GB. There is no way to cancel the clustering job without restarting the Server.

Recommendations

The number of iterations (number of times the data is scanned) and the convergence threshold that you configure, grossly affects the clustering performance. The following table provides a broader guideline that you can follow:

Number of Clusters	Algorithm	Iterations	Convergence Threshold	Normalization
6	Kmeans	25,50	1e-3	Min-Max

Number of Clusters	Algorithm	Iterations	Convergence Threshold	Normalization
6	Kmeans	25,50	1e-6	Min-Max
6	Kmeans++	50	1e-6	Min-Max

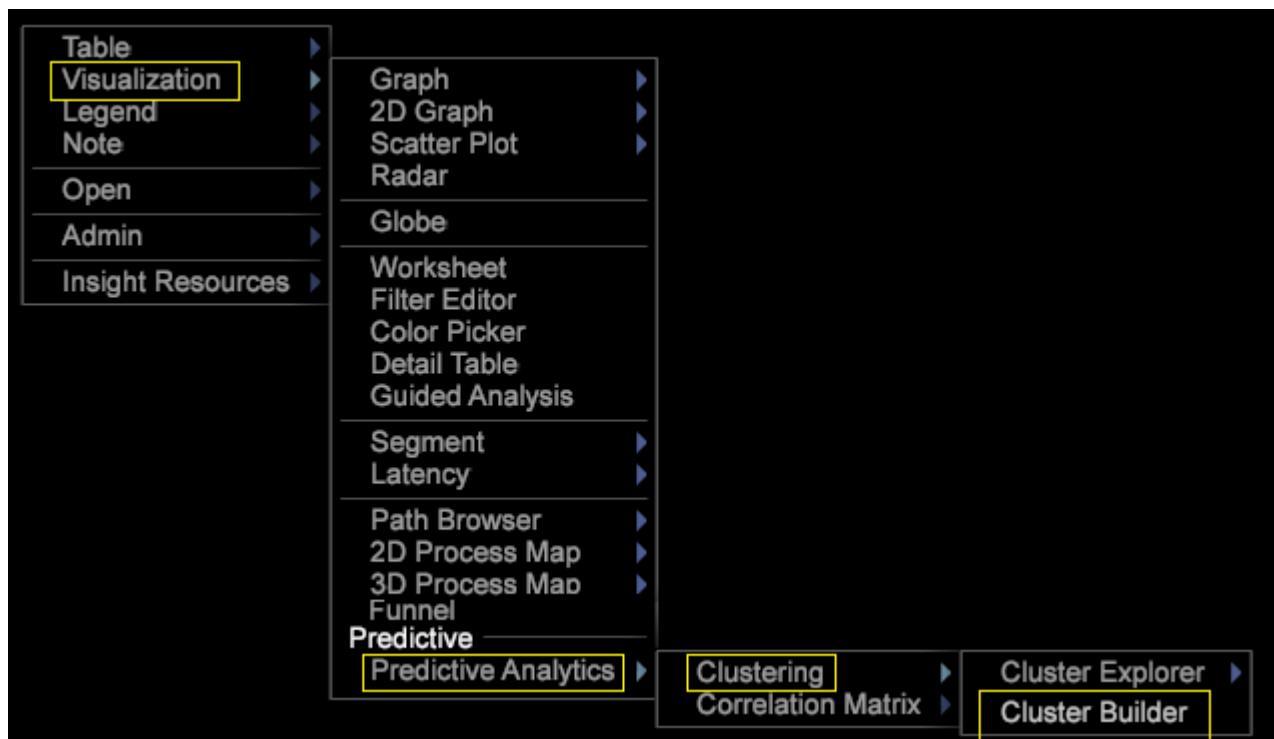
Building Clusters

Select input variables, the number of clusters, and a target population (if desired) to define clusters in your dataset.

Building Clusters

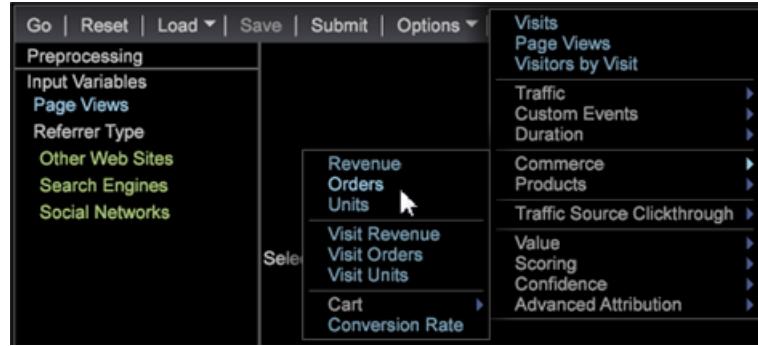
1. Open the **Cluster Builder**.

Click **Visualization > Predictive Analytics > Clustering > Cluster Builder**.



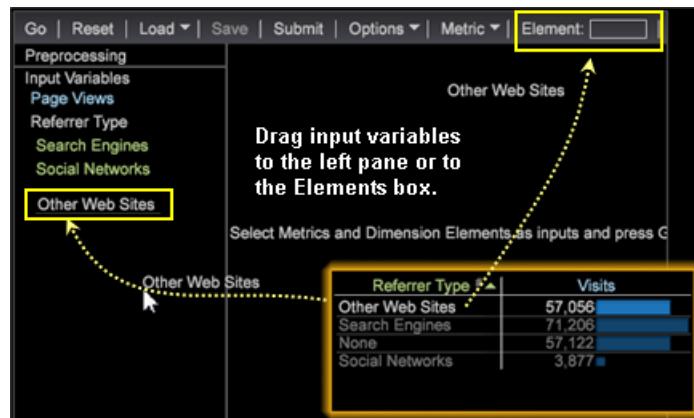
2. Select input variables.

- Add metrics to the **Input Variables** list by selecting from the **Metric** menu in the toolbar.



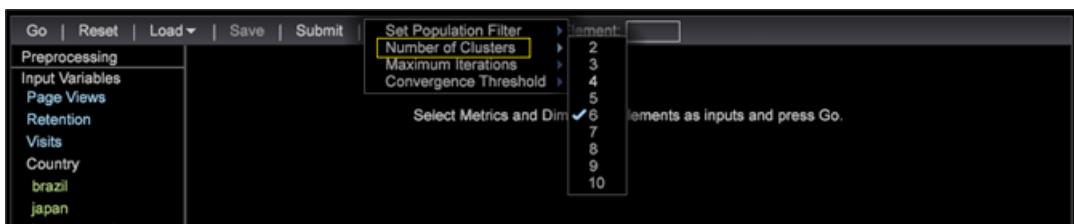
- Add dimension elements to the **Input Variables** list by dragging them from a Dimension's table.

Press **Ctrl + Alt** and drag selected dimension elements to the **Input Variables** list or to the **Element** box in the toolbar.



By default, clustering is performed on the entire dataset. You can see all input variables in the left **Preprocessing** pane.

3. Use the **Options** menu to select the desired number of clusters.



4. If you want to cluster a subset of the Visitors in your dataset, you can define a Population Filter.

The screenshot shows the SAS Data Studio Cluster Builder interface. The top menu bar includes 'Go', 'Reset', 'Load', 'Save', 'Submit', and 'Options'. The 'Options' menu is open, with 'Set Population Filter' highlighted. A tooltip for 'Visitors with at least 1 Purchased Product' is overlaid on a dimension element distribution bar. The main workspace displays a table of visitors by day of the week.

Day of Week	Visits
Sun	8,810
Sat	8,710
Fri	10,658
Thu	10,567
Total	37,745

Start by defining the desired subset using selections in your Workspace or by using the **Filter Editor**. Once you have the desired subset selected, set the Target Population in the **Options** menu. It is recommended that you give the targeted group an identifying name.

The **Options** menu also has settings to control the maximum number of passes and the acceptable threshold for center convergence.

- After inputs and options have been configured, click the **Go** button to run the clustering locally or press **Submit** to send the task to the Predictive Analytics Server. Submissions to the server will save the resulting dimension to the dataset when convergence is complete.

When running locally, you will see the Cluster Builder move through four canopy clustering stages as it defines intelligent centers based on the inputs.

Once the centers of the clusters stop changing more than the specified convergence threshold, the Cluster Dimension is converged and the Cluster Builder displays additional information about how relevant an input was to each cluster.

- Customize the clusters.

Right-clicking on the statistics' color bar opens a context menu allowing you to customize the relevance thresholds, and in the case of the dimension element distributions, to choose which test is displayed.

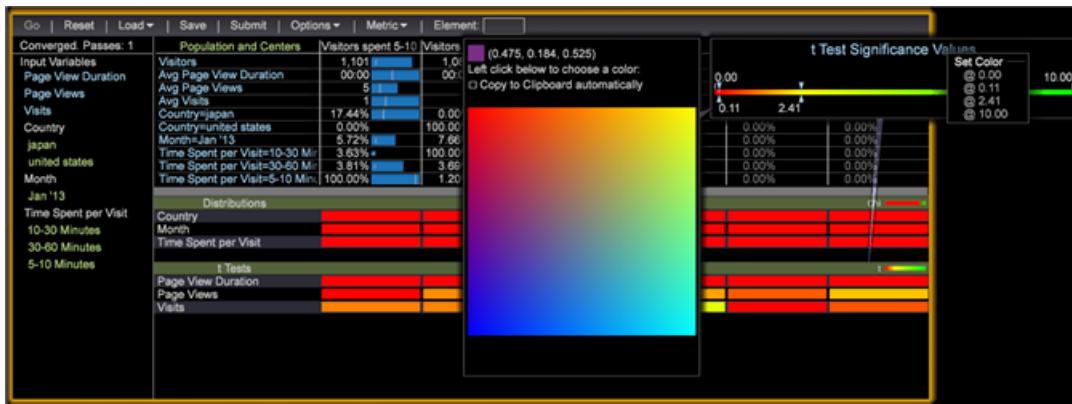


Metric inputs provide a t-test for each cluster, while dimension element inputs provide three distribution tests (Chi squared, an entropy U statistic, and Cramer's V statistic) for each cluster.

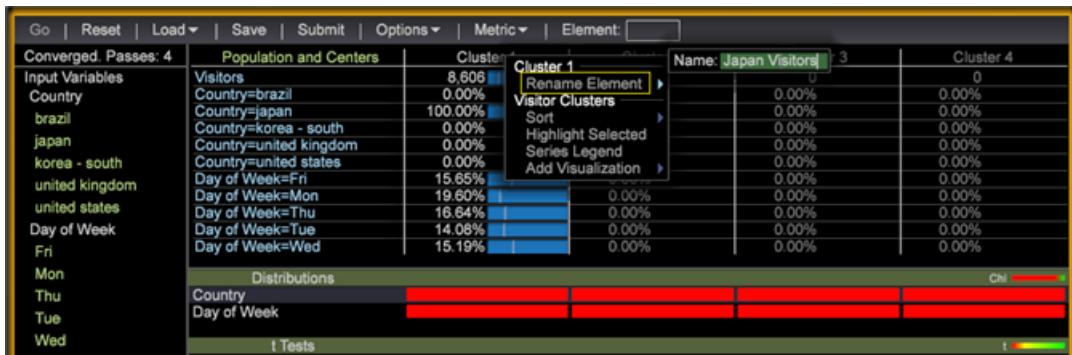


Note: If you add or remove inputs during convergence, the process will pause until you press **Go** again.

After building clusters, you can open the color picker to assign colors for different distribution results.



- With the Cluster Dimension converged, you can add metrics to the table and make selections as normal. You can also right click on the element names (Cluster 1, Cluster 2, etc) to open the context menu to rename them to something more meaningful.



- If you wish to use this cluster dimension in other visualizations, you can **Save** it locally or **Submit** it to the server.

If you wish to run convergence again or see the relevance of inputs, Cluster Builder can also load existing cluster dimensions.

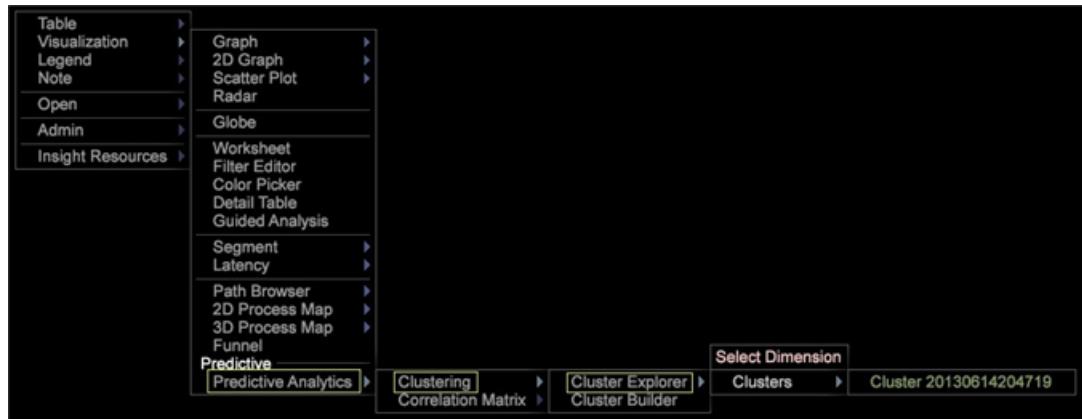


Tip: When selected, **Reset** will completely release all the input variables and give you a blank cluster builder visualization to define new clusters.

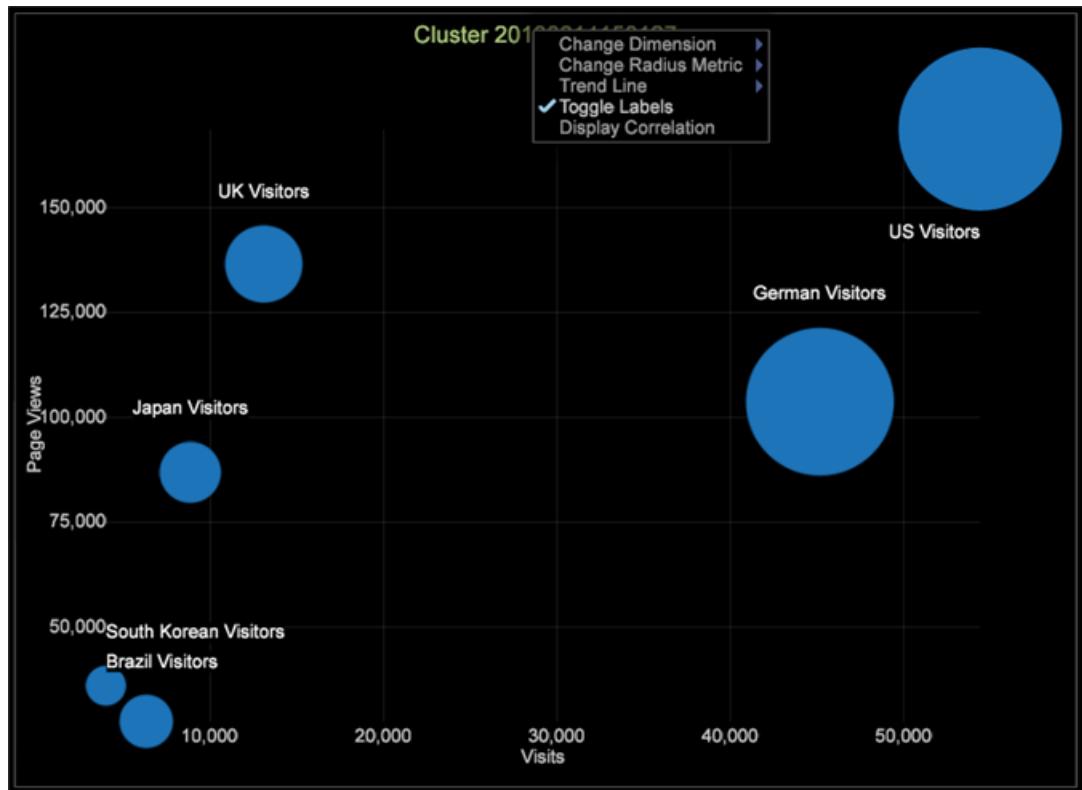
Exploring Clusters

After a Cluster Dimension is built, you can save it and examine it using the Cluster Explorer.

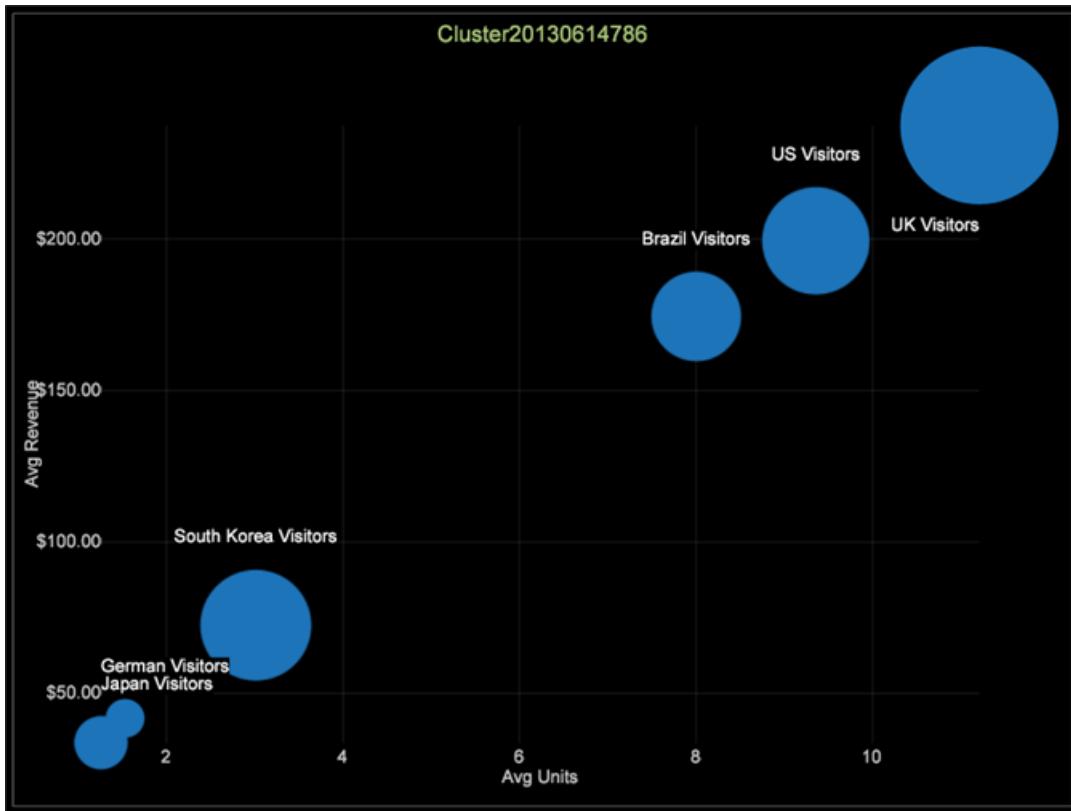
- Select **Visualization > Predictive Analytics > Clustering > Cluster Explorer > Clusters**. Choose the saved cluster dimension that you want to explore.



- Right-click the visualization title and click Toggle Labels to turn them on. You can use this menu to change the cluster dimension you are exploring, change the radius metric, add trend lines, and display correlation.

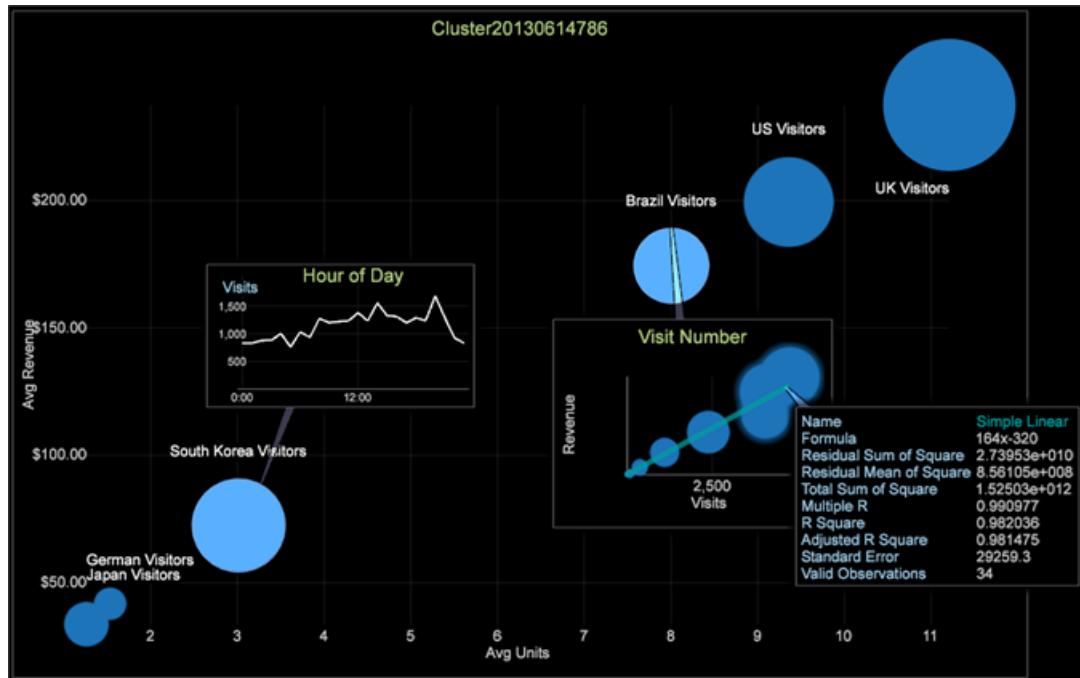


- To modify the Cluster Dimension you are exploring, change the radius metric, add trend lines, and display the correlation using this menu.



The Cluster Explorer displays the centers of the clusters along any two input axes at a time. This allows you to examine their separation in multidimensional space.

4. Right click the cluster to display the menu and choose from callout types: Image or Text Annotation, Metric Legend, Table, Line Graph or Scatter Plot.



Correlation Matrix

Statistical correlations measure meaningful relationships to identify opportunities through advanced data mining.

Employing the [About the Pearson Correlation Coefficient](#) on page 168, the Correlation Matrix furnishes you with relevant information to better identify the next steps in a marketing campaign, to improve site design, or to continue in-depth customer analysis for additional correlation dependencies.

Building a Correlation Matrix

The Correlation Matrix compares metrics over a countable or non-countable dimension. The matrix can then be modified to highlight correlations within the visualization through color picking or to render it as a text map, heat map, or both.

1. Open a Correlation Matrix.

Right-click **Visualization > Predictive Analytics > Correlation Matrix**. The dimension table will open.



Select a dimension, such as Time > Day of the Week from this menu. The correlation table will open with the dimension identified in the corner of the matrix and its associated metric placed in the row and column. For the Day of the Week dimension, **Visits** is the associated metric.



The correlation is 1.000 because you are comparing a metric against itself (which reflects a perfect, but unusable, correlation.)

2. Change one of the metrics.

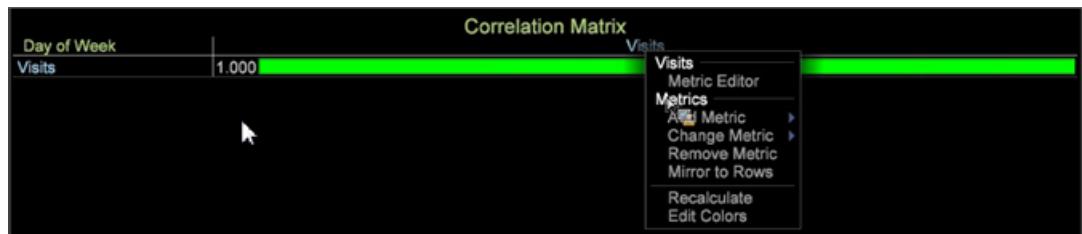
Right-click and select **Change Metric** to change a metric in either the row or column. This sets up a correlation between two metrics of value.

For this example, change the **Visits** metric in the column to **Internal Searches**. Right-click and select **Metric** > **Custom Events** > **Custom Event 1-10** > **Internal Searches**.

Correlation Matrix	
Day of Week	Internal Searches
Visits	0.828

3. Add more metrics to the Correlation Matrix.

Right-click in a metric column or row. For example, from the Metric menu, add Metric > Custom Events > Custom Event 1-10 > Sign in Error.



The screenshot shows the Correlation Matrix with two rows: 'Day of Week' and 'Visits'. The 'Visits' row has a context menu open, listing options like 'Metric Editor', 'Metrics', 'Add Metric', 'Change Metric', 'Remove Metric', 'Mirror to Rows', 'Recalculate', and 'Edit Colors'. The 'Add Metric' option is highlighted.

Day of Week	Visits
Visits	1.000

The new metric will appear in a column with a correlation number. You can add other metrics, such as **Email Signups**, to build out the table.



The screenshot shows the Correlation Matrix with three columns: 'Day of Week', 'Visits', and 'SignIn Errors'. The 'Visits' row has a correlation value of 1.000, and the 'SignIn Errors' row has a correlation value of 0.681. The 'SignIn Errors' column is currently selected.

Day of Week	Visits	SignIn Errors
Visits	1.000	0.681

Or add metrics to rows to compare against metrics in columns.

Correlation Matrix			
Day of Week	Internal Searches	SignIn Errors	Email Signups
Visits	0.795	0.815	0.769
Visitors by Visit	0.865	0.889	0.869
Visitors by Visit			
Metric Editor			
Metric Legend			
Edit Metric Details			
Metrics			
Add Metric			
Change Metric			
Label			
Remove Metric			
Mirror to Columns			
Recalculate			
Edit Colors			

4. (optional) Constrain a metric by adding a dimension element.

Right-click in the workspace and select **Table**. From the open dimension table, press Ctrl + Alt and drag the element over a metric in a column or row. The element will appear next to the metric in brackets.

For example, for the **Visits** metric, you can constrain it by selecting the **Country** as **New Zealand**.

Correlation Matrix			
Day of Week	Internal Searches	SignIn Errors	Email Signups
Visits[Country="new zealand"]	0.599	0.542	0.283
Visitors by Visit	0.606	0.587	0.334
Country # ▲ Visits			
guatemala		32	
new zealand		330	
ecuador		191	
croatia (local name - hrvatsko)		81	
thailand		363	
finland		138	
p puerto rico		182	
i indonesia		124	
philippines		150	

Notice that when you select a dimension element, the correlation changes in all metrics based on the selected dimension element. Only the Visit metric will be constrained for "New Zealand" once the dimension window is closed.



Note: If changing a metric with a dimension constraint (by right-clicking and selecting **Change Metric**), the dimension element constraining the metric will be lost. You will need to add the dimension element again.

5. Create a *Binary Filter in the Correlation Matrix* on page 169 to further constrain the metric. Right-click the metric in the table and select Binary Filter from the menu.

Correlation Planning and Analysis Goals

The following are general goals for building a Correlation Matrix.

Identify the relationship between two metrics against a specified dimension. In the example, the matrix was built around the core dimension, Day of the Week, with the metrics Visit, Email Signups, and SignIn Errors compared against Internal Searches, Login, and Survey Displayed metric events.

Develop hypotheses to focus analysis. After running a correlation analysis, your next step is to look for dependencies and correlation of the metrics. For example, understanding that internal searches has an effect on email sign-ups provides a path to predict that relationship and to modify marketing campaigns or web site navigation design.

Identify metrics to include more advanced data mining algorithms. In most cases, the key metrics will be identified because they will be seen affecting multiple correlations. You can now take those key metrics and apply them to additional data mining analysis for deeper insight.

Correlation Matrix Feature Notes

Filtering and selecting on dimension elements within a table compares like values. For example, using Day of the Week dimension and then clicking into an element of that core dimension, such as clicking on a specific day within the Day of Week dimension table, renders a one to one match at 100% that provides no usable correlation. Because the root dimension was Day of the Week, any selection within the Day of the Week dimension table will alter the matrix to be a one-to-one correlation.

Day of Week	Visits
Mon	29,273
Tue	22,477
Wed	23,200
Thu	30,602
Fri	29,066
Sat	24,400
Sun	24,837

Day of Week	Correlation Matrix		
	Visits	Email Signups	SignIn Errors
Internal Searches	1.000	1.000	1.000
Login	1.000	1.000	1.000
Survey Displayed	1.000	1.000	1.000

However, the 1 to 1 correlation (when a single selection is made of all elements) is only on that specific day. If you make multiple selections then it does not necessarily remain a 1 to 1 correlation, and will not always yield a 100 percent match regardless of selecting 1 or 1+ days of the week.

Statistical correlations are not equal to the Correlated Data Model, the historical reference of Adobe Analytics products. The statistical correlation in data workbench is based on the [About the Pearson Correlation Coefficient](#) on page 168.

Display Correlation in a Scatter Plot. Right click the title on a Scatter Plot and choose **Display Correlation** from the **Visualization** menu. The Correlation value will display in the upper right section of the Scatter Plot.



Note: The Scatter Plot and Pearson's matrix will display "Calculation Error" if the application is unable to run the Pearson's correlation calculation. This is usually due to insufficient data, which can cause the equation to attempt to divide by 0.

About the Pearson Correlation Coefficient

The Pearson Correlation Coefficient is used in the Correlation Matrix as the algorithm to display the strength of the linear dependence between two variables.

This linear correlation is a statistical measure of the linear dependence, or correlation, between two variables to render a value between +1 and -1 inclusive, representing either a positive or negative dependence.

Here is the Pearson Correlation Coefficient

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

The Pearson's value is visualized in the Correlation Matrix, which depicts the correlation between two defined metrics. These metrics can be compared to one another over any countable or non-countable dimension in the dataset.

You can highlight these comparisons through contrasting colors using the color picker, or by comparing values in a text map and heat map, or both.

Correlation Analysis Use Cases

Defining dependent correlation points that make sense in your market is the essence of correlation analysis.

These use cases highlight the art of identifying relationships as correlation points applied to the science of the [About the Pearson Correlation Coefficient](#) on page 168.

Social Media and Website Correlation

Digital publishers want to maximize their understanding of the potential relationship between social media activity and visits to their website. For example, the digital publisher runs the correlation report between hourly Twitter mentions and visits for a two week period. The correlation is found to be $r = 0.28$, which indicates a medium, positive relationship between Twitter mentions and website visits.

Optimization for E-retailers

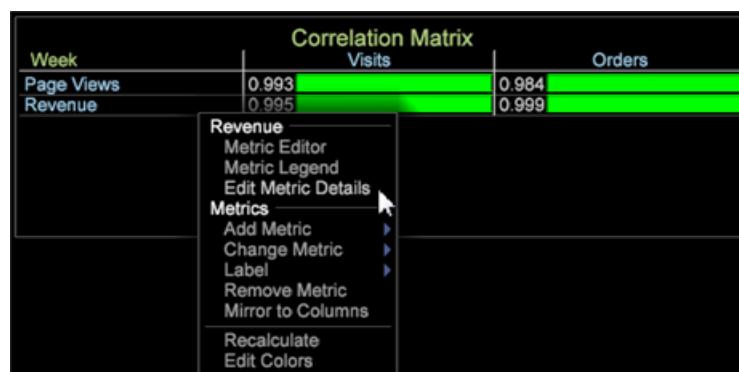
E-retailers are interested in driving increased revenue. For example, an e-retailer wants to compare a number of secondary success events (e.g., file downloads, product detail page views, internal search click-throughs, etc.) with weekly web revenue. They can quickly identify internal search click-throughs as having the highest correlation ($r = 0.46$), which may indicate an area for optimization.

Binary Filter in the Correlation Matrix

A **Binary Filter** in the Correlation Matrix lets you constrain values for one or both of the correlated metrics to better focus the comparison.

To set a binary filter on a Correlation Matrix:

1. From the Correlation Matrix, right-click a metric name.
2. Select **Edit Metric Details**.



The **Edit Correlation Metric Details** window will open.

Week	Correlation Matrix	
	Visits	Orders
Page Views	0.993	0.984
Revenue	0.995	0.999

Edit Correlation Metric Details
 Metric: Revenue
 Label:
 Binary Filter: Inactive
 Comparison: >
 Value: 0.000
 Element: No Element

3. Set up a Binary Filter.

First, click the **Inactive** setting. It will toggle to set the filter as **Active** and display the **Comparison** and **Value** fields.

Then, select a **Comparison** operator and set its **Value** to set up a filter for the selected metric.



Important: The Binary Filter for Data Workbench 6.2 has been updated with new features, requiring you to rebuild any correlation matrix with a binary filter built in previous versions.

Adding Dimension Elements

You can also add a dimension element to constrain a metric. A metric can have only one element associated with it.

Week	Correlation Matrix	
	Visits	Orders
Page Views	1.000	1.000
Revenue[Week='01/21/13']	1.000	1.000

Edit Correlation Metric Details
 Metric: Revenue[Week="01/21/13"]
 Label:
 Binary Filter: Inactive
 Comparison: >
 Value: 0.000
 Element: Week="01/21/13"

Week	Visits
01/21/13	13,782
01/28/13	40,825
02/04/13	42,010
02/11/13	35,254

Right-click in the workspace and select **Table**. Open a dimension with its elements and drag to the **Element** setting in the Edit Correlation Metric Details window, or drop on a metric in the Correlation Matrix.

Data Workbench 6.0.4 Release Notes

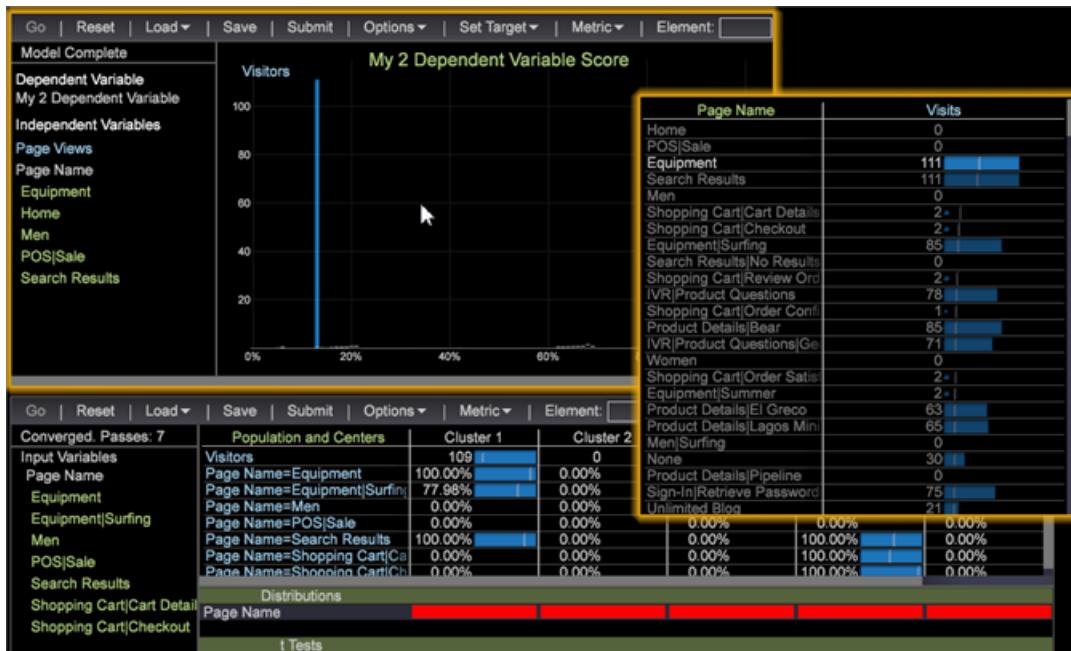
New features introduced in Data Workbench 6.0.4, including bug fixes and known issues.

To view previous features and fixes based for each past release, see the [release note archives](#).

New Features

Data Workbench 6.0.4 includes these new features and visualizations for added reporting capabilities and predictive analysis tools.

Propensity Scoring visualization. Data workbench calculates scores for each visitor as an estimated probability that a specified event may happen. The Visitor Scoring visualization allows you to create a score dimension that gives a probability of a specified event for every visitor of interest based on the input variables.



See [Propensity Scoring](#) on page 124 for additional information about this feature.

Upgrade Requirements

Log Source ID must be defined. Starting in version 6.04, if the Log Source ID is not defined then you will get the following error:

Missing Log Souce ID in log processing.cfg. Log Source ID must be defined for all log sources.

The Recording of Rows per Log Source was added in Data Workbench 6.0 and can be defined in the custom profile Log Processing.cfg by adding a uniquely named Log Source ID. If you have a blank Log Source ID, then you could see Log Processing issues such as incomplete reading of the log source data and other discrepancies.

```
Log Processing.cfg
Log Sources = vector: 2 items
0 = VisualSensor:
    Compressed = bool: false
    Log Paths = vector: 1 items
        0 = Path: \some path\
    Log Server = serverInfo:
        Address = string:
        Name = string:
        Port = int: 80
        Proxy Address = string:
        Proxy Password = string:
        Proxy Port = int: 8080
```

```
Proxy User Name = string:  
SSL Client Certificate = string: Certificates\\server_cert.pem  
SSL Server Common Name = string:  
Use SSL = bool: false  
Log Source ID = string: <Name your ID Here>  
Name = string:  
Recursive = bool: false
```

Ability to Delegate FSU Resources

In `Profiles/<profilefilename>/dataset/Cluster.cfg`, you can now specify separate File Server Units (FSU) for the Normalize and Source List servers. These services are no longer tied to the Master FSU.



Note: If the List Server is not specified, then the List Server will inherit the Normalize Server's configuration settings.

Example in the `cluster.cfg` file.

```
Cluster = ClusterConfig:  
    Normalize Server = serverInfo:  
        Address = string: normalizeserver.domain.com  
        Port = int: 80  
        Use SSL = bool: false  
    List Server = serverInfo:  
        Address = string: sourcelistserver.domain.com  
        Port = int: 80  
        Use SSL = bool: false
```

Fixed Bugs

- In Data Workbench 6.0, the Correlation Matrix and Cluster Builder did not support Compute in Background. This is now fixed in version 6.0.4.
- Previously, if you had a selection on the Funnel and removed a step, an access violation could occur. This has been resolved.
- Fixed a potential locking condition in Segment Export that may cause problems under heavy load conditions.

Installing Data Workbench

Navigation title:

This chapter describes how to install the Data Workbench software and configure the connection to the Server.

Workstation requirements

The following are requirements and recommendations for installing the Workstation (Client) in Data Workbench.

See [server system requirements](#) for additional Data Workbench system requirements.



Important: The server, report server, and client components are upgraded to run on 64-bit Windows operating systems.

Before You Begin

Make sure you have the these tasks completed before installing the Data Workbench Workstation (Client):

- Add **Excluded Processes** for *MS System Center Endpoint Protection in Windows 2012 Servers* for the following executables:
 - **InsightServer64.exe**

-
- **ReportServer.exe**
 - **ExportIntegration.exe**

This will allow "white list" rights for these interfacing executables.

- **Install Microsoft Excel to export analysis data.** To export data from workspaces as Microsoft Excel (.xls or .xlsx) files, the computer on which you install Data Workbench must have Excel installed and registered. If Excel has not been registered and Data Workbench tries to access it for the first time, Excel displays a registration dialog box. If you are not sure whether the copy is registered, start Excel manually, and if a registration dialog box appears, complete the registration process.



Note: With the release of Data Workbench 6.4, support for Excel 2007 has been discontinued. Also, because Data Workbench only runs on Microsoft Windows for 64-bit architecture, it is recommended that you also install a 64-bit version of Microsoft Excel.

- **Installing Adobe Acrobat for printing scaled workspaces to PDF.** To print scaled workspaces to Adobe PDF format, the computer on which you installed Data Workbench must have Adobe Acrobat installed.
- **Providing access to a printer for printing workspaces.** To print workspaces from Data Workbench, the computer on which you install Data Workbench must have access to a printer. Data Workbench can print workspaces to color or monochrome printers and does not require postscript or other advanced printer features. For optimal results, Adobe recommends printing workspaces in color.
- **Implement security measures.** You should follow your company's normal enterprise security policies for Data Workbench computers. To serve its primary purposes, Data Workbench requires only the ability to connect to a server (via ports 80 and 443) and to any servers collecting data. You can use the Data Workbench hardware for any other purpose as long you maintain the Data Workbench software and allocate at least 10 GB of storage space for Data Workbench.
- To render visualizations accurately, the computer on which you install the workbench must have an appropriate **graphics adapter** installed (see Graphic Adapter requirements below).

Data Workbench Client Requirements

Operating System

- Microsoft Windows 7 64-bit
- Microsoft Windows 8.1 64-bit
- Microsoft Windows 10 64 bit



Note: Windows XP is not supported for Data Workbench 6.1 and later versions.

Resolution

- Required: 1024 x 768 (XGA)
- Recommended: 1920 x 1200 (WUXGA)

Graphics Adapter

- Required: OpenGL hardware acceleration to support OpenGL 3.2
- Recommended: Dedicated video adapter (e.g., NVIDIA or ATI adapter)

Processor

- Required: 1.2 GHz or higher Intel or AMD
- Recommended: ICore 2 Duo-Class

RAM

- Required: 2 GB
- Recommended: 4 GB

Connectivity

- Required: 512 Kbps link to the DPU
- Recommended: 2Mbps or faster link to the DPU

File System

NTFS

Disk Storage

At least ten (10) GB or greater of free hard disk drive space

Printing

Printer access (color or gray scale printers) for printing workspaces and reports

Other

- Dedicated mouse
- Low-glare working environment
- Matte-surfaced monitor

Workstation setup

Workstation Setup Wizard

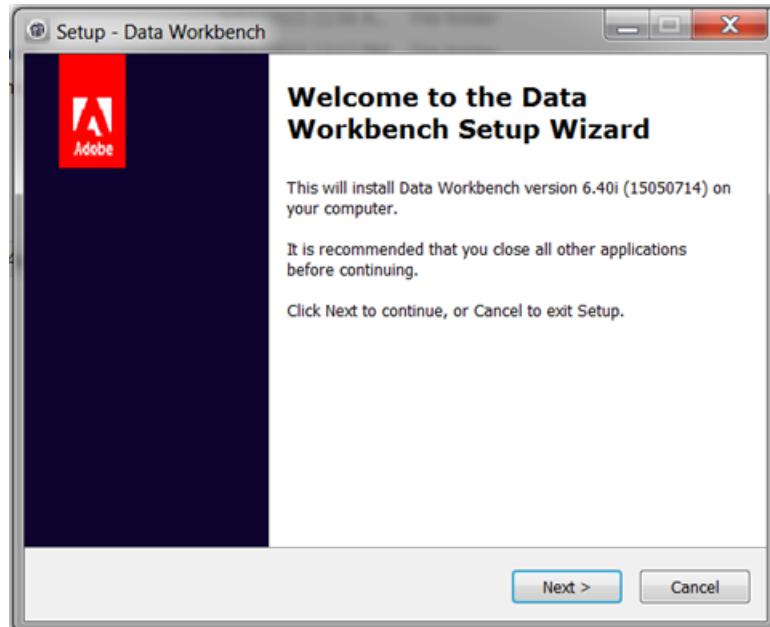
Data Workbench provides a set up wizard to install the workstation (client) application.

Installing the Workstation using the Setup Wizard

Launch the installation wizard executable and walk through each step to install the workstation client program. After installation of the workstation, you can connect to servers and profiles.

1. Double-click the workstation installer executable.
2. Click **Yes** to allow the program to install on Windows.
3. Select a **Language** for the setup wizard.

The wizard will open:



4. Click **Next** on the **Welcome to Data Workbench Setup Wizard** dialog.
5. Select to install a **New Installation** or to **Upgrade or repair** an existing installation.

New Installation overwrites any previously installed files.

Upgrade updates your Workstation to the latest version or lets you repair an existing installation. Data Workbench will compare installed **Insight.exe** files and run the Workstation Setup Wizard if a newer version of the client is available.

6. Select install location:

Typical installs to a default folder and location.

- Program files are saved by default to:

```
C:\Program Files\Adobe\Adobe Analytics\Data Workbench
```

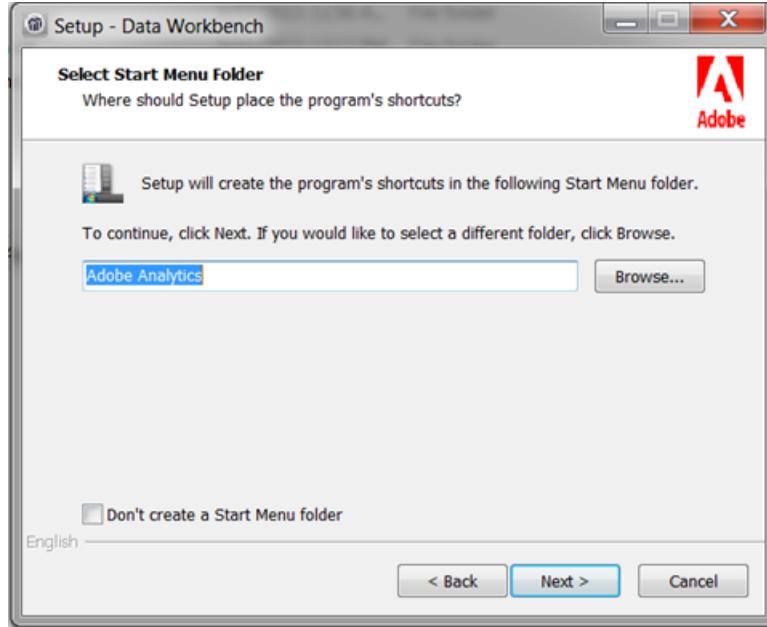
- Data Files (profiles, certificates, trace logs, and user files) are saved by default to:

```
C:\Users\<username>\AppData\Local\Adobe\Adobe Analytics\Data Workbench\
```



Important: A generic **Insight.cfg** file without server details will be installed initially. It is recommended that you use the newly installed **Insight.cfg** file and customize it rather than moving a file from a previous installation. Because the path for installing the workstation has changed, the addition of fonts, removal of the *User Folder*, and the removal of the *TraceFileComponent* is recommended.

7. (optional) Select **Custom** to choose the language package and the location of the program and data files.
8. Select location for **shortcuts in the Start Menu**.



Click **Don't create a Start Menu folder** to not install a shortcut on the Windows Start Menu.

9. Click **Next**. A summary of selected file location paths and languages will display. Click **Install**.
10. Locate the **Data Workbench Certificate**.

If the setup wizard cannot find the Data Workbench certificate during installation, it will open a dialog to browse to the location of the certificate (a **.pem** file located by default in the client **Certificates** folder), or click **Skip** to find the certificate after installation.

Click **Install** after locating the certificate.

11. After the setup wizard is complete and Data Workbench installed, click **Finish** to complete setup.

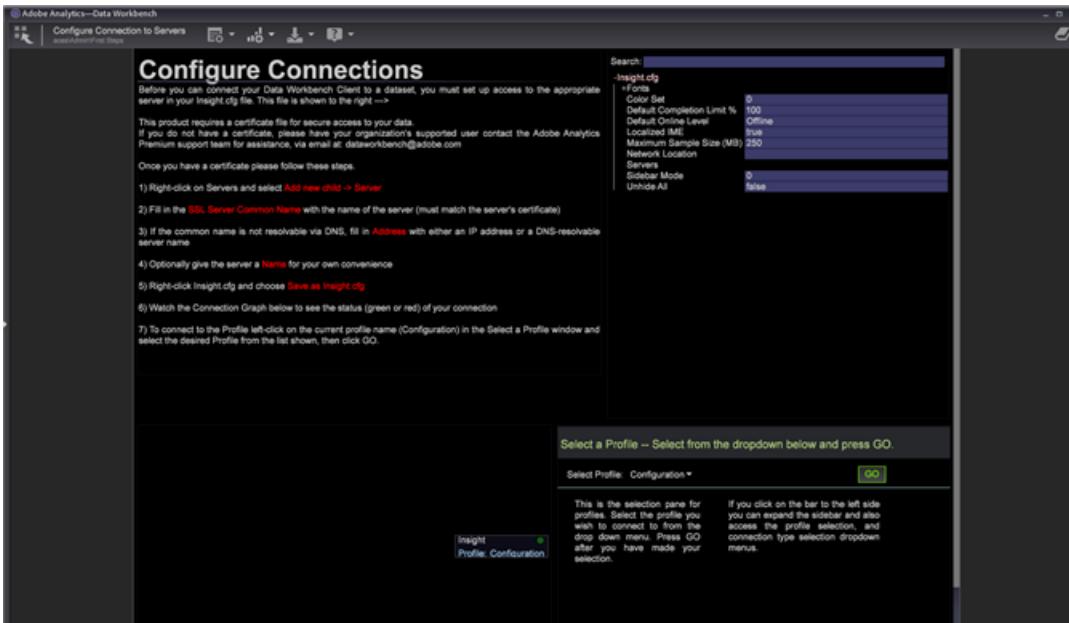


Note: The default log location for the Workstation Set up Wizard at
C:\Users\<userName>\AppData\Local\Temp.

Select the **Launch application** checkbox to open the workbench after setup.

12. **Configure connections** to servers in **Insight.cfg** file.

After installation of the workstation, the Enhanced Workstation Configuration Experience workspace will open with additional information about [entering server connection information](#) in the *Insight.cfg* file and an option to select a profile from the drop-down. You can also view the connection status to your servers.



Installation Folders

The Data Workbench folder structure has two installation locations:

- **Program Files**

The **Insight.exe** and supporting client files (**Insight.ini**) are now located by default at
C:\Program Files\Adobe\Analytics\DataWorkbench

- The **Appdata** folder.

Insight.cfg, profiles, certificates, trace logs, and user files are now located by default at
C:\Users\<Winuser>\AppData\Adobe\Analytics\DataWorkbench

You can set the path for the **Appdata** folder in the **Insight.ini** file:

```
[InitialSettings]
AppDataFolder=C:\Users\mhiatt\AppData\Local\Adobe\Analytics\Analytics\Local
Locale=en-us
```

Uninstalling the Workstation

Data Workbench now includes an executable to uninstall the workstation (located by default at **Program Files\Adobe\Adobe Analytics\Analytics\unins000.exe**).

Launch and follow steps to remove the Data Workbench Workstation files from your hard drive.



Note: You can launch the **unins000.exe** executable from the folder, using the **Uninstall Data Workbench** shortcut from Start Menu, or from **Control Panel > Program and Features**.

Files Included in the Installation Package

Navigation title: Files Included in the Installation Package

Files included in the Data Workbench installation package.

The following directories are included in the Insight package.

Program Files

The Workstation executable (**Insight.exe**) and supporting files are installed by default to this folder.

C:\Program Files\Adobe\Adobe Analytics\Data Workbench

Directory	Description
Insight.exe	The Data Workbench client executable.
insight.ini	Set language and the path for the Appdata folder.
Insight.zbin	Data workbench now supports an Input Method Editor (IME) as a secondary text entry process that allows you to enter international characters from your keyboard using a floating text box. Data workbench will support English by default but also allows you to load other files to support international languages, such as a virtual Chinese keyboard (Pinyin IME). A new dictionary file (Insight.zbin) is required by the client application to support the IME.
unins000.exe	Executable to uninstall Workstation and delete files.

AppData Files

Data Files (**Insight.cfg**, Profiles, Certificates, trace logs, and user files) are saved by default to:

C:\Users\<Winuser>\AppData\Adobe\Analytics\DataWorkbench\

Directory	Description
Insight.cfg	The Data Workbench configuration file. Defines parameters within which Data Workbench operates. See Configuring the Connection to Insight Server on page 181.
Base	Contains the program files for Data Workbench.  Note: You should not remove or alter any of these files.
Certificates	Contains the certificate file, <code>trust_ca_cert.pem</code> , and the named user digital certificate for Data Workbench. See #unique_99 .
Configuration	Contains the files that are used for Workstation configuration.
Geography	Files for graphic rendering of geographical visualizations.
Trace	Log files generated from the Workstation.
Profiles	<i>AdobeSC</i> , <i>Predictive Analytics</i> and other profile configuration files.
InsightSetup.exe	Setup wizard to install additional client computers in the Software/Insight folder.

Installing the Input Method Editor

Data workbench now supports the Input Method Editor (IME) as a secondary text entry process for international languages.

IMEs allow you to enter international characters using a variety of methods suited for your local language. Data workbench provides an input dialog box that allows you to open and use your desired IME for text fields.



Note: For the data workbench 6.1 release, only the virtual Simplified Chinese keyboard will be supported. Inputting other languages through the IME could result in unexpected behavior.

Using an IME

To use the floating IME text input feature:

1. Click **Alt + Space** for any text input area.
2. Enter values using your system's IME.
3. Close the input dialog by selecting the **Enter** key or clicking the **OK** button.

The dialog will disappear and the characters will then appear in the selected field.

Updating the Insight.cfg file

To employ the IME, you must update the `Insight.cfg` file with this setting:

```
Localized IME = bool: true
```

If this setting does not exist in the configuration file, then pressing **Alt + Space** will not engage the IME feature.

Starting Insight in another language: To better support localized assets like a splash screen and to support multiple languages in the future, data workbench requires command-line arguments identifying the language to load. The default language is English.

Starting data workbench in Chinese requires you to invoke `Insight.exe` with the `-zh-cn` argument:

```
Insight.exe -zh-cn
```

(These command line arguments are not case sensitive.)

Installing the Terrain Images.cfg File

Navigation title: Installing the Terrain Images.cfg File

You install the `Terrain Images.cfg` file only if you want to make new terrain image layers to display on the globe visualization.

See [Globes](#).

To install the Terrain Images.cfg file on Insight Server

1. Copy the `Terrain Images.cfg` file from the `.zip` file provided to you by Adobe.
2. Paste the file in the **Components** folder in your Insight Server installation directory.

For information about configuring `Terrain Images.cfg`, see [Terrain image layers](#).

Setting up Localized Languages

Set up `insight.zbin` file to set the language of the client application.

Update the data workbench server components

The administrator must first complete these tasks to update these server components:

1. **Update to data workbench server 6.x.**

You need to update the data workbench server for localization by updating the `base\localization*.zbin` file. This `insight.zbin` file will then be copied to the client.

An `insight.zbin` file is included in the installation folder alongside the `insight.exe` file. If you connect to a server that doesn't provide you with language-specific `.zbin` files, then data workbench will proceed to use this file.

The backup `insight.zbin` file can be provided in any language. As a result, if you use data workbench in Chinese and connect to a server that doesn't support this language, then your data workbench client will still be in Chinese, even if the server changes your base profile and removes your `.zbin` files from the `Base/Localization` folder.

2. Update the data workbench report server.

The `insight.zbin` at the root folder of data workbench report server will be in English by default. As the administrator, you will be required to select and copy the `.zbin` file from the updated report server package and place it in the root directory of the data workbench report server. Like the client, the report server also requires the proper arguments for the selected language, such as `Insight.exe -zh-cn`

- a. Stop the report server services.
- b. Copy the `Localization` folder from the new report server package.
- c. From the `Localization` folder, copy the `Insight.zbin` file and place it in the root directory of the report server where the `Insight.exe` is located.
- d. Add any required arguments, such as `insight.exe -zh-cn`
- e. Restart the report server.

Update the data workbench client

After updating the server, follow these steps to update each client.

1. To make sure the client does not get updated from the server during this update, set your `Insight.cfg` argument to False.

```
Update Software = bool: false
```

2. Restart the client.
3. Navigate to the Software and Docs profile (SoftDocs profile) and download the required `insight.zbin` file from the client package: `Software\Insight Client\Insight_6.1.zip`
4. Move the `insight.zbin` file to the folder where `insight.exe` is located.
5. To make sure that the client files now gets updated from the server, change the `Insight.cfg` file argument to True:

```
Update Software = bool: true
```

I



Note: Your client will synchronize with the server and you will see a message stating that it is updating. At the conclusion of the download, you will get a message asking if you want to restart your client.

6. Click **OK** to restart the client.

If you get the following message, then it means the `zbin` file was not placed in the same location as the `Insight.exe`.

```
Insight Terminated: The backup dictionary file insight.zbin  
is missing.
```

Localized Splash Screens

Data workbench looks for the following splash screen files:

- English (default): `Base/Images/<version_product>_Splash.png`

-
- Chinese (when started with -zh-cn): Base/Images/<version_product>/Splash_zh-cn.png.

If a splash screen is requested but missing, data workbench will access the English splash screen by default.

Downloading and Installing the Digital Certificate

After you have installed the Insight program files, you must download and install the digital certificate provided to you by Adobe.

Configuring the Connection to Insight Server

Navigation title:Configuring the Connection to the Server

After you have installed your Insight software and digital certificate, you must start Insight and configure its connection to Insight Server.



Note: In some cases, the connection to Insight Server may have been pre-configured by Adobe Consulting Services or your system administrator. If so, you do not need to complete this task.

When you start Insight for the first time, it automatically connects to the Adobe License Server to register your digital certificate. To complete the registration process successfully, your computer must be connected to the Internet when you execute the following steps.



Note: If you have already requested, downloaded, and installed a pre-locked certificate as described in *Downloading and Installing the Digital Certificate* on page 181, Insight will not attempt to connect to the License Server and you will not receive an error.

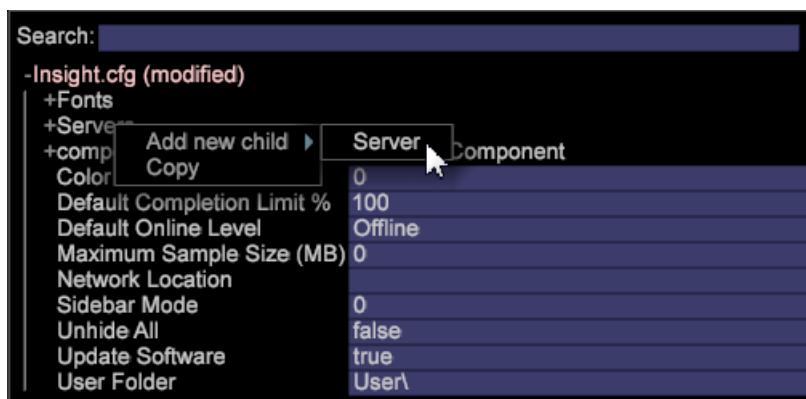
To configure the connection to Insight Server

When working in a clustered environment, Insight should be configured to access the master Insight Server to avoid synchronization issues. In Insight you can view information about the processing Insight Servers in your cluster using the **Related Servers** menu item in the *Servers Manager*.

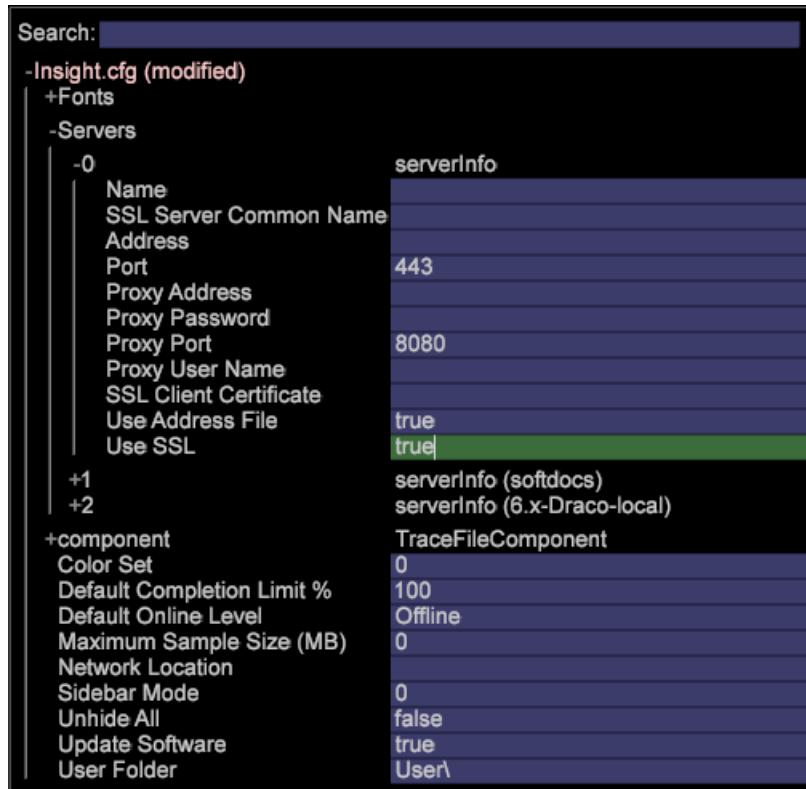
1. Launch Insight.
2. On the **Worktop**, click **Admin**, then **First Steps**.
3. Click the **Configure Connection to Servers** thumbnail.

The **Servers Manager**, the `Insight.cfg` file, and instructions for configuring your `Insight.cfg` file are displayed.

4. In the `Insight.cfg` window, right-click **Servers** and click **Add new child> Server**.

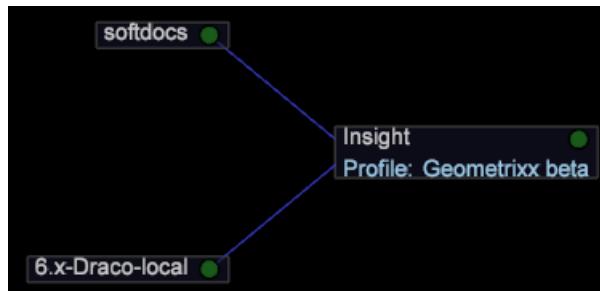


5. Complete or modify the server parameters to provide Insight with access to your master Insight Server. For detailed descriptions of the parameters in the `Insight.cfg` file, see *Configuration parameters*.



6. Repeat Step 4 and Step 5 for each Insight Server to which you want to configure a connection.
7. To save your configuration changes, right-click **Insight.cfg (modified)** at the top of the window and click **Save as Insight.cfg**.

Insight attempts to connect to the Insight Server(s) using the settings that you have specified. If a connection is established, a green node appears in the **Servers Manager** as shown on the following page.



- **Green:** Indicates that the connection to the Insight Server is active.
- **Light Red:** Indicates a potential problem with the server, such as a drain on server processing, high memory usage, or low disk space.
- **Red:** Indicates that the connection to the Insight Server is not active.

If Insight cannot connect using the specified settings, a red node appears in the **Servers Manager**. If this happens, see *[Connection Troubleshooting](#)* on page 185.

Directories and Files Created after Setup

When you start Insight for the first time and configure it to access an Insight Server, the list of profiles available on the Insight Server displays in the **Switch Profile** menu.

When you select a profile to use, the profile information (including related data and any specific workspaces or visualizations defined for the profile) is downloaded to your computer. As you download each profile, Insight creates a folder within the installation directory using the profile name.

For example, if you select a profile named Sales, a folder named Sales appears in your Insight directory. This folder contains the metrics, dimensions, workspaces, and visualizations defined in the Sales profile. After the initial loading of the profile, the profile can be used when working offline. See [Working offline and online](#).

In addition, when you connect to Insight Server for the first time from Insight, Insight Server creates the following directories in the Insight installation directory.

- **Trace directory:** Within the **Trace** directory is the Insight log file (`insight.log`). When the size of the `Insight.log` file reaches 100 MB, the file is renamed to `insight-1.log`. If a file of the name `insight-1.log` already exists, then `insight-1.log` is renamed to `insight-2.log`, and so forth, with a maximum of `insight-9.log`. The file `insight.log` always contains the most recent log information, and `insight-max.log` contains the oldest.
- **User directory:** Within the **User** directory are folders that correspond to each profile used to date, and within each profile folder are folders named **Work** and **Workspaces**. The directory `User\profile name\Workspaces` is the default location in which Insight workspace files are saved. `User\profile name\Work` is the default location in which Insight visualizations and other custom work performed by the Insight user are saved.

The following table lists the default locations of commonly accessed components.

Insight Default Locations

Component	Directory Location
Saved visualizations	<i>Insight\User\profile name\Work\</i>
Saved Workspaces	<i>Insight\User\profile name\Workspaces\tab name\</i>
Saved .png files	<i>Insight\User\profile name\Work\</i>
Data cache	<i>Insight\User\Cache.db</i>
Insight.log file	<i>Insight\Trace\</i>

Searching for Entries in a Configuration File

The configuration tree in the `Insight.cfg` (or any `.cfg` file) includes a **Search** field.

You can search by key name, key type, or value to quickly locate an entry, to remove the need to scroll through expanded, large files for nested information. You can locate dimension names, server names, and so on. The following example shows matches for a search on the phrase map.

Search: map

Detailed Status for Standard Profile Repository vsdevtest02

- Component Status	map
Addressing	OK
Administrative Alerts	OK
Administrative Event Logs	OK
AsyncEngine	OK
Communications	OK
Disk Files	OK
Profile Manager	OK
SamplingActivity	OK
+ Disk Space	
+ Execution Engine	
+ Hardware	
- Memory Status	ServerMemoryStatus
+ Address Space Map	6408 KB
+ Heap Fragmentation	
+ Histogram	
+ Page File	
+ Physical	
+ Process Address Space	
+ SDAlloc	
+ Small Blocks	
Address Space Load	95% (heavy)
+ Network Connections	
- Performance	
- CPU Usage	
+ Log Processing	
- Transformation	
+ Bank - vsdevtest	5130.9 sec
+ WalMart D	4144.9 sec
- Dimensions	
- Disk Usage	
- Bank - vsdevtest	map
Browser	7.989 MB
Browser Type	7.989 MB
Campaign	7.989 MB

Type a search phrase into this field to locate the data. Depending on the success of a match, the color of the field changes. Matches are shown highlighted and non-matches are dimmed. If there are no matches, the background of the search field turns red. When you press Enter, the config tree expands every place where there is a match and collapses where there is not a match.

You can also use regular expressions in the **Search** field. For example, you can use re: *zip.* for any entry containing the word “zip.”

To clear a search, press **Escape**.

Connection Troubleshooting

Navigation title: Connection Troubleshooting

If Insight cannot connect to the Insight Server(s) using the specified settings, a red node appears in the **Servers Manager**.

This might occur, for example, if you configure the connection incorrectly or you do not have permission to view the Insight Server's status.

To determine why Insight is not able to establish a connection

1. Right-click the red server node and click **Detailed Status**.
2. In the **Detailed Status** interface, click **Network Connections** and expand the numbered items. The **Status** parameter provides information about why Insight is not able to establish a connection:
 - A status code in the 500s indicates a configuration error.
 - A status code of 403 usually indicates that you do not have permission to view the server's status.

You can view additional status information in the `insight.log` file. This log file is located in the **Trace** folder in the directory where you installed Insight. To view the file, open it in a text editor such as Notepad.

If you need assistance understanding the information in the `insight.log` file, first contact your system administrator. If you require further assistance, contact Adobe Customer Care.

Upgrading Data Workbench

Navigation title: Upgrading Data Workbench

Information on upgrading your data workbench software.

Upgrading 6.4 to 6.5

Follow these steps to upgrade to Data Workbench v6.5.

Upgrade Requirements and Recommendations

Follow these requirements and recommendations when upgrading to Data Workbench 6.5.

- Changes in the **Components for Processing Servers\Communications.cfg** file require you to update this file for the DWB 6.5 release. The `SourceListServer`, `SegmentExportServer`, and `NormalizeServer` entries were removed. (DPU's should not be running `sourcelist`, `segment export`, or `normalize servers`.)
- Correlation Chord, Correlation Matrix, Association Chord, Association Matrix, Propensity Score and Best Fit Attribution visualizations are now multi-pass visualizations.

When there are more than one multi-pass visualizations in a workspace, Report Server will fail to generate reports by default with the error:

`Too many Multipass visualizations in workspace (has #, 1 allowed).`

Avoid this error by updating your `ReportServer.cfg` file or add this line to your existing file in the `Reporting` section.

`Max Multipass Per Slice = int: n`

where `n` is the maximum number of multi-pass visualizations supported by Report Server in a workspace.

Upgrading 6.3 to 6.4

Follow these steps to upgrade to Data Workbench v6.4.

Upgrade Requirements and Recommendations

Follow these requirements and recommendations when upgrading to Data Workbench 6.4.



Important: It is recommended that you use the newly installed default configuration files and customize them, rather than moving files from a previous installation—with these exceptions:

- **Add Excluded Processes** for *MS System Center Endpoint Protection in Windows 2012 Servers* for the following executables:

- **InsightServer64.exe**
- **ReportServer.exe**
- **ExportIntegration.exe**

This will allow "white list" rights for these interfacing executables.

- **Update the *Trust_ca_cert.pem* certificate on the servers.**
- **Reorganization of Attribution Profiles.**

- The *Attribution* folder was renamed to ***Attribution - Premium*** (found in the default installation at *Profiles\Attribution - Premium*).

- The *Premium* profile was removed and the workspace moved to the new ***Attribution - Premium*** folder.

- **Update *Attribution-Premium* settings.** If you have customized profiles with parameter settings that override the default *Adobe SC* profile, then you need to update the custom fields in these configuration files:

- **Decoding Instructions.cfg**
- **SC Fields.cfg**

- Because of this reorganization, you will want to remove the old *Attribution* and *Premium* folders from your server installation.

Change these settings

```
Profile = profileInfo:  
    Active = bool: true  
    Directories = vector: 6 items  
        0 = string: Base\  
        1 = string: Geography\  
        2 = string: Predictive Analytics\  
        3 = string: Adobe SC\  
        4 = string: Attribution\  
        5 = string: Premium\  
    
```

to these settings:

```
Profile = profileInfo:  
    Active = bool: true  
    Directories = vector: 5 items  
        0 = string: Base\  
        1 = string: Geography\  
        2 = string: Predictive Analytics\  
        3 = string: Adobe SC\  
        4 = string: Attribution - Premium\  
    
```

- **Update custom Meta.cfg files (if necessary).**

The **Meta.cfg** files in **Base\Context** and **AdobeSC\Context** folders have been updated in this release.

If you override the **meta.cfg** file during installation, then your profile copy needs to be updated with these parameters and the **metadata vector** appropriately entered:

```
94 = meta:  
    path = string: SegmentExport:CRS Configuration/CRS Attributes  
    acceptable children = vector: 1 items  
        0 = Template:  
            name = string: CRS Attributes  
            value = CRSAttributeConfiguration:  
                Attribute Name = string:  
                Attribute Type(int,string) = string:  
                Field Name = string:  
  
95 = meta:  
    path = string: SegmentExportQuery:CRS Configuration/Report Suite  
    acceptable children = vector: 1 items  
        0 = Template  
            name = string: Add Report Suite  
            value = string:
```

- **Set Report Server permissions** to generate Microsoft Excel reports On Windows 2012 servers.

1. Set permission of the root folder (**E:\ReportServer**) to *Everyone = full control*.
2. Create the following folders with appropriate permissions:

```
C:\Windows\SysWOW64\config\systemprofile\AppData\Local\Microsoft\Windows\INetCache  
C:\Windows\System32\config\systemprofile\AppData\Local\Microsoft\Windows\INetCache  
C:\Windows\System32\config\systemprofile\Desktop  
C:\Windows\SysWOW64\config\systemprofile\Desktop
```



Note: If you are running Report Server on Windows Server 2012, you need to have Windows Server 2012 R2 installed.

3. Assign "SYSTEM" as the owner for these folders.

- **Add fonts to the Report Server.**

In the **ReportServer.cfg** file, add these fonts (for all languages):

```
Fonts = vector: 3 items  
    0 = string: Arial  
    1 = string: SimSun  
    2 = string: MS Mincho
```

- **Update your version of Microsoft Excel** (if necessary).

With the release of Data Workbench 6.4, support for Excel 2007 has been discontinued. Also because Data Workbench only runs on Microsoft Windows for 64-bit architecture, it is recommended that you also install a 64-bit version of Microsoft Excel.

- **64-bit architecture** required for Workstation (Client) installation.

- **Run the Workstation Setup Wizard.**

Install the new version of the workstation (client) by downloading and launching **InsightSetup.exe** and stepping through the setup instructions. The setup wizard will install your files to a new location by default:

Program files are now saved by default to:

```
C:\Program Files\Adobe\Adobe Analytics\Data Workbench
```

Data Files (profiles, certificates, trace logs, and user files) are now saved by default to:

```
C:\Users\<username>\AppData\Local\Adobe\Adobe Analytics\Data Workbench\
```

-
- Add fonts to the Workstation.

In the **Insight.cfg** file, add these fonts (for all languages):

```
Fonts = vector: 3 items
  0 = string: Arial
  1 = string: SimSun
  2 = string: MS Mincho
```

DWB Server upgrade: 6.2 to 6.3

Upgrading server components for Data Workbench 6.3.

Upgrade Server

If you have customized profiles that take precedence over the default files provided in the Base package, then you will need to update these customized files:

- **Update the Meta.cfg file** (E:\..\Profiles\<your custom profile>\Context\meta.cfg) to set updated password encryption for the File System Unit (FSU server), and to add entries for the Name Value Pair transformations to take advantage of [DWB Server upgrade: 6.2 to 6.3](#) on page 188.

1. Open the meta.cfg file on the FSU.
2. Change the data type for **Proxy Password** from "string" to "EncryptedString" in the *Workstation Configuration* section.

```
Proxy User Name = string:
Proxy Password = EncryptedString:    (from Proxy Password = String)
Use Address File = bool: true
```

3. Add new entries to enable the new Name Value Pair transformations: *BuildNameValuePair* and *ExtractNameValuePairs*.

Open a workspace and right-click **Admin > Profile Manager**.

Under **Context**, click the **meta.cfg** file in the **Base** column and then click **Make Local**. From the User table column, right-click and select **Open > in Workstation**.



- In the new window, click **metadata** and add acceptable children templates.

+90	meta
-91	meta
-acceptable children	
-0	Template
-value	Column
Column Name	
Field Name	
name	Name Value
path	BuildNameValuePair:Input Columns
-92	meta
-acceptable children	
-0	Template
-value	Column
Column Name	
Field Name	
name	Name Value
path	ExtractNameValuePair:Output Columns
+93	meta

- Open transformation and add new templates.

-transformations	
+0	Template
-1	Template
-value	BuildNameValuePair
Name	
Comments	Comment
Condition	AndCondition
Delimiter	
Input Columns	
Output	
name	BuildNameValuePair
+2	Template
+3	Template
+4	Template
+5	Template
+6	Template
-7	Template
-value	ExtractNameValuePair
Name	
Comments	Comment
Condition	AndCondition
Delimiter	
Input Field	
Output Columns	
name	ExtractNameValuePair
+8	Template

-
- **Update for Fast Merge improvements.** Add parameters or change values to the following configuration files to take advantage of speed improvements in Data Workbench during a transformation.

- **Communications.cfg** (E:\Server\Components\Communications.cfg)

```
18 = SourceListServer:  
    URI = string: /SourceListServer/  
    Listing Interval = int: 10 (new)
```

- **Disk Files.cfg** (at E:\Server\Components and E:\Server\Components for Processing Servers)

```
Disk Cache Size (MB) = double: 1024 (from double: 256)  
Disk Cache Read Limit (MB) = double: 768 (new)
```

- **Log Processing Mode.cfg** (E:\Server\Profiles\<your profile>\Dataset\Log Processing Mode.cfg)

```
(changed)  
Batch Bytes = int: 268435456  
Cloud Bytes = int: 268435456  
Real Time FIFO Bytes = int: 268435456  
  
(new)  
Cache Bytes = int: 32000000  
Fast Input Decision Ratio = double: 200  
Fast Input FIFO Bytes = int: 268435456  
FIFO Hash Mask = int: 16383  
Fast Merge Buffer Bytes = int: 536870912  
Slow Merge Buffer Bytes = int: 268435456  
Fast Merge Fan In = int: 64  
Key Cache Size Logarithm = int: 21  
Max Seek = int: 512  
Output Old Buffer Bytes = int: 536870912  
Overflow FIFO Bytes = int: 67108864  
Paused = bool: false
```



Note: To take advantage of the Fast Merge improvements, make sure you have at least 8 GBs of RAM per DPU.

- **Adobe Target with DWB integration update.** A new export file, ExportIntegration.exe, replaces the existing TnTSend.exe file on the Insight Server (E:\Server\Scripts\TnTSend.exe). This new export file supports both *Adobe Target* integration and coordination with the new Master Marketing Profile (MMP) and *Adobe Audience Manager*.

You will need to update the following commands for Adobe Target exports.

```
Command = string: TnTSend.exe  
to  
Command = string: ExportIntegration.exe
```



Note: This will only affect exports created prior to version 6.3.

You can also try the following to employ the old export process:

- Create a new Test And Target Export in the workstation.
- Modify the old Test and Target export found in Server/Profiles/<your profile>/Export.

-
- **Update the Adobe SC profile.** Changes to the `Exclude Hit.cfg` file require a field to be declared in the associated `Decoding Instructions.cfg` file.



Note: If your Adobe SC profile includes a customized `Decoding Instructions.cfg` file, you will need to include a `DelimitedDecoder` parameter to your customized file.

```
0 = DelimitedDecoder:  
  Delimiter = string: \t  
  Fields = vector: x items  
  ...  
  5 = string:  
Changed to:  
  5 = string: x-hit_source
```

Adding the `DelimitedDecoder` field allows you to take advantage of feature updates and to avoid possible Log Processing problems resulting from these updates.

DWB Server upgrade: 6.1 to 6.2

Upgrading server components for Data Workbench 6.2 and 6.2.2.

Upgrade Issues for 6.2

- The Attribution profile is configured for users who have implemented the Adobe SC profile to employ the Analytics (SC/Insight) data feed. By default, the Marketing and Conversion events are employed as the default interactions evaluated in the rules-based models. See [Deploying the Attribution Profile](#) for additional information.
- For users of the Adobe SC profile upgrading to Data Workbench 6.2, if you are not using the default configurations, verify that the `x-bot_id` value in the `SC Fields.cfg` file is being decoded properly and that the `x-bot_id` field is listed properly in the `Decoding Instructions.cfg` and the `Exclude Hit.cfg` files. This will only be an issue if you have modified the configuration file from the default configuration.
- If you have deleted unused fields in the `Dataset > Log Processing > SC Fields.cfg` file for the Adobe SC profile, you will need to update to accommodate updated field values used for the Attribution profile (see [Deploying the Attribution Profile](#)).

Upgrade Issues for 6.2.2

- The **Browsers and Operating Systems** lookup files will not be updated within the legacy **Traffic** profile (for example, `Lookups\Traffic\Browsers.txt`). Instead, configuration of the **Traffic** profile will utilize the DeviceAtlas bundle (`Lookups\DeviceAtlas\DeviceAtlas.bundle`) to provide this configuration information.
- Data Workbench 6.2.1 will be the last release to provide a download of the 32-bit client application. All future client application downloads will be 64-bit and continue to require Windows 7 or newer. Memory limitations of the 32-bit application are addressed with the introduction of the 64-bit application beginning with the 6.1 release.



Note: The 32-bit version of the Data Workbench client application may experience potential issues related to memory limitations when running predictive models using the clustering and scoring features.

Data Workbench 6.0 to 6.1 Upgrade

Follow these steps to update to data workbench v6.1 from your data workbench v6.0x installation.

Step 1: [Server Upgrade](#) on page 192

Step 2: *Report Server Upgrade* on page 193

Step 3: *Client Upgrade* on page 194



Important: The server, report server, and client components are upgraded to run on 64-bit Windows operating systems.

Server Upgrade

Follow these steps to update the **Server v6.1** components:

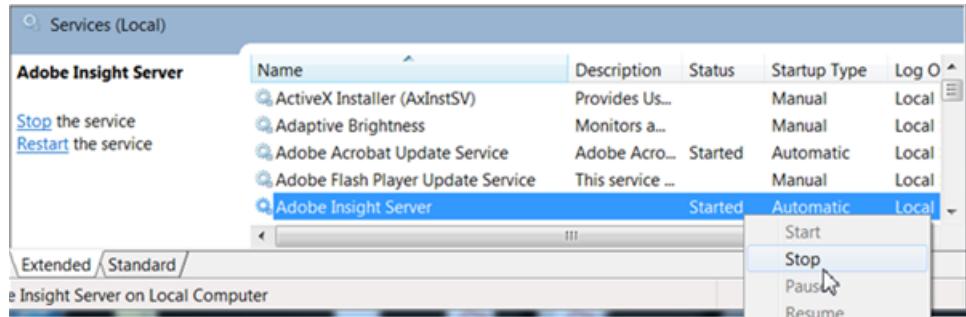
1. Using the **Software and Docs** profile, open the **Start Here** workspace and download all needed server packages to a local folder.

- Download **Server Packages \ v6.1** zip folders and extract all files.

The Server package includes **Lookup** and **Profile** folders with **Base** and **Transform** profiles to update the server.

- Download the **Profiles** folders.
- Download the **Lookup** folders.
- Download the **Report Server \ v6.1** package.
- Download additional **Sensor**, **Documentation**, and **Dashboard** files as needed for your system.

2. Stop the **Adobe Insight Server** service.



3. From the downloaded **Server** package:

- a. Replace the **Server\Bin** folder to update the **InsightServer64.exe** and supporting files.
- b. Replace the **Server\Profiles** folder. You can overwrite all files.
- c. Update the **Server\Lookups** folder. You will want to add the newly downloaded files to the custom files already located in the folder.
- d. Replace the **Server\Software** folder to update **Insight.exe** and **ReportServer.exe**.
- e. Update the **Server\Scripts** folder to update **TnTSend.exe**.

4. If you employ **DeviceAtlas**, then you will need to *update the bundle* located in the **Server\Lookups** folder.

5. Configure the **Profile.cfg** file to ensure that the vector is updated to reflect the number of items for each profile.

For example, to enable the **Predictive Analytics** profile you will need to update this setting.

```
Directories = vector<string>: 5 items
  0 = string: Base\\
```

```
1 = string: Geography\\
2 = string: Predictive Analytics\\
3 = string: Adobe SC\\
4 = string: Profile Name\\
```

6. Configure and save the PAServer.cfg file for the Predictive Analytics feature.

If you want to submit Predictive Analytics jobs to the servers, then you will need to configure the Server > Predictive Analytics > Dataset > PAServer.cfg file to manage server-side clustering submissions.

The custom profile should inherit the settings from the Predictive Analytics configuration profile, allowing you to configure and save the PAServer.cfg file based on your site's implementation.

7. Define the Log Source ID.

The **Recording of Rows per Log Source** was added in **v6.04** and defined in the custom profile's Log Processing.cfg file by adding a uniquely named **Log Source ID**.

```
Log Processing.cfg
  Log Source ID = string: <Name your ID Here>
```

If you do not have the Log Source ID defined, then you will get the following error:

```
Missing Log Source ID in log processing.cfg.
Log Source ID must be defined for all log sources.
```

8. Because the EventMessages.dll has been updated, it is required that you unregister and then register the **Adobe Insight Server** across the cluster.

- InsightServer64.exe /unregserver
- InsightServer64.exe /regserver

9. Start the **Adobe Insight Server** service across the cluster.

The server installation is now complete.

Report Server Upgrade



Important: Before upgrading to **Report Server v6.1**, you must first upgrade to **Server v6.1**.

1. Using the **Software and Docs** profile, download **v6.1** from the **Report Server** package to a local folder.
2. Copy **Report Server 6.1** from the downloaded package and replace the profile packages.



Note: The Insight.zbin file in the install folder is a backup file used for localization, and must be present in the install directory. This file or other .zbin files will be used depending on the command-line settings passed when starting up.

3. (optional) Data workbench currently supports English (-en-us) and Chinese (-zh-cn). You need to set a font to support single and double-byte characters:

```
Report Server.cfg - Add Fonts
  Fonts = vector: 2 items
  0 = string: SimSun
  1 = string: Arial
```

The Windows operating system must also have the listed fonts installed.

4. Configure Report Server v6.1 for localization.

- a. Stop the **Adobe Insight Report Server** service.

-
- b. Launch a command prompt as an "Administrator".
 - c. Navigate to the Report Server `install` folder.
 - d. Delete the Report Server service using the following command:

```
ReportServer.exe /unregserver
```

- e. Start the service based on language settings:

```
ReportServer.exe -RegServer -Locale -en-us (English)  
ReportServer.exe -RegServer -Locale -zh-cn (Simplified Chinese)
```

- 5. To verify that Report Server is running with the correct settings, open up **Windows Service Manager** and right-click **Adobe Insight Report Server - Properties**. The path to the executable will display the updated command-line settings.

The report server installation is now complete.

Client Upgrade



Important: Before upgrading to **Client v6.1**, the administrator must first upgrade to **Insight Server v6.1**.

- 1. Launch `Insight.exe` but do not connect to any profiles.
- 2. Edit the `Insight.cfg` file.

```
Update Software = bool: true
```

- 3. Connect to your profile.

Allow the client to synchronize with the server and your client will be upgraded with the latest v6.1client profiles, executables, and configuration files.



Note: The `Insight.zbin` file in the `install` folder is a backup file used for localization and must be present. This file or other `.zbin` files will be used depending on the command-line settings passed when starting up.

See [Installing the Input Method Editor](#) on page 133 to add an `insight.zbin` file required for localized settings.

Additional Client Settings

Before configuring `Insight.exe` and supporting files, you must exit the client application.

To install Simplified Chinese:

- 1. Create a shortcut that passes in the command-line setting to the `Insight.exe` file.

```
Insight.exe -zh-cn
```

- 2. Configure `Insight.cfg` to support single and double-byte font characters.

Data workbench currently supports both English and Simplified Chinese. You can select fonts to support both of these languages:

```
Fonts = vector: 2 items  
0 = string: SimSun  
1 = string: Arial
```

The Windows operating system must also have the requested fonts installed.

- 3. Launch the shortcut that you created to synchronize profiles and the updated `.zbin` file.

To employ the Input Method Editor (IME).

IME allows you to input international characters.

1. Update the Insight .cfg file with these settings:

```
Localized IME = bool: true
```

2. Launch the shortcut that you created to synchronize profiles and the updated .zbin file.

The client installation is now complete.

Data Workbench 5.5 to 6.1 Upgrade

Follow these steps to update to data workbench v6.1 from your Insight v5.5x installation.

Step 1: [Server Upgrade](#) on page 195

Step 2: [Report Server Upgrade](#) on page 196

Step 3: [Client Upgrade](#) on page 197



Important: The server, report server, and client components are upgraded to run on 64-bit Windows operating systems.

Server Upgrade

Follow these steps to update the **Server v6.1** components:

1. Using the **Software and Docs** profile, open the **Start Here** workspace and download all needed server packages to a local folder.

- Download **Server Packages \ v6.1** zip folders and extract all files.

The **Server** package includes **Lookup** and **Profile** folders with the **Base** and **Transform** lookup files to add and replace to update the server.

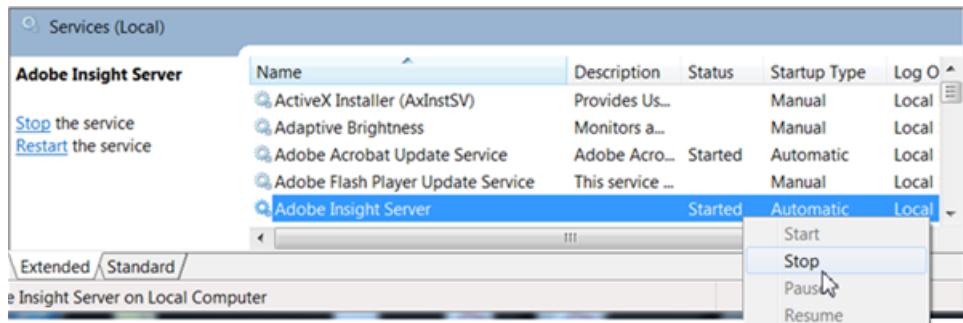
- Download new **Profiles** folders.

- Download updated **Lookup** folders.

- Download the **Report Server \ v6.1** package.

- Download additional **Sensor**, **Documentation**, and **Dashboard** files as needed for your system.

2. Stop the **Adobe Insight Server** service.



3. From the downloaded **Server** package:

- a. Replace the **Server\Bin** folder to update the **InsightServer64.exe** and supporting files.

-
- b. Replace the Server\Profiles folder. You can overwrite all files.
 - c. Update the Server\Lookups folder. You will want to add the newly downloaded files to the custom files already located in the folder.
 - d. Replace the Server\Software folder to update Insight.exe and ReportServer.exe
 - e. Update the Server\Scripts folder to update TnTSend.exe.
4. If you employ **DeviceAtlas**, then you will need to *update the bundle* located in the Server\Lookups folder.
 5. Set Directories in the Profile.cfg file to ensure that the vector is updated to reflect the number of items for each profile.

For example, to enable the **Predictive Analytics** profile you will need to update this setting.

```
Directories = vector: 5 items
  0 = string: Base\\
  1 = string: Geography\\
  2 = string: Predictive Analytics\\
  3 = string: Adobe SC\\
  4 = string: Profile Name\\
```

6. Configure and save the PAServer.cfg file to upgrade the Predictive Analytics feature.

If you want to submit Predictive Analytics jobs to the servers, then you will need to configure the Server > Predictive Analytics > Dataset > PAServer.cfg file to manage server-side clustering submissions.

The custom profile should inherit the settings from the Predictive Analytics configuration profile, allowing you to configure and save the PAServer.cfg based on your site's implementation.

7. Define the **Log Source ID**.

The **Recording of Rows per Log Source** was added in **v6.04** and defined in the custom profile's Log Processing.cfg file by adding a uniquely named **Log Source ID**.

```
Log Processing.cfg
  Log Source ID = string: <Name your ID Here>
```

If you do not have the Log Source ID defined, then you will get the following error:

```
Missing Log Source ID in log processing.cfg.
Log Source ID must be defined for all log sources.
```

8. Because the EventMessages.dll has been updated, it is required that you unregister and then register the **Adobe Insight Server** across the cluster.
 - InsightServer64.exe /unregserver
 - InsightServer64.exe /regserver
9. Start the **Adobe Insight Server** service across the cluster.

The server installation is now complete.

Report Server Upgrade



Important: Before upgrading to **Report Server v6.1**, you must first upgrade to **Server v6.1**.

1. Using the **Software and Docs** profile, download **v6.1** from the **Report Server** package to a local folder.
2. Copy **Report Server 6.1** from the downloaded package and replace the profile packages.



Note: The Insight.zbin file in the `install` folder is a backup file used for localization, and must be present in the `install` directory. This file or other .zbin files will be used depending on the command-line settings passed when starting up.

3. (optional) Modify the report server configuration file to support double-byte characters.

Data workbench currently supports English (-en-us) and Chinese (-zh-cn). You need to set a font to support single and double-byte characters:

```
Report Server.cfg - Add Fonts
  Fonts = vector: 2 items
    0 = string: SimSun
    1 = string: Arial
```

The Windows operating system must also have the listed fonts installed.

4. Configure Report Server v6.1.

- a. Stop the **Adobe Insight Report Server** service.
- b. Launch a command prompt as an "Administrator".
- c. Navigate to the Report Server `install` folder.
- d. Delete the Report Server service using the following command:

```
ReportServer.exe /unregserver
```

5. Start the service based on the language settings:

```
ReportServer.exe -RegServer -Locale -en-us (English)
ReportServer.exe -RegServer -Locale -zh-cn (Simplified Chinese)
```

6. To verify that Report Server is running with the correct settings, open up **Windows Service Manager** and right-click **Adobe Insight Report Server - Properties**. The path to the executable will display the updated command-line settings.

The report server installation is now complete.

Client Upgrade



Important: Before upgrading to **Client v6.1**, the administrator must first upgrade to **Server v6.1**.

1. Launch `Insight.exe` but DO NOT connect to any profiles.
2. Edit the `Insight.cfg` file to not update software automatically.

```
Update Software = bool: false
```

3. Connect to **Software and Docs** profile (`softdocs`).
4. Download `Software\Insight Client\v6.10`.
5. (optional) Modify `insight.cfg` to support double-byte characters.

Data workbench currently supports both English and Simplified Chinese. Select fonts to support both of these languages:

```
Fonts = vector: 2 items
  0 = string: SimSun
  1 = string: Arial
```

6. Exit out of the client.
7. Copy the files in the downloaded **v6.1** client package to the `Install` folder.



Note: The `Insight.zbin` file in the install folder is a backup file used for localization, and must be present in the install directory. This file or other `.zbin` files will be used depending on the command-line settings passed when starting up.

For example, to launch Simplified Chinese, create a shortcut that passes in the command-line setting.

```
Insight.exe -zh-cn
```

If you want to launch in English (default), then no command-line change is necessary.

8. Launch `Insight.exe` for English or the shortcut that you created for another language.
9. Connect to your profile and allow the client to synchronize with the server.
10. (optional) To employ the IME, make these changes to the `Insight.cfg` file:

```
Localized IME = bool: true
```

The Input Method Editor (IME) allows you to input international characters.

11. (optional) Edit the `Insight.cfg` file to automatically update software:

```
Update Software = bool: true
```

See instructions for implementing the IME.

12. Restart again after the profile synchronization to employ the most recent `.zbin` file.

The client installation is now complete.

Data Workbench Client

Navigation title:

An overview of the basic elements, procedures, and terminology that you need to understand when working with Data Workbench.

For details about working with the specific types of visualizations, see [Analysis visualizations](#) on page 304 and [Administrative interfaces](#) on page 446.



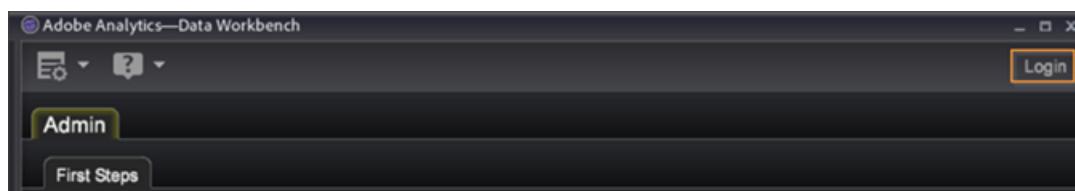
Note: This guide provides you with information about the functionality and user interface delivered with Data Workbench. Because your implementation of Data Workbench can be fully customized, it may differ from what is documented in this guide.

Opening Data Workbench

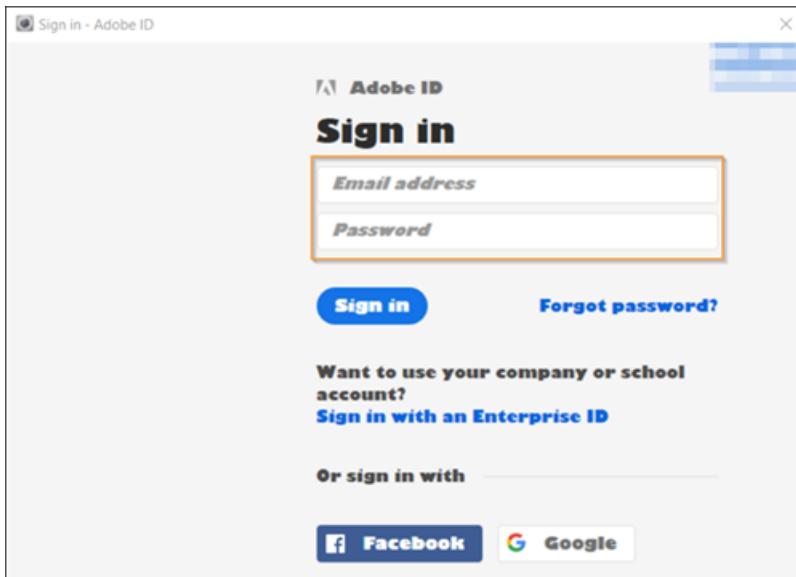
Navigation title:Opening Data Workbench

After Data Workbench has been installed and is connected to the server, you can use it to analyze your processed data.

1. Navigate to the Data Workbench installation directory.
2. Double-click `insight.exe` to start the application.
3. Click **Login**.



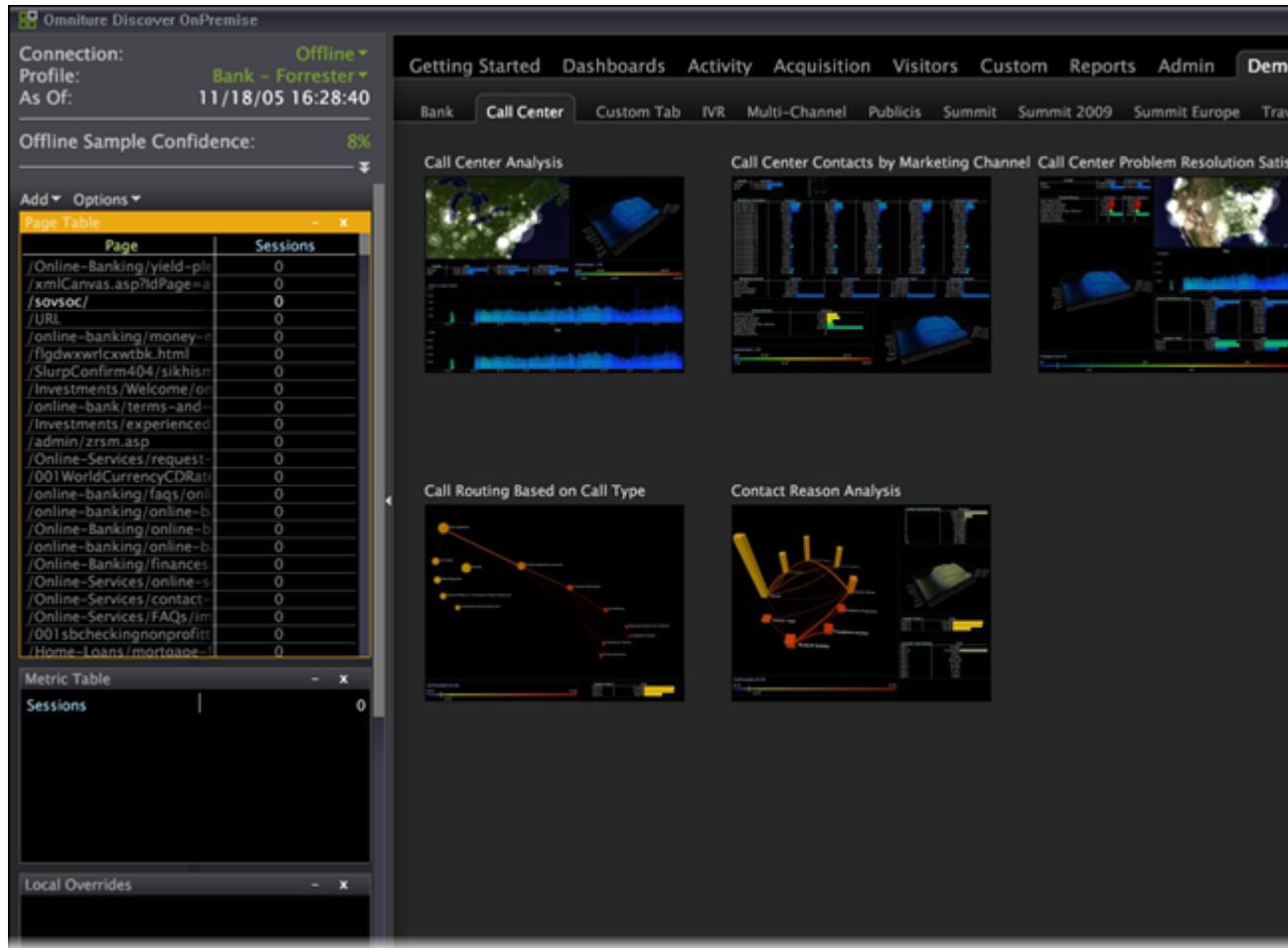
-
4. Enter your credentials.



The first time you start, the system displays the **Start** page and begins downloading profile data and the data cache.

5. After the data cache begins to fill, click **Start**.

The system displays the **Worktop**.



The next time you start Data Workbench, the system skips the **Start** page and displays the **Worktop**, unless a new profile has been added to the server.

Updates to the profile and its data are downloaded only if you are working online and have a network connection to the Data Workbench server. For more information about working online, see [Working offline and online](#) on page 201.



Note: It is not recommended for multiple users to share a single machine for using the Workstation.

Self-provisioning of users

Use Workstation to manage your Data Workbench users.

You can use Workstation to connect to the Data Workbench Server by setting up the required certificate from Adobe. This certificate aids in both SSL communication and authorization to use the licensed resources and features. In certificate-based authentication, you need to acquire and setup multiple certificates for using the Workstation on multiple machines. Also, user provisioning and entitlements are managed by Adobe and you must contact Adobe for new certificates or certificate revocation.

Starting in DWB 6.7, the Workstation supports user authentication through username and password.

While the certificate-based authentication/authorization will still work for your setup, it is highly recommended to migrate to the newer credentials-based authentication. In the newer approach, your Workstation users

authenticate themselves through the Adobe Identity Management System (IMS). Before they can use the Workstation, they need to be given access to the features through the [Admin Console](#) (formerly known as Enterprise Dashboard) by the administrator of the organization.

The new authentication and user provisioning model helps in:

- Self-provisioning of users and groups through Admin Console. You don't have to contact Adobe to add, remove, or modify license entitlements for users.
- Accessing Workstation from multiple machines without losing the configuration state by logging in using credentials. The local cache is deleted on log out, the current profile is closed, and Configuration profile becomes active.

Getting started

Before you begin, contact Adobe to add your organization in the Admin Console. Depending on the services you have purchased, Adobe will provision the organization for you. For example, organizations can have access to the Attribution service or Beta builds, or both. Once an organization is configured, the organization administrator can add users and groups. See [Managing Users and Groups](#) in Marketing Cloud for more information. The organization administrator can also configure usage restrictions for different users depending on their roles. For example, non-pre-release users do not need access to the Beta builds.

Each provisioned user added to this organization through the Admin Console will have access to use the Data Workbench. The sub-services can only be enabled or disabled for each user depending on their product access. When a user is upgraded from certificate to IMS, all local data will be copied to the new IMS user directory.



Note: A session lasts 6 hours on Server and 23 hours on Client unless the access token is refreshed. When the token is refreshed, you can use Client without logging in again.

At least one Product Level Configuration needs to be created in Admin Console by the administrator before giving access to any user.

The boolean flag **Use IMS** can be added to `Insight.cfg` to fallback to certificate mode. For information on configuring Access Control for IMS users, see [Updating the Access Control File](#).

Conflict resolution

When a user is logged on to multiple machines with same IMS account on the same profile, and is in offline mode on one of the machines, a `.conflict` may form and a pop-up window will inform you. This occurs when there is a difference in content with any files (workspaces, dimensions, filters, etc.) synced on both machines in `User\Profile\` on server and client. A backup will be created in the `.conflict` file and no data will be lost. A boolean flag in `Insight.cfg` gives you the ability to disable this conflict pop-up.

Flag: Conflict Notifications

This is applicable for workspaces, metrics, dimension, etc. in User Folder.

Working offline and online

Navigation title:

Information about working with the Data Workbench server either offline or online.

Data Workbench automatically downloads updates to the profile and its data from the Data Workbench server if you have a network connection to the server and are working online. If you have not specified to work online, Data Workbench loads the profile and its data from your computer's cache. In this case, you are viewing the version of the profile and its data that was downloaded the last time that you worked online with the profile.

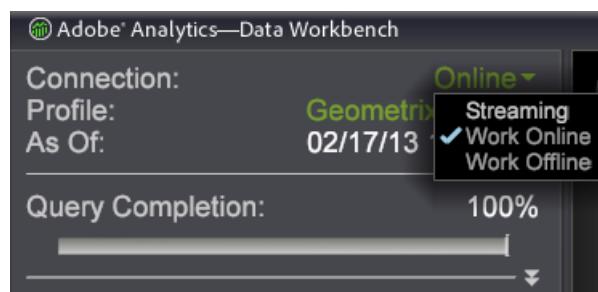
Working offline offers a processing speed advantage because you are working from the local cache and querying the data on your own computer. When working online, every query must go back to the Data Workbench server, which takes longer and forces you to compete with other online users for server resources. As long as you have a network connection to the Data Workbench server, working offline stops the Data Workbench server from

updating the profiles or data on your Data Workbench, but it does not stop you from saving items to the Data Workbench server.

Because of the ability to work offline, the Data Workbench server is sized to handle some amount of real-time traffic input and some amount of data in the dataset, along with some number of Data Workbench users, but it does not have to be sized to support the maximum number of concurrent users (which in practice does not happen often). Because users are usually looking for trends and ratios, exploring the data as you go, in most cases you do not need exact counts. If there is a need to query and resolve to exact counts using current data, you can work online and get that, but the queries take longer to resolve to 100 percent.

To toggle online, offline, or streaming

In the side bar, click the **Connection** setting and click **Work Online**.



When you work online, Data Workbench connects to the Data Workbench server and synchronizes the information on your machine with the profile and its dataset information residing on the Data Workbench server.

The default configuration for Data Workbench is to work offline, but as described in the following section, each user can change their `Insight.cfg` file to make their instance of Data Workbench work online by default.

To work online by default

1. Navigate to your Insight installation directory.
2. Open the `Insight.cfg` file in a text editor.
3. Add the highlighted line to the file as shown in the following example:

```
...
Update Software = bool: true
Default to Online = bool: true
Color Set = int: 0
```

The next time that you open Data Workbench, it connects to the Data Workbench server and works online by default.

Profiles

Navigation title:

Data Workbench downloads profiles to your machine.

If you are loading a profile for the first time, you must have a network connection to the Data Workbench server and be working online so that Data Workbench can download the necessary files from the Data Workbench server.

Downloading the profile might take several minutes. You should not start working with the profile until the Data Cache begins to fill, but you do not have to wait until it is full. You can track the progress of the data cache, the progress of the profile synchronization, and the date and time of the most recently processed data, by looking at the status bars as the profile loads.



Note: You do not see data in visualizations that you add until the data cache begins to fill.

The next time that you load the profile, updates to the profile and its data are downloaded only if you have a network connection to the Data Workbench server and are working online. If you are working offline, the profile and its data are loaded from your machine's cache. In this case, you are viewing the version of the profile and data that was downloaded the last time that you worked online with the profile. For more information about working online versus offline, see [Working offline and online](#) on page 201.

When you need to change your profile (using the **Profile Manager** or the **Server Files Manager**), you should work online to ensure that you have the most up-to-date version of the profile. For more information about the **Profile Manager** and the **Server Files Manager**, see [Administrative interfaces](#) on page 446.

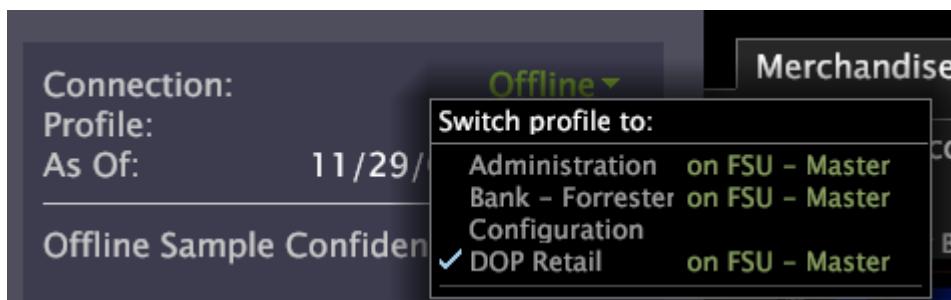
If you are unable to access or load a profile, you may need to confirm the following:

- You have a network connection to the Data Workbench server machine on which the profile resides.
- You have the appropriate permissions to access the profile.

For assistance, contact your system administrator.

Loading or switching profiles

1. Launch Data Workbench.
2. In the side bar, click the profile name and click **Switch Profile > <profile name>**, where *profile name* is the profile with which you want to work.



If this is your first time loading the selected profile, it may take several minutes to download enough data to populate a visualization.

Accessing a profile on a cluster

Data Workbench users who access a profile running on a Data workbench server cluster identify only the master Data Workbench Server in the Data Workbench configuration file (`Insight.cfg`). From the perspective of the Data Workbench user, the profile is accessible on only one Data Workbench Server (the master Data Workbench Server). However, query requests from analysts can be directed to any of the Data Workbench Servers in the cluster.

For more information about profiles running on a Data Workbench Server cluster, see the *Server Products Installation and Administration Guide*.

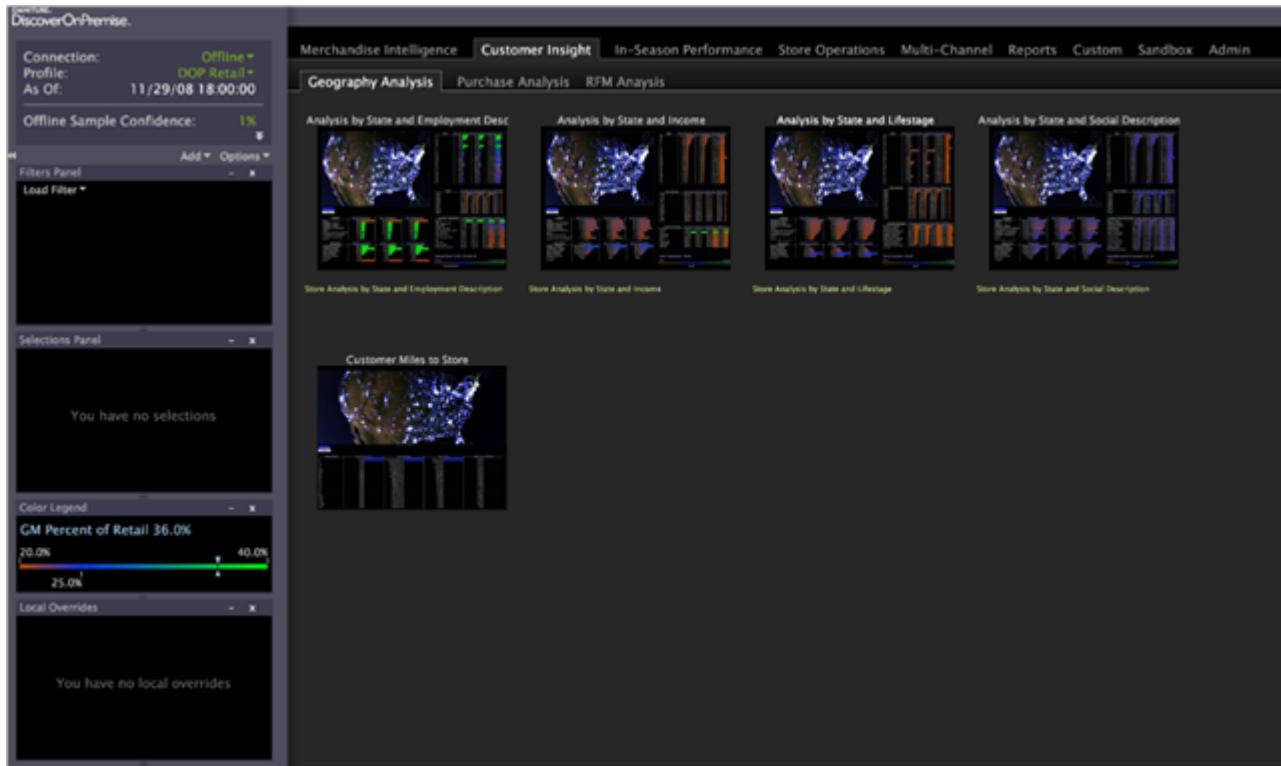
Worktops

Navigation title:

The **Worktop** is where you organize and access all of your workspaces and reports.

In most cases, the **Worktop** is displayed immediately after you open Data Workbench. The features of the **Worktop** include the sidebar and tabbed interface. Additionally, the sidebar lets you add visualizations and access regularly-used functions.

Worktop



The **Worktop** also enables you to create and save new and updated workspaces and reports, as well as publish workspaces and reports to the Data Workbench server for others to access.

The **Worktop** elements table below describes each element of the **Worktop**.

Sidebar	<p>The sidebar provides context and control for workspaces, as well as awareness of the current state of a workspace and access to regularly-used functions. The following functions are available in the sidebar:</p> <p>Connection: A status indicator showing the online status. Click the connection status to enable or disable Work Online. See Working offline and online on page 201.</p> <p>Profile: An indicator of the current profile in use.</p> <p>As Of: A status indicator showing how up-to-date the data is in the profile's data set. This data is downloaded and processed from the DPU server, which can occur only when you are working online.</p> <p>Query/Sample Confidence: An indicator of the query completion. When the status queries to 100 percent, all of the data has been queried.</p> <p>Add: Lets you add visualizations like panels, legends, and tables to the sidebar. See Add visualizations to the sidebar on page 209.</p> <p>Options: Lets you revert to a previous sidebar setting, and to automatically hide the sidebar.</p> <p>Sidebar settings are saved in the <code>sidebar.vw</code> file when you close Data Workbench.</p>
Tab and Subtab or Drop-down	Each tab that appears on the Worktop corresponds to the tab's <i>working profile name\Workspaces\tab name</i> folder within the Data Workbench installation directory and represents a particular type of information, such as

subdirectories (not shown)	Dashboards, Activity, Acquisition, Visitors, and so on. The subfolders in the tab name folder display as subtabs by default, but they also can be displayed as subdirectories. See Customize a worktop tab on page 490.
	 Note: Each Data Workbench profile is delivered with a standard set of tabs. Because your implementation can be fully customized, the workspaces (and, therefore, tabs) that appear may differ from what is documented in this guide.
Profile Status	Provides the connection status to the Data Workbench server and the name of the currently loaded profile. The As Of date and time for the data in the profile's dataset display below the online indicator.
Minimize, Maximize, Close	Standard Windows functions.
Thumbnails	<p>A thumbnail is a snapshot of a workspace that appears on the Worktop. A new snapshot is taken every time you save the workspace. Thumbnails enable you to quickly identify a particular workspace on the Worktop.</p> <p>To open a workspace, click the thumbnail.</p>  Note: Each Data Workbench profile is delivered with a standard set of workspaces. Because your implementation can be fully customized, the workspaces (and, therefore, thumbnails) that appear may differ from what is documented in this guide.
	<p>For more information about workspaces, see Configure the sidebar on page 209.</p>
Error messages	<p>Error messages display in red below the status. For status code descriptions, see http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html.</p> <p>For example: 403_Forbidden.</p>

File versioning

Navigation title:

The **Worktop** makes it easy to determine where each particular workspace is stored, whether it be on the Data Workbench server, your local machine, or both.

Identifying file versions

Server

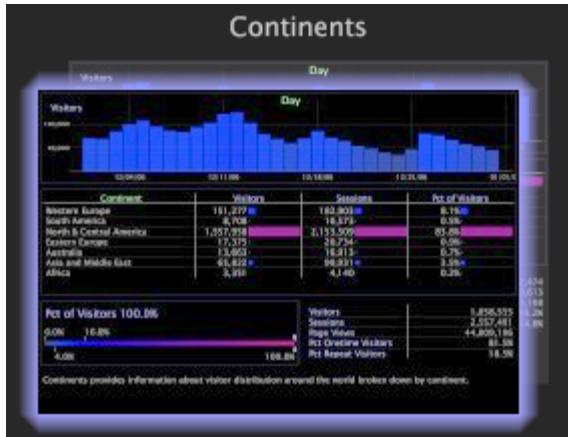
A server workspace is stored on the connected Data Workbench server and is available to all users who have access to this profile and tab. A server workspace displays as a single thumbnail.



Server workspaces are stored by default in the appropriate sub folder within the Workspaces folder on the connected Data Workbench server.

Local

A local workspace is the local version of a server workspace. A local workspace displays as two overlapping thumbnails. The thumbnail on the top initially is surrounded by a glow, which indicates recent changes were made locally to the server workspace. This glow dissipates over time.



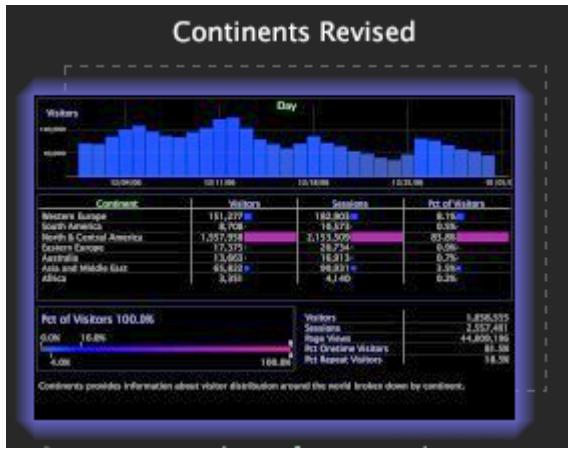
Local workspaces are stored by default in the `User\working profile name\Workspaces\tab name` folder within the Data Workbench (or Insight) installation directory.



Note: When you have a local version of a server workspace, you must revert to the server version before you can download an updated version of the server workspace. To revert back to the server version without local changes, right-click the thumbnail of the local workspace and click **Revert to server version**.

User

A user workspace is a workspace that was created on and exists only on the local machine. A user workspace displays as a single thumbnail with a dotted outline of a blank workspace behind it, indicating that there is no source workspace on the connected Data Workbench server.



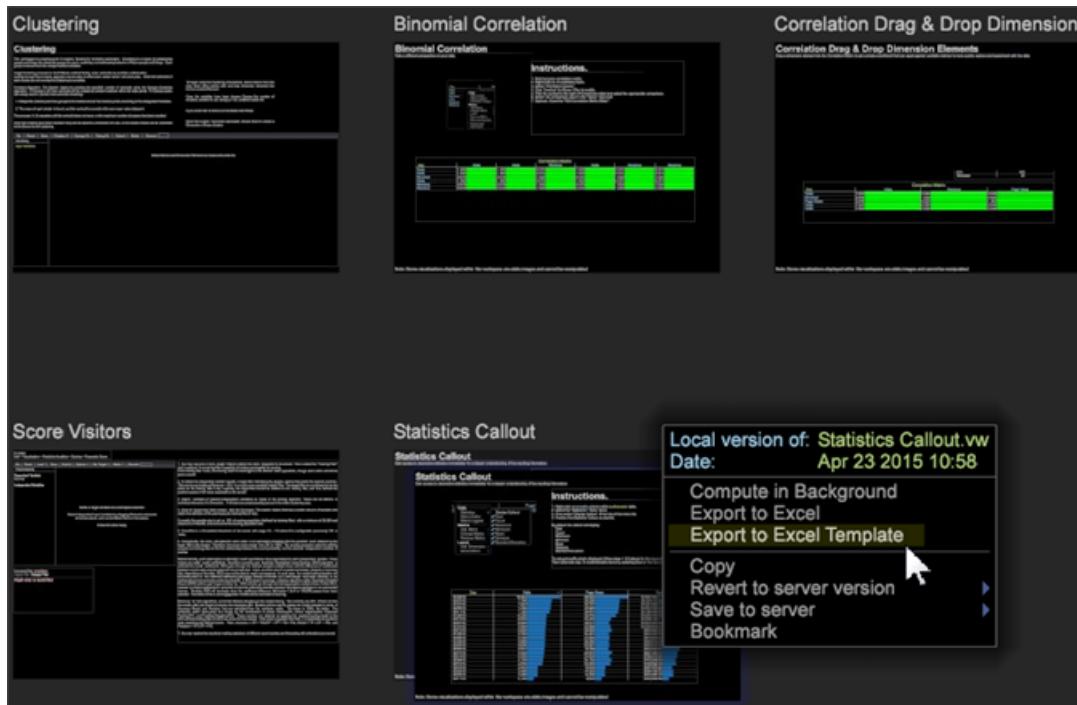
User workspaces are stored by default in the `User\working profile name\Workspaces\tab name` folder within the Insight installation directory.

Using the Worktop thumbnail menu

Navigation title:

How to export, copy, and bookmark from the Worktop.

Right-click a Workspace to export, copy, and bookmark functions from the Worktop.



Interface descriptions

The following elements are available in the **Worktop** thumbnail menu:

Server workspace: *name*

Appears only for unedited server workspaces. Identifies the named workspace as the same as the workspace stored on the server.

Date: *day and time*

The date and time the workspace was last opened.

Local version of: *name*

Appears only for local versions of server workspaces. Identifies the named workspace as an edited, local version of a workspace stored on the server.

User workspace: *name*

Appears only for user workspaces. Identifies the named workspace as a workspace that exists only on the local machine.

Compute in Background

Appears only when working online. Keeps the queries in the selected workspace running in the background while you continue working. When selected, the thumbnail displays the following information, which indicates the progress of the queries:

- Working: *n%* - indicates that the query is processing and the percentage of the processing that is complete.

-
- n MB Query Load - total size of the query result. Query Load is proportional to the total memory load of your Data Workbench server, but does not correlate directly. As a guideline, a 10 MB or higher query load may strain your system. The query load listed does not take clustering into account.



Note: If Compute in Background remains selected, the queries in the selected workspace becomes standing queries, continuing to be updated and use memory load. Make sure to deselect Compute in Background when you are finished working in the workspace.

Export to Excel

Export workspace data to table in Microsoft Excel (.xls and .xlsx files). When exporting a workspace to Excel, Data Workbench exports data from certain visualizations, dimension and value legends, and text annotations to a new Excel workbook with one visualization per worksheet.

Export to Excel Template

Export to an Excel template (.xltx).

Copy

Copies the workspace. For more information about pasting a copied workspace, see [Copying and Pasting Existing Workspaces](#) on page 215.

Revert to server version

Appears only for local versions of server workspaces. Deletes the local copy of this workspace. The original remains on the server.

Delete

Appears only for user workspaces. Deletes the user workspace, which exists only on the local machine. For information about deleting workspaces from the connected Data Workbench server, see [Delete files from your working profile](#) on page 460.

Save to server

Appears only for local versions of server workspaces and user workspaces and works only for those users with the appropriate permissions. Saves the local copy of the workspace to the server. By default, workspaces are saved to the appropriate working <profile name>\Workspaces\<tab name> folder.

Bookmark

Bookmark a workspace to quickly retrieve later.

A bookmark icon will appear above the workspace on the worktop and the name of the workspace will appear in the Bookmark panel.



Configure the sidebar

The sidebar provides access to regularly-used functions and preserves visualizations as you move between Workspaces.

Administrators can customize a sidebar to make it appropriate for different user groups, then deploy the sidebar with a profile.

The sidebar is ideal to help you keep track of filters and local overrides. If you prefer not to use the sidebar, you can hide it.

Add visualizations to the sidebar

1. Launch Data Workbench.
2. In the sidebar, click **Add** > *<item>*. For example, **Selections Panel**, **Filters Panel**, or **Table**.

The following sidebar panels are available in the standard installation of Data Workbench. More items might be available in your specific profile:

- **Selections Panel:** Lets you understand what selections are active in the current workspace. The **Selections Panel** updates whenever you make a new selection. You can clear selections by clicking **x**. See [Make selections in visualizations](#) on page 281 for information about how to select data.
- **Filters Panel:** Makes it easy to load and apply saved filters. You can load multiple filters, and enable or disable each one independently by clicking the check box next to it. See [Filter editors](#) on page 348.
- **Local Override Panel:** This panel displays which metrics, dimensions, and filters that are present in the profile have been modified in your personal copy of the profile. This helps alert you to possible differences between the way data appears in your client and that of other users. When you save changes in a metric, dimension, or filter to the server, the override is removed from the **Local Overrides panel**. If you click an override and then click **Revert to Server**, the local override is removed and the item reverts to the shared version.
- **Metric Legend:** Adds a metric legend. **Metric legends** enable you to see baseline metrics related to your profile and statistics related to the dataset (or to the current selection, if one has been made). See [Metric legends](#) on page 380.
- **Color Legend:** Adds a color legend. You can color-code visualizations by metrics, such as Conversion and Retention, and use them in almost every Workspace. Linking business metrics to color makes it easy to spot anomalies, exceptions, and trends. See [Color legends](#) on page 378.
- **Text Annotation:** Adds a notes panel. **Text annotations** are windows into which you can enter arbitrary text to add descriptive information or comments to a **Workspace**. See [Text annotations](#) on page 308.
- **Table:** Adds a table. A table can display one or more metrics across one or more dimensions of data. See [Tables](#) on page 414.
- **Open:** Opens a saved file.

Open a Sidebar Panel

You can open a sidebar visualization file from a saved location or from the clipboard.

1. In the sidebar, click **Add** > **Open**.
2. Click **File** to locate the .vw file of the panel you want to add, or click **Last Closed Window**, which pulls the visualization from the clipboard.

Additionally, you can click **From Clipboard** to paste a visualization that has been copied to the clipboard. See [Copying a Sidebar Panel](#) on page 210.

Copying a Sidebar Panel

1. Right-click the panel's top border, then click **Copy > Window**.
2. To paste the panel, click **Add > Open > From Clipboard**.

Saving a Sidebar Panel

On an sidebar panel, right click in the title bar and click **Save**.

Similarly, you can open a saved sidebar visualization. Data Workbench saves the visualization as a .vw file at the location you specify.

Revert to the Default Sidebar

In the sidebar, click **Options > Revert**.

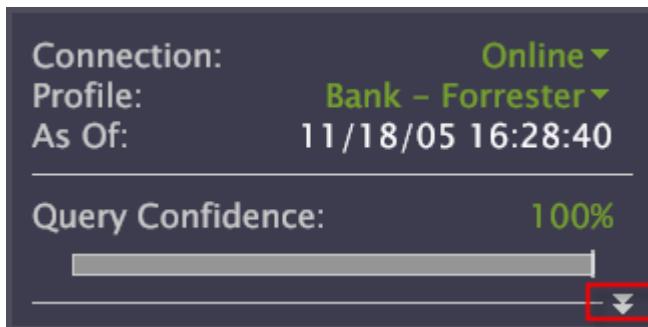
When you close Data Workbench, the system saves the current sidebar configuration in the sidebar.vw file in the user profile. When you open Data Workbench, the system loads the sidebar.vw file from the user profile, rather than a parent profile.

You can revert to a default or previously saved sidebar, which deletes the sidebar from the user profile and reloads the sidebar from the parent profile. Administrators can replace the default (parent) sidebar with a local sidebar by uploading it from the **Profile Manager**.

Customize the More Status Panel File

System administrators can build formulas in the More Status Panel.vw. This places contextual words around metric and dimension values, and displays the results in the **More Status panel** in the sidebar.

To display the **More Status panel** in the sidebar, click the arrows shown in the following example.



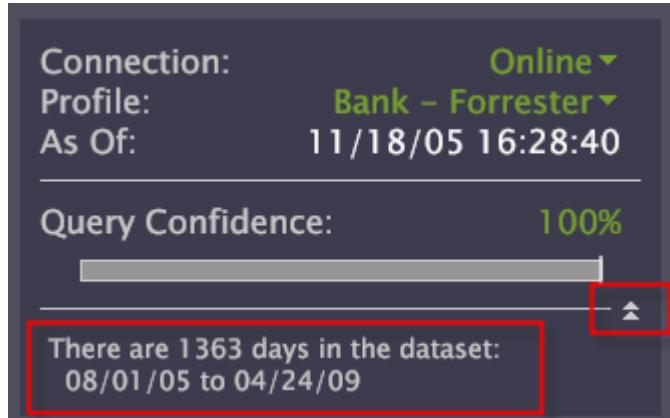
The following procedure shows a simple example of how to create a customized status that tells you how many days are in a dataset:

1. In the **Profile Manager**, click **Sidebar**.
2. In the **Base_5_3*** column, make a local copy of the More Status Panel.vw file.
To do so, right-click the file check mark and click **Make Local**.
3. Open the More Status Panel.vw file in the .vwEditor or in Notepad.

Search:	
-More Status Panel.vw	
-panel	SidebarContextVis
Name	More Details
-Context Rows	
-0	contextRow
-Items	
-0	contextString
Value	
-1	contextString
Highlight	false
Value	There are
-2	contextDim
Dim	Day
Use Words	false
-3	contextString
Highlight	false
Value	days in the dataset:
-1	contextRow
-Items	
-0	contextString
Value	
-1	contextDimElement
Dim	Day
First or Last	FIRST
-2	contextString
Highlight	false
Value	to
-3	contextDimElement
Dim	Day
First or Last	LAST
owner	parent
size	(320,60,0)

4. Complete the **Context** and **Items** fields in the **Editor**. See *Query language syntax* on page 515 for guidelines about syntax.
5. Save the file.

The values in the preceding example result in a status formula displayed as follows:



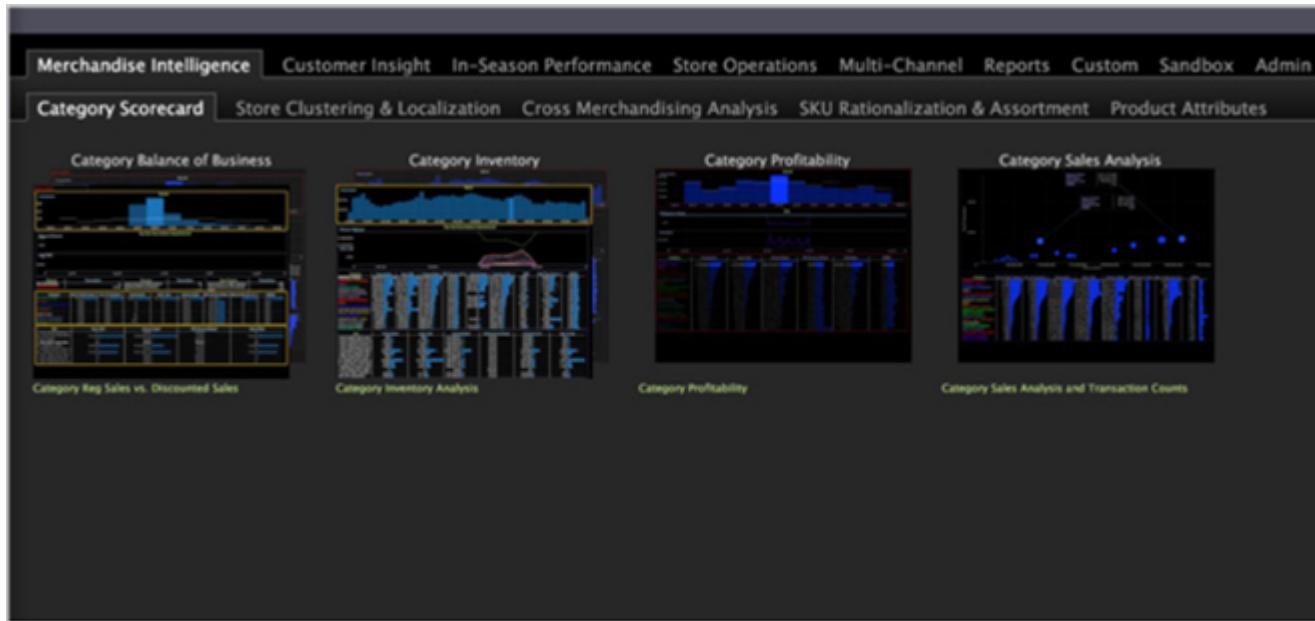
Display workspaces

Navigation title:

A workspace is a container for a particular analysis and visualization task.

A workspace can contain multiple visualizations, all of which operate against a common set of data (that is, all visualizations in the workspace render the same set of query results). When you perform a filtering operation in one visualization, the selected subset is reflected throughout the entire workspace.

Below is an example of several workspace thumbnails on the **Worktop**.



Note: Because your implementation of Data Workbench can be fully customized (that is, because the tabs, workspaces, and menu items that appear in your implementation might differ from what is documented in this guide), it is important to note that by right-clicking almost anywhere, you can view a list of options related to the workspace or visualization in which you are working.

You analyze your data by combining different types of visualizations in a workspace. Powerful queries are created by making selections from the data displayed in those visualizations, the results of which are presented visually within the workspace across all visualizations. You use workspaces to control the scope of each query:

every selection made within a visualization changes the current query and affects the queries in other visualizations in the workspace. See [Make selections in visualizations](#) on page 281.



Attention: After opening up a new workspace, you may need to click **Add > Temporarily Unlock** to unlock the screen.

Display a workspace

Navigation title:

After working in Data Workbench, you can display one of the many pre-configured workspaces delivered with Data Workbench.

The sidebar is a persistent feature of the workspace, unless you hide it. See [Configure the sidebar](#) on page 209.

To display an existing workspace

On the desired **Worktop** tab, click the thumbnail of the workspace you want to display.



Note: After opening a new workspace, you may need to click **Add > Temporarily Unlock**.

Below is an example of an open workspace.



To scroll within a workspace

With your cursor outside of the workspace, use the scroll wheel on your mouse to scroll through the workspace.

Create a workspace

Navigation title:

You can customize existing workspaces by adding or modifying existing visualizations or creating new workspaces.

You can then save these workspaces for later use. See [Create a workspace](#) on page 214.

You can create new workspaces by creating a blank workspace, copying and pasting an existing workspace, or saving a workspace to another location or using another name.

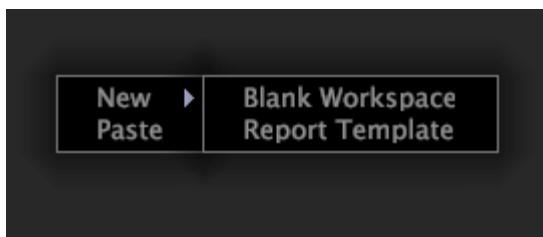
This section describes the following tasks:

- [Adding New Workspaces](#) on page 214
- [Copying and Pasting Existing Workspaces](#) on page 215
- [Changing a Workspace Name](#) on page 216

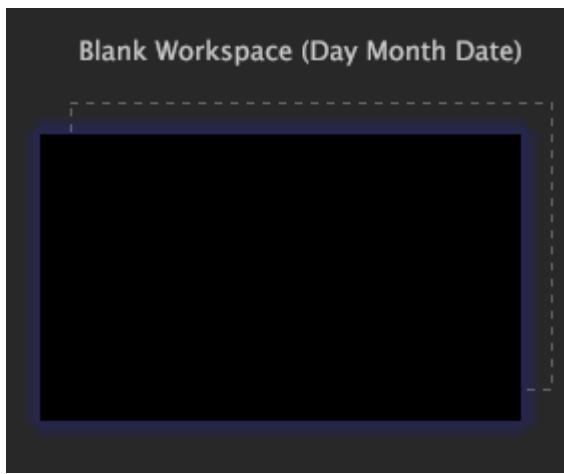
Adding New Workspaces

On the desired **Worktop** tab, right-click a blank area within the tab and click one of the following options:

- **New > Blank Workspace.** This option displays a standard size, blank workspace.
- **New > Report Template.** This option displays a standard size template that you can use to create reports.



The thumbnail for the new type of workspace appears on the tab.



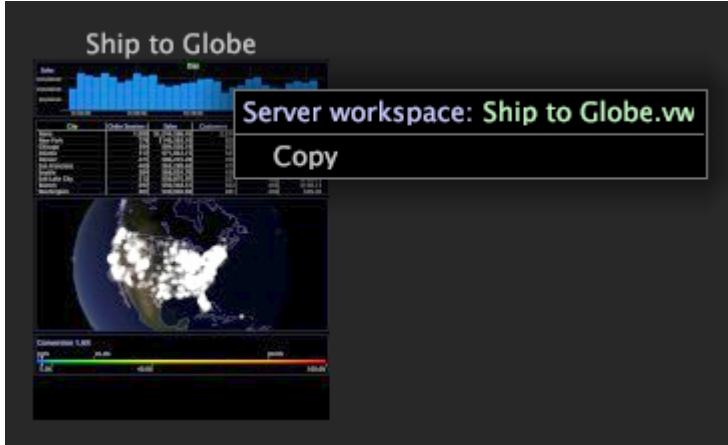
Note: After opening a new workspace, you may need to click **Add > Temporarily Unlock**.

Copying and Pasting Existing Workspaces

You can copy and paste an existing workspace onto to any other workspace. If you paste the workspace on a tab that contains a server version of a workspace with the same name, the workspace is pasted as a local version of that workspace. If you paste the workspace on a tab that contains server and local versions of a workspace with the same name, the workspace is pasted as a user version of that workspace with the same name followed by (Copy). If you paste the workspace on a tab that does not contain a workspace with the same name, the workspace is pasted as a new user workspace.

To copy and paste an existing workspace

1. On the desired **Worktop** tab, right-click the workspace that you want to copy and click **Copy**.

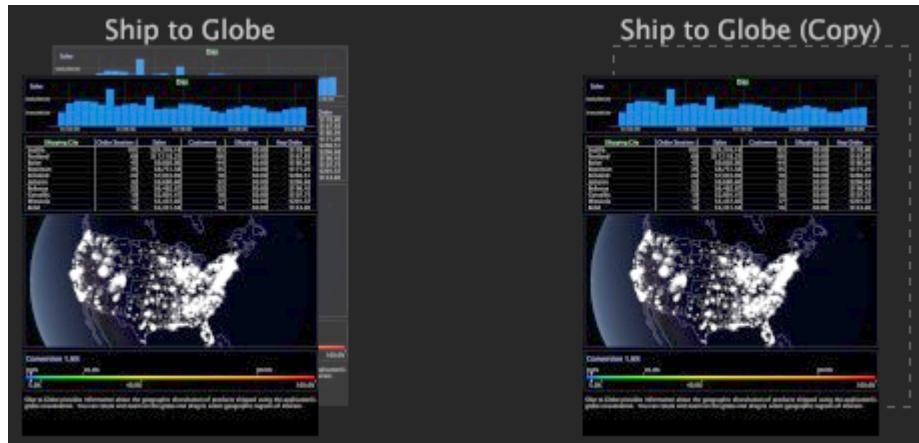


2. Go to the tab in which you want to paste the workspace, right-click a blank area within the tab, then click **Paste**.

The following example is a workspace pasted on a workspace tab that contains a server version:



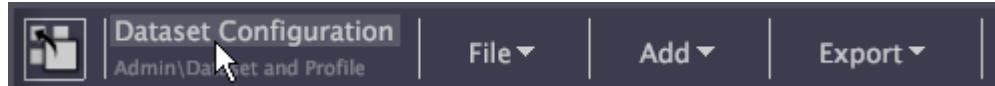
The following example is a workspace pasted on a workspace tab that already contains a server and local version. The workspace is pasted as a user version of that workspace with the same name followed by (Copy):



Changing a Workspace Name

To change the name of a workspace

1. On the desired **Worktop** tab, click the thumbnail of the workspace you want to display.
2. In the workspace, click the title text in the title bar and type the new title, as shown in the following example:



Deleting a Workspace

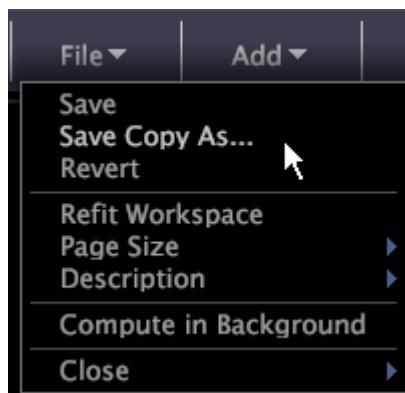
To change the name of a workspace

Workspace File menu

Navigation title:

Information about the menu options available in the Workspace file menu.

In the workspace, click **File**.



The following table provides descriptions of each menu item.

Save

Saves the workspace. See [Save a workspace](#) on page 218.

Save Copy As

Saves the workspace with a different name or in a different location. See [Save a workspace](#) on page 218.

Revert

Reverts to the last saved version of the current workspace.

Refit Workspace

Resizes visualizations to the display resolution you are using. This is helpful to quickly adjust your display for presentations.

Page Size

Sets a working page size for a workspace. You can select Full Screen, Standard, or any of the specific sizes for computer screens, printed pages, and reports. Full Screen fills your particular screen and Standard is a setting that can be configured to fit your organization's standard screen size. The default Standard setting fills the screen when using 1024 x 768 resolution.

Workspaces that are copied, saved as .png files, or printed also use this page size for output. Page sizes larger than the Data Workbench visualization use scroll bars, while smaller sizes are centered on the screen and display a light gray border around the workspace.

Description

Enables you to create or edit a textual description of the workspace. This text appears on the **Worktop** below the thumbnail. See [Add a description to a workspace](#) on page 222.

Compute in Background

(Appears only when working online.) Keeps the queries in the selected workspace running in the background while you continue working. When selected, the thumbnail displays the following information, which indicates the progress of the queries:

- Working: *n%* - indicates that the query is processing and the percentage of the processing that is complete.
- *n MB* Query Load - total size of the query result. Query Load is proportional to the total memory load of your Data Workbench server, but does not correlate directly. As a guideline, a 10 MB or higher query load may strain your system. The query load listed does not take clustering into account.



Note: If Compute in Background remains selected, the queries in the selected workspace become standing queries, continuing to be updated and use memory load. Make sure to clear the selection for Compute in Background when you are finished working in the workspace.

Close

Closes the workspace. Click **Close > Save** to save the changes you have made in the workspace, or click **Close > Don't Save** to return to the **Worktop** without saving the changes you have made in the workspace.

You also can save your changes, close the workspace, and return to the **Worktop** using any of the following methods:

- Click the Data Workbench logo in the upper-left corner of the workspace.
- If your mouse has navigation buttons, click the back button on your mouse.

You also can void your changes by closing the workspace without saving by pressing **<Ctrl> + <Backspace>**.

Export to Excel from Template

If Microsoft Excel is installed on the Data Workbench machine, automatically starts Microsoft Excel and exports the data from certain visualizations, certain legends, and text annotations to the template Excel file (.xls or .xlsx) that you select. See [Export to a template Excel file](#) on page 221.

Refit a workspace

Navigation title:

Refit Workspace resizes visualizations to the display resolution you are using.

In a Workspace, click **File > Refit Workspace**.

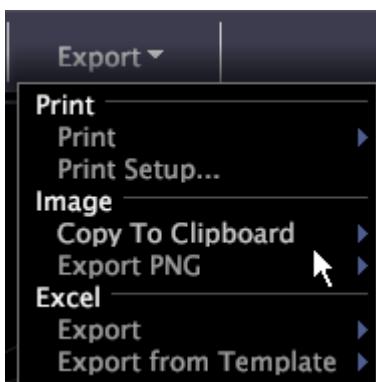
This is helpful to quickly adjust your display for presentations.

Share a workspace

Navigation title:

You can use the **Export workspace** menu to print and export workspaces to .png files or to Excel.

In a workspace, click **Export**.



Workspace Export Menu

Menu Option	Description
Print	Prints the workspace. When you print a workspace, the background prints as white. If you are printing to a color printer, the workspace prints in color on a white background. If you are printing to a black and white printer, the workspace prints in grayscale on a white background. Printed output is centered horizontally on the page by default.
Print Setup	Sets the page orientation (portrait or landscape) and margins.
Copy to Clipboard	Copies the workspace or sidebar to the clipboard. The following options are available when copying workspaces, sidebars, or both: <ul style="list-style-type: none">• Dark Background option copies the workspace as displayed.• Light Background option copies the elements of the workspace in color and displays them on a white background.• Monochrome option copies the elements of the workspace in grayscale and displays them on a white background.
Export PNG	Saves a snapshot of a workspace in Portable Network Graphic format (.png file). See Export a workspaces as a PNG file on page 220.
Excel	If Microsoft Excel is installed on the Data Workbench machine, automatically starts Microsoft Excel and exports the data from certain visualizations, certain legends, and text annotations to a new Excel file (.xls or .xlsx). See Export workspace data to Microsoft Excel on page 220.
Export (Excel) from Template	If Microsoft Excel is installed on the Data Workbench machine, automatically starts Microsoft Excel and exports the data from certain visualizations, certain legends, and text annotations to the template Excel file (.xls or .xlsx) that you select. See Export to a template Excel file on page 221.

Save a workspace

Navigation title:

By default, closing an unlocked workspace saves any changes that have been made to the workspace.

If the workspace is a server workspace, your changes are saved only locally unless you specifically save the updated workspace to the Data Workbench server. For more information about locked workspaces, see [Unlocking a workspace](#) on page 222.

Save a workspace locally

The default save location is the `User\profile name\Workspaces\tab name` folder within the Data Workbench installation directory. For example, if you are working with the Movies profile and you save a workspace locally from the **Custom** tab, the workspace is saved to the `User\Movies\Workspaces\Custom` folder in your Data Workbench installation directory.

To save changes to a workspace

- In the workspace, click **File**, then **Save**.

To save an existing workspace as a new workspace

1. On the desired **Worktop** tab, click the thumbnail of the workspace you want to display.
2. In the workspace, click **File**, then click **Save Copy As**.
3. In the **Save Workspace As** dialog box, specify the name and location where you want to save the copied workspace and click **Save**.

Save a workspace to the Data Workbench server

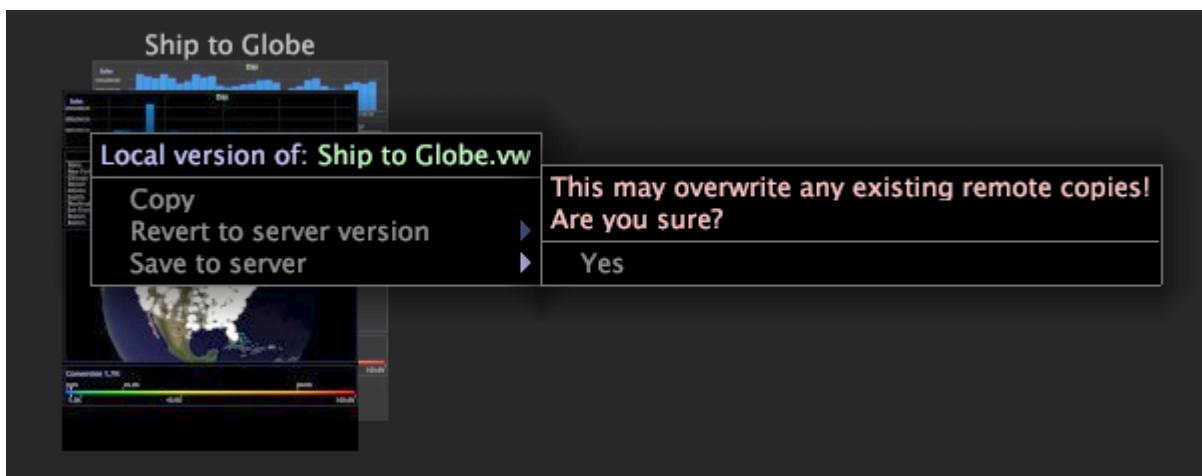


Note: Only users with the appropriate permissions can save workspaces to the Data Workbench server. For more information, contact your system administrator.

Saving workspaces to the connected Data Workbench server is also referred to as publishing a workspace because it makes the workspace available to other users. By default, workspaces are saved to the `working profile name\Workspaces\tab name` folder of the Data Workbench server. For example, if you are working with the Movies profile and you save a workspace to the connected Data Workbench server from the **Custom** tab, the workspace is saved to the `Movies\Workspaces\Custom` folder of the Data Workbench server.

To save a workspace to the Data Workbench server

- On the desired **Worktop** tab, right-click the thumbnail of the workspace that you want to save to the Data Workbench server and click **Save to server**.



Export a workspace

Navigation title:

You can export a workspace as a .png image file or export the data from certain windows to an Excel (.xls or .xlsx) file.

Export a workspaces as a PNG file

You can save a snapshot of a workspace in Portable Network Graphic format (.png files). The following color options are available when saving workspaces as .png files:

- **Black background** copies the workspace as displayed.
- **White background** copies the elements of the workspace in color and displays them on a white background.
- **White background (B&W)** copies the elements of the workspace in grayscale and displays them on a white background.

To export a workspace as a .png file

In the title bar menu of a workspace, click **Export > Export PNG > <color option>**.

The **Save Image As** dialog box appears.

Navigate to the directory in which you want to save the file, change the name of the file if necessary, and click **Save**.

Export workspace data to Microsoft Excel

When exporting a workspace to Excel, Data Workbench exports data from certain visualizations, dimension and value legends, and text annotations to a new Excel workbook with one visualization per worksheet.

To export workspaces and individual windows to Microsoft Excel, the following requirements must be met:

- Microsoft Excel must be installed on the same machine as Data Workbench.
- The user account under which the Data Workbench process is running must have permission to access Microsoft Excel.



Note:

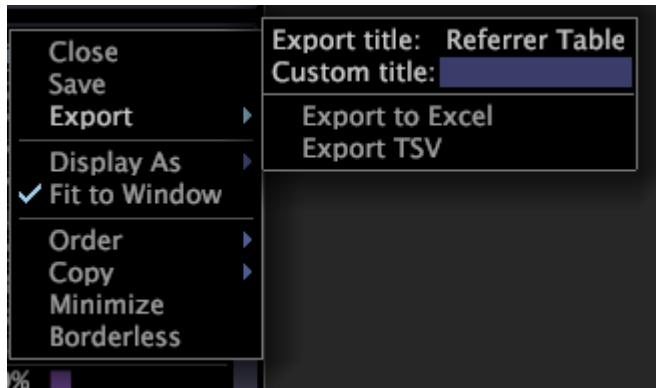
- When you export data as Excel files, you are opening a new instance of Excel. For more information about this process, see <http://support.microsoft.com/kb/257757>.
- Although Data Workbench supports more than 256 columns and 65,536 rows of data, versions of Microsoft Excel prior to 8.0 do not.

If these requirements are met, Data Workbench automatically starts Microsoft Excel and export the data to a new Excel workbook. Data is not exported from the following visualizations: graphs, path browsers, process maps, scatter plots, and globes.

Apply custom titles

Unless you have specified a Custom title for the window on the **Export** menu, the **Export title** listed (for example, City Table) is used as the worksheet name.

1. Right-click the top border of the window and click in the **Custom title** field.
2. Type the title that you want to apply to the window.



Note: If you enter a hyphen (-) in the **Custom title** field, this visualization is not exported with the workspace.

When you export the workspace to Excel, the worksheet containing the data for this window is named using the title that you specified instead of the title in the **Export title** field.

Export a workspace or sidebar to Excel

To export workspace data to a new .xls or .xlsx file

1. In title bar of the workspace, click **Export > Export**.
2. Specify whether to export the workspace, sidebar, or both.

Export to a template Excel file

You can export data in your workspace to a template Excel file (.xls or .xlsx). Using a template file can reduce the amount of time that you spend formatting your data each time the workspace is exported.



Note: This template file must be an .xls or .xlsx file, not an .xltx file.

When the data is exported, the existing tabbed sheets in the template (each representing one visualization) are repopulated with the most recent data from the workspace, while any new windows that are not present in the template as tabbed sheets are ignored. Any other tabbed sheets in the template file remain unchanged.

In addition, if you have a macro defined in the template Excel file that you would like to run automatically when the report is generated, name the macro “VSExport.”

Let's say that you want to use exported campaign data from a table visualization in a pie chart on another tabbed sheet in an Excel file and want to update this information every week. You can use a template so that you do not have to recreate your references from the table's tabbed sheet to the pie chart's tabbed sheet each time you want to update the data. The table data is updated upon export, which automatically updates the pie chart.

To export workspace data to a template .xls or .xlsx file

1. Right-click the title bar of the workspace and click **Export > Export to Excel from Template**.
 2. Specify whether to export a workspace, sidebar, or both.
- The **Select a template worksheet** dialog box opens.
3. Complete one of the following steps as appropriate:
 - If you are using a .xls template file:

-
1. Browse to and select the template .xls file.
 2. Click **Open**.
 - If you are using a .xlsx template file:
 1. Browse to the location of the template file. The .xlsx file name is not displayed.
 2. In the **File name** field, type *.xlsx and click **Open**. All .xlsx file names display in the file list.
 3. Select the template .xlsx file.
 4. Click **Open**.

Import a workspace

You can import a workspace to your worktop.

To import a workspace

1. Right-click in the Worktop (outside of a workspace thumbnail view) to open a menu.



2. Click **Import** to open a file-selection window.
3. Select a .vw file and press **Open** or **Open as read-only**.



Note: Visualizations cannot be imported like workspaces. When you right-click in the worktop and select **Import**, you can import an existing workspace but not a visualization residing outside of the workspace. If you attempt to import a visualization, you will receive a "not a valid workspace" error.

Add a description to a workspace

Navigation title:

Steps to add a description to a workspace.

1. On the desired **Worktop** tab, click the thumbnail of the workspace that you want to display.
2. Click **File > Description**.
3. Click in the **Edit workspace description** field and type a description.

When you save and close the workspace, the description displays below the workspace thumbnail.

Unlocking a workspace

Navigation title:

Administrators can configure Data Workbench to allow only certain users to change certain workspaces.

While a workspace is locked, users can make selections in most visualizations and sort the data in tables but cannot make other changes to the workspace.

Certain users have the ability to temporarily unlock a locked workspace using the **Temporarily Unlock** option on the workspace title bar menu, make the necessary changes, then save the workspace. The next time that user opens the workspace, it is locked again.

If you cannot make and save changes to a workspace and the **Temporarily Unlock** option is not available in the **File** menu, this workspace has been locked by your Adobe application administrator and you cannot change it. Please contact your administrator for more information.

To lock or unlock a workspace

In the title bar of the workspace, click **File > Temporarily Unlock**.



Note: For information about configuring individual workspaces and entire tabs or folders to be locked, see [Configure a locked workspace](#) on page 496.

To delete server workspaces, you must use the **Profile Manager**. For more information, see [Delete files from your working profile](#) on page 460.

Deleting a workspaces

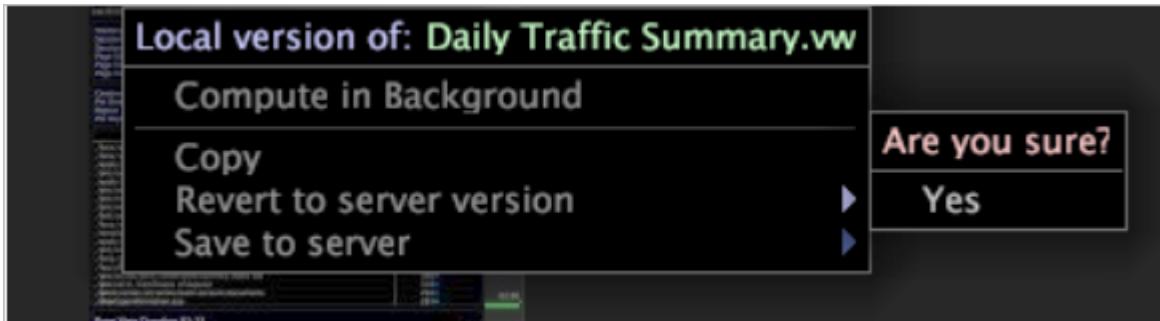
Navigation title:

Information about deleting workspaces from the **Worktop**.

A local workspace is a local version of a workspace that exists on the server. A user workspace exists only on the local machine.

To delete local workspaces from the Worktop

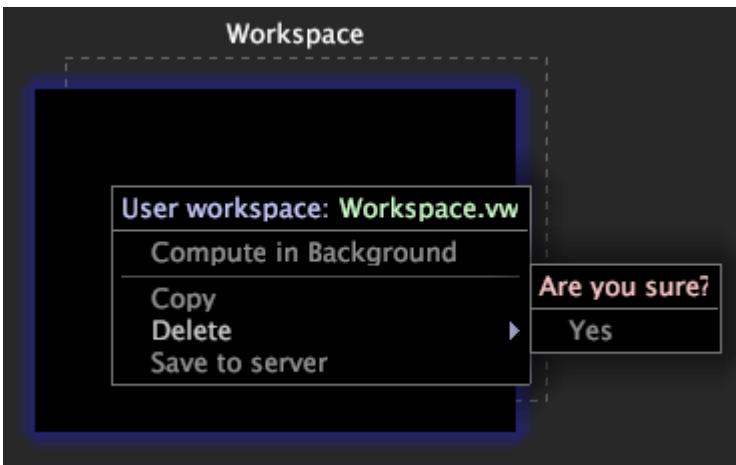
Right-click the workspace that you want to delete and click **Revert to server version > Yes**. The local version is deleted, but the server version remains.



Note: When you have a local version of a server workspace, you must lose your changes by reverting to the server version before you can download an updated version of the server workspace.

To delete user workspaces from the Worktop

Right-click the workspace that you want to delete and click **Delete**.



Using the workspace Window menu

Navigation title:

This menu enables you to add available visualizations to your workspace.

Right click in the workspace window (anywhere in the background of your workspace).



 **Note:** Because your implementation of Data Workbench can be fully customized, the menu items that appear may differ from what is documented in this section.

The following table contains descriptions of each menu item.

Workspace Window menu

Menu Option	Description
Table	Adds a table to your workspace based on the selected dimension. See Visualizations on page 279 and Tables on page 414.
Visualization	Adds a visualization to your workspace based on the selected visualization type (for example, a graph or table) and the selected dimension(s). See Visualizations on page 279 and Analysis visualizations on page 304.
Legend	Adds a legend to your workspace. See Legends on page 378.
Note	Adds an image, text, or dynamic title window to your workspace. See Annotations on page 308.
Open	<ul style="list-style-type: none">The File option enables you to navigate to and open an existing visualization.The From Clipboard option enables you to open a window that you have previously copied. This is equivalent to pasting a window that you have copied.
Admin	Heading for the administrative interfaces. See Administrative interfaces on page 446.
Data	<ul style="list-style-type: none">(Available only with Site) The Broken Session Filter option enables you to filter out all visitors with only one page view. Because most of these tend to be robots, spiders, or users who have turned off cookies, more accurate analysis can be done by filtering these sessions out of the analysis. You can toggle the filter on and off by clicking Data > Broken Session Filter. When the Broken Session Filter is active, an X appears to the left of the Broken Session Filter menu option. The Broken Session Filter is on by default, but its state is maintained separately for each workspace that you load; therefore, you can turn it on or off for each workspace.The Subset option enables you to select the data subset that you want to view by clicking Data > Subset > <subset name>. See Subsets on page 289.The Auto Generate Report Time option keeps your Report Time metric in sync with the dataset. It is enabled by default.

Keyboard shortcuts for the Data Workbench Client

Use these shortcut keys for workspaces and configuration files.

Keyboard shortcuts from the workbench thumbnail view

Ctrl + N

Opens a new workspace.

If you are limited to creating a single type of blank workspace, then the new workspace will open immediately.

If you have rights to open multiple types of workspaces based on your implementation, then Ctrl + N displays a menu allowing you to choose the type of workspace to be created.

Ctrl + O

Imports a workspace

Resize and refit workspaces between window and full page views

Commands	Quick Keys	Combined menu commands
Full screen view. Workspace fills the screen and refits to the new size.	Ctrl plus Ctrl + (on keypad) <i>or</i> Ctrl Shift + (on keyboard)	• File > Page Size > Fill Screen <i>followed by</i> • File > Refit Workspace
Window view. Workspace displays in a standard window view and refits to the new size.	Ctrl minus Ctrl -	• File > Page Size > Standard <i>followed by</i> • File > Refit Workspace

General keyboard shortcuts for text

Action	Result
Ctrl + X	Cuts text
Ctrl + Z	Undo action for previously typed characters
Ctrl + Shift + Z	Redo action for typed characters
Ctrl + A	Selects all

Keyboard shortcuts within a workspace

Action	Result
Ctrl + P	Prints a workspace.
Ctrl + W	Closes a workspace without saving.
Ctrl + S	Saves a workspace.
Ctrl + O	Open a visualization file within a workspace. This is the same feature accessed by clicking Add > Open from the workspace menu.
Ctrl + E	Open a configuration (.cfg) file in a workspace. By default the .cfg files appear, but you can also open other .dim, .metric or .vw files as needed (when you select the All Files option).

Action	Result
Add > Open and Add > Edit quickly open recently opened or pinned files in a workspace.	Open files recently accessed (the last ten files are save) or files saved or "pinned" to the new menu. You can click the pin icon next to the menu item to save the link in the menu.

Keyboard shortcuts within an open configuration (.cfg) file

Action	Result
Right-click a node and click Paste	Pastes a node into a .cfg file
Right-click the vector and click Paste as child	Pastes a vector into a .cfg file at a vector level

Working with windows in a workspace

Navigation title:

The windows in Data Workbench (visualizations, legends, annotations) display meaningful information about your data.

Manipulate windows in a workspace

Navigation title:

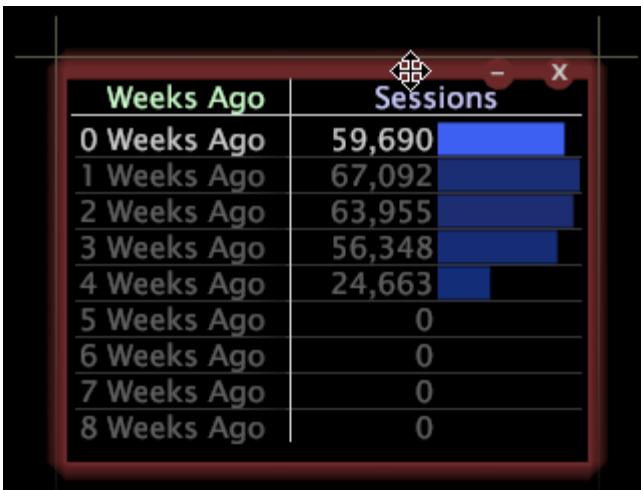
The types of windows available depend on the profile and application with which you are working.

To add a window

1. Right-click in the workspace and click the desired menu option.
2. From the menus that appear, select the type of window that you want to view.

To move a window

Rest the mouse pointer on the top border of the window until it becomes a four-pointed arrow, then click+drag it to the desired location in your workspace.



To resize a window

Rest the mouse pointer on any side or corner of the window until it becomes a two-pointed arrow, then click+drag the window corner or side until the window is the desired size.



To copy and paste a window

You can copy and paste windows within and between workspaces.

1. Right-click the top border of the window and click **Copy > Window**.
2. In the desired workspace, click **Open > From Clipboard**.

To copy and paste a window as a callout for a visualization

1. Right-click the top border of the window and click **Copy > Window**.
2. In the desired visualization, right-click the desired element and click **Add Callout > From Clipboard**.

Work with text in windows

Navigation title:

Wherever you can enter text in a window, you can perform many of the functions that you can perform in any text editor, such as Notepad or Microsoft Word.

Data Workbench supports basic editing features, including cut (Ctrl+x), copy (Ctrl+c), paste (Ctrl+v), undo (Ctrl+z), redo (Ctrl+Shift+z), select section (click+drag), and select all (Ctrl+a).

In addition to basic copying and pasting of text, you can copy and paste text within and between configuration files (.cfg).

To copy and paste parameters and nodes within a configuration file (.cfg)

When copying and pasting a parameter, both must be of the same type, such as string to string.

1. Right-click a node or parameter label and click **Copy**.
2. Complete one of the following steps:
 - Right-click the vector and click **Paste as child**. The node is pasted above the first node in the vector.
 - Right-click the node below which you want to paste the node and click **Paste**.

To copy and paste nodes into a text editor

1. Right-click a vector or node and click **Copy**.
2. In the appropriate location within the text editor file, press **Ctrl+v**.

Window Controls menu

Navigation title:

The **Window Controls** menu enables you to control an individual window.

Right-click the top border of the window.



These menu options vary by window type. The table in the following procedure contains descriptions of the basic menu items available for most types of windows. Additional options available for particular types of windows are addressed in their respective sections in this guide.

Window Controls

Menu Option	Description
Close	Closes the current window.
Save	Saves the visualization (as a .vw file) in its current state. The dialog box that displays enables you to save the visualization using a different name or to a different location. The default location is the User\profile name\Work folder within the Data Workbench installation directory.
Display As	(Available for tables only.) Displays the data in a table as numbers, bars, or both at the same time. See Change the table display on page 417.
Display	(Available for graphs only.) Displays the data in a graph as lines or bars. See Change a graph display on page 364.
Fit to Window	Allows the contents of the window to automatically expand or shrink to fit the window each time you resize it. See Image annotations on page 310 and Fit data to a window on page 416.
Order	Organizes the layers of windows in the workspace. Order commands include the following: <ul style="list-style-type: none">• Front: Brings the window to the front of the display.• Normal: Leaves the window in the position in which it opened.• Back: Places the window behind any other windows open in the current display. An X appears to the left of the active option.
Copy	Copies individual windows to the clipboard for use in third-party applications. After you have copied a window to the clipboard, it can be pasted into common office applications including Microsoft Word, Microsoft PowerPoint, or a Microsoft Outlook email message. To copy a window, right-click its top border, click Copy , and click the appropriate option: <ul style="list-style-type: none">• Black background copies the workspace as displayed.• White background copies the elements of the workspace in color and displays them on a white background.• White background (B&W) copies the elements of the workspace in grayscale and displays them on a white background.
Minimize	Reduces the overall size of the window, while all of the window's contents remain visible but are not being updated. Right-click the minimized window to restore it to its original size.
Borderless	Removes the borders around the selected window. This command is very useful when creating reports.

Export window data

Navigation title:

You can export the data in certain windows to an Excel file (.xls or .xlsx) or to a tab separated values file (.tsv).

Data is not exported from graphs, path browsers, process maps, scatter plots, and globes.

Export window data to Microsoft Excel

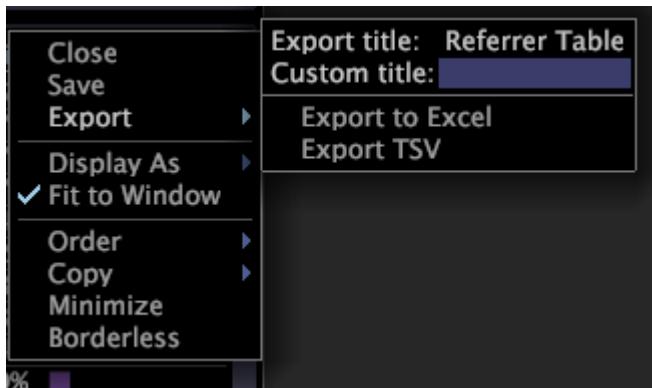
To export individual window data to Microsoft Excel, the following requirements must be met:

- Microsoft Excel must be installed on the same machine as Data Workbench.
- The user account under which the Data Workbench process is running must have permission to access Microsoft Excel.
- When you export data as Excel files, you are opening a new instance of Excel.
- Although Data Workbench supports more than 256 columns and 65,536 rows of data, versions of Microsoft Excel prior to 8.0 do not.

If these requirements are met, Data Workbench automatically starts Microsoft Excel and export the data to a new Excel workbook when you select the **Export To Excel** menu option.

To export window data to an .xls or .xlsx file

Right-click the top border of the window and click **Export > Export to Excel**.



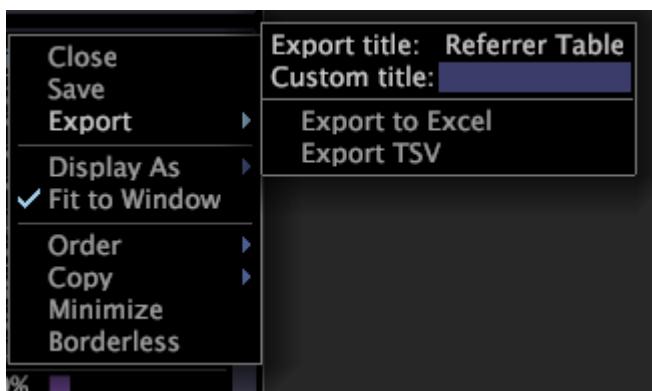
Excel opens a new workbook containing the exported data. Unless you have provided a Custom title (as described in the following section), this workbook is named using the **Export title** (Day Table in the example above).

Apply custom titles

If you provide a custom title for a window (using the **Custom title** field on the **Export** menu) the worksheet to which Data Workbench exports the data is named using this custom title instead of the title in the **Export title** field (Day Table in the example above).

To apply a custom title to a visualization

1. Right-click the top border of the window and click in the **Custom title** field.
2. you



When you export the visualization to Excel, the worksheet containing the data for this window is named using the title that you specified instead of the title in the **Export title** field.

Export window data to a TSV file

To export a window to a .tsv file

Right-click the top border of the window and click **Export > Export TSV**.

1. The **Save Window** dialog box appears.

-
2. Navigate to the directory in which you want to save the file. Change the name of the file if necessary, and click **Save**.

Exporting data using segment export

Navigation title:

The segment export feature enables you to create a segment of the elements of any countable dimension, then output data for that segment on a batch or ongoing real-time basis into a tab-delimited file.

Each time that you export a segment, you output metric or dimension data for all of the dimension elements included in that segment. You can control how the output data is formatted so that other systems can easily load the data. For information about configuring a segment for export, see [Configure segments for export](#) on page 234.

Segment export

You can easily create a **Segment Export** definition from the Detail Table visualization in the Data Workbench Client.

In addition, **Segment Exports** automatically combine their results to a single server, rather than producing partial results on each DPU that you must combine using an external process. You can create a segment export file, save it to the **Profile Manager**, and upload the output file to a server of your choice.

To configure the segment export server

The **Segment Export** feature creates a single output file on the segment export server, rather than separate output files created on each DPU. The segment export server is usually configured to run on the FSU.

In the Dataset\ directory in the **Profile Manager**, open the Segment_Export.cfg in Workstation, and specify your server's address. (Your address could be an IP or fully qualified domain name.):



This is IP of the Data Workbench server receiving the results of the segment export. This is a one-time setup. If the Segment_Export.cfg is not present, exports do not run.

To configure export directories

For security purposes, executables or batch files that run after a segment export must reside in the configurable Scripts\ directory of the segment export server.

The .part and final output must reside in the configurable Exports directory. The command to run exists in Command and Command Arguments. Instances of the %file% in the Command Arguments will be replaced with the path of the output file.



Note: New to Data Workbench 5.4, the \Exports folder is created automatically. Previous export directories set up before version 5.4 required an Exports\ prefix before the filename for each segment export. Adding this prefix is now redundant.

1. In Communications.cfg on the destination server for **Segment Exports**, add a SegmentExportServer to the list of servers. (Example shown in red).

```

component = CommServer:
  Access Control File = Path: Access Control\Access Control.cfg
  IP Interface = string:
  Port = int: 80
  SSL Port = int: 443
  Servers = vector: 23 items
    0 = InitServer:
      Client Type = string: Sensor
      URI = string: /SensorInit.vsp
    1 = LoggingServer:
      Flush Interval = int: 10000
      Log Directory = string: Logs\\
      URI = string: /Submit.vsp
    2 = SendServer:
      URI = string: /
...
    19 = NormalizeServer:
      URI = string: /Cluster/
    20 = FileServer:
      Local Path = string: Archive\\
      URI = string: /Archive/
    21 = SegmentExportServer:
      URI = string: /SegmentExportServer/
      Scripts Directory = string: Scripts\\
      Exports Directory = string: Exports\\
    22 = SourceListServer:
      URI = string: /SourceListServer/

```

Exports Directory: Specifies where to put .part and output files. This can be a shared directory.

Scripts Directory: Specifies the directory from where all executable or batch files are run.

2. Access Control.cfg, on the same server, add read-write access to the URI /SegmentExportServer/ to the Cluster Servers AccessGroup:

```

...
4 = AccessGroup:
Members = vector: 3 items
0 = string: IP:192.168.5.128
1 = string: IP:192.168.5.129
2 = string: IP:192.168.5.130
Name = string: Cluster Servers
Read-Only Access = vector: 8 items
0 = string: /Components for Processing Servers/
1 = string: /Addresses/
2 = string: /Profiles/
3 = string: /Lookups/
4 = string: /Users/
5 = string: /Access Control/
6 = string: /Bin/
7 = string: /Logs/
Read-Write Access = vector: 3 items
0 = string: /Cluster/
1 = string: /SegmentExportServer/
2 = string: /SourceListServer/
...

```

3. Change your .export files:

```

Query = SegmentExportQuery:
Command = string: flipcommand.bat      <-- optional name of command to run from
Command Arguments = string: "%file%"   <-- optional arguments for command, %file% is replaced by output file
Filter = string: My_Filter
Level = string: rating
Output File = string: My_Segment_test.txt
Output Format = string: %=Average_Score by User%\t%Rating by User%\t%Score%
Schedule End Time = string:
Schedule Every = string:
Schedule Start Time = string:
Time Limit (sec) = double: 1800

```

4. For each profile, the Segment Export.cfg is located in the Dataset\ directory, with the following contents:

```

Segment Export = SegmentExport:
Segment Export Server = serverInfo:
Port = int: 80
Address = string: 192.168.5.128 (for example) Use SSL = bool: false

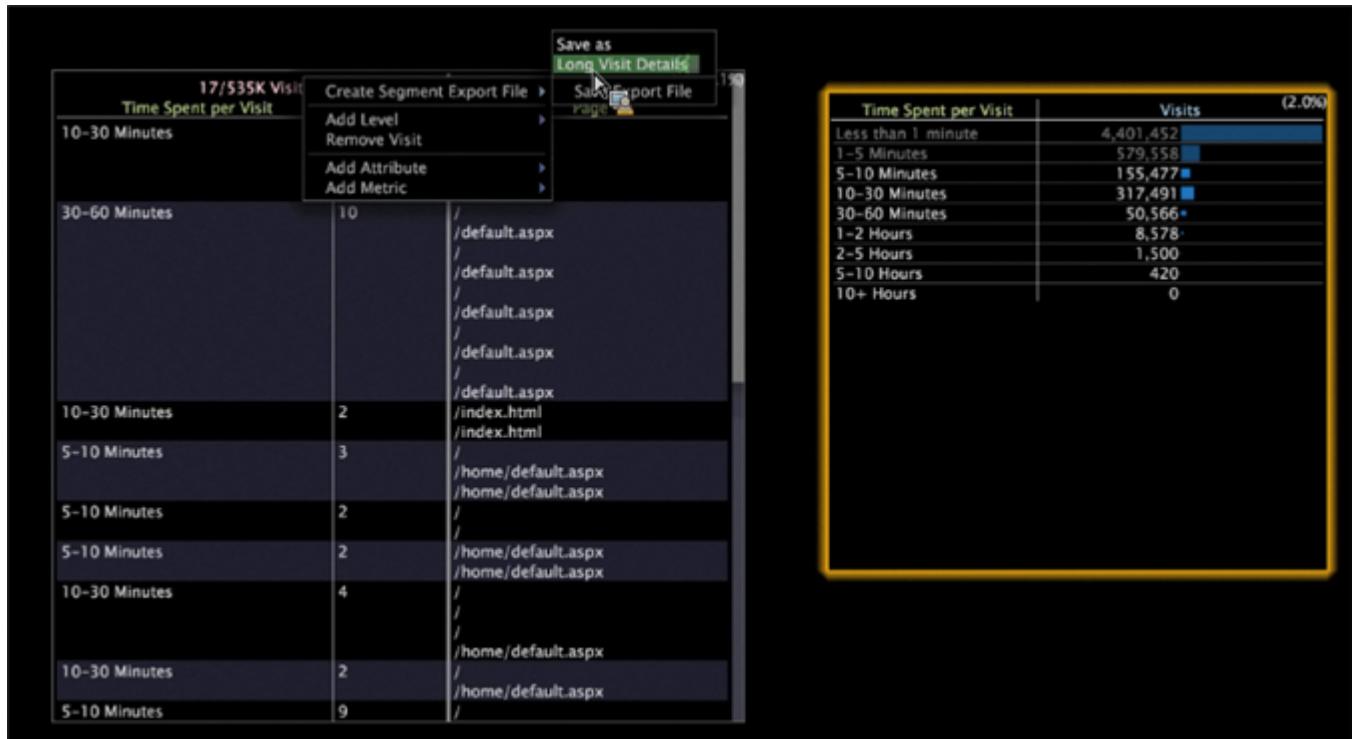
```

5. Ensure that the directories referred to in Exports Directory and Scripts Directory exist.

Only executables and batch files in the Scripts directory can be run as the command of a segment export.

To create a segment export file

1. In a workspace, create a Detail Table showing subsets of data (Visualization > Detail Table) and add attributes.
2. If desired, make selections in the workspace. (Any selections or filters are applied to the export.)



3. In the Detail Table header, right-click and select **Create Segment Export File**.

4. In **Save as**, type a name for the .export file.

5. On the .export file, configure the parameters as necessary.

Any selections or filters in the workspace are incorporated in the export file.

6. Save the .export file.

The saved file displays in the **Profile Manager** for you to save to the server. When you save the file to the server, the export begins.

Output format

Navigation title:

Guidelines for specifying the output format.

- Each metric or dimension name must start and end with a percent sign (%).
- %XYZ% specifies the element of dimension XYZ corresponding to the element of Level. An error is generated if dimension XYZ is not one-to-one or one-to-many with Level.
- %=XYZ% specifies the value of metric or metric formula XYZ for the given element of Level.
- Dimension names that are two words or longer do not require underscores.
- Metric names that are two words or longer require underscores.



Note: If you hold down the Ctrl key and right-click within the **Output Format** field, an **Insert menu** appears. This menu contains a list of special characters (for example, Tab) that often are used as delimiters.

If you want to export session duration data for your segment, you must create a new metric based on the Session Duration metric. The new metric, which is for use only with the **Output Format** field of segment export, enables Microsoft Excel to correctly interpret sessions lasting less than one hour. To create a new session duration metric,

open the Session Duration.metric file (from the **Profile Manager**) and insert a pound sign (#) into the ftime string: ftime = string: %#H:%M:%S

The pound sign causes a leading “0” to be appended to session durations of less than 1 hour. As a result, Excel interprets 0:53:21 as 53 minutes and 21 seconds. Save the new metric with a different name and upload it to the appropriate profile for others to use by right-clicking the check mark for the file in the **User** column and clicking **Save to > <profile name>**.

Configure segments for export

Navigation title:

You can create a segment of the elements of any countable dimension, then output data for that segment on a batch or ongoing real-time basis into a tab-delimited file.

Each time that you export a segment, you output metric or dimension data for all of the dimension elements included in that segment. You can control how the output data is formatted so that other systems can easily load the data.



Note: You cannot export reporting dimensions, because they use a report_time.metric file for reference. As a workaround, if you place a hard-coded report_time.metric in the profile, the segment export can use it as a reference point for reporting dimensions. However, the report_time.metric does not automatically update based on the profile's As Of Time, so when you want to change the reporting dimension reference, you must change the hard-coded report_time.metric file.

To configure a segment for export, you must open and edit a .export file.

1. In the **Profile Manager**, click the **Export** directory in the **File** column to show its contents.

If the Export directory does not exist, create it as follows:

1. Navigate to your Data Workbench installation directory.
 2. Open the directory for the profile with which you are working.
 3. Within the Profile directory, create a new directory named “Export.”
2. In the **Profile Manager**, right-click the empty cell in the **User** column for the Export directory, then click **Create > New Segment Export**.

A file named New_Segment_Export.export appears in the **File** column for Export.

3. Rename the new file by right-clicking in the **User** column for the file and typing the new name in the File parameter.
4. Open the new file by right-clicking in the **User** column for the file and clicking **Open > from the workbench**.

The configuration window for the .export file appears.

5. Click **Query**, then modify the fields of the .export file as described in the following table:

For this parameter...	Provide this information...
Command	Optional. A program to be executed after the Output File is created. This field must reference an executable (an .exe file), not a shell command.  Note: The segment export will fail if there is a space in the command parameter.
Filter	Optional. A named filter or a filter expression. You can either create a named filter using a filter editor, then type the name of that filter here, or you can type a filter expression itself.

For this parameter...	Provide this information...
	<p>For more information about filter editors, see Filter editors on page 348. For more information about filter expression syntax, see Syntax for filter expressions on page 521.</p> <p>Elements of Level that match the filter are exported, while all other elements are not.</p>
Level	<p>The countable dimension whose elements are to be exported.</p> <p>Example: A level of Visitor exports one row of data for each visitor.</p>
Output File	<p>Path and file name of the exported data. If the profile is running on a Data Workbench server cluster, each Data Workbench server writes an Output File containing a portion of the data.</p> <p>The Data Workbench server installation directory contains an Exports directory where you can save the output file. For example, you could enter <code>Exports\Visitor Segment.txt</code>, where <code>Visitor Segment.txt</code> is the name of the file containing the exported data.</p>
Output Format	<p>The metric or dimension data to be exported for each Level element. If the output is a tab-delimited file, the fields should be separated by Tab characters, and the format should end with the appropriate new-line characters. For more information, see Output format on page 233.</p>
Schedule End Time	<p>Optional. The end date and time for the schedule, including the time zone.</p> <p>Format: YYYY-MM-DD hh:mm time zone</p> <p>Example: 2013-08-01 12:01 EDT</p> <p>Scheduled exports stop at this time; however, the Output File is still regenerated any time its definition is changed. This field is meaningless without defining Schedule Every. For more information about time zone settings, see the Dataset Configuration Guide.</p>
Schedule Every	<p>Optional. The frequency at which to regenerate the Output File. Supported values are hour, day, week, and month. The Output File is still regenerated any time its definition is changed.</p>
Schedule Start Time	<p>Optional. The start date and time for the schedule, including the time zone.</p> <p>Format: YYYY-MM-DD hh:mm time zone</p> <p>Example: 2013-08-01 12:01 EDT</p> <p>Scheduled exports start at this time, and the schedule is relative to this time. This field is meaningless without defining Schedule Every. For more information about time zone settings, see the Dataset Configuration Guide.</p>
Time Limit (sec)	<p>Optional. The maximum time permitted to elapse while a segment export is being generated. If the specified interval is exceeded, then the export starts over. Setting this value to 0 (zero) removes the limit. The default value is 600 seconds.</p>

6. Right-click (**New**) at the top of the window, then click **Save**.
7. To make this file available to all users of the working profile, right-click the check mark for the created `.export` file in the **User** column, then click **Save to > <working profile name>**.



Note: Saving the `.export` file to the Data Workbench server causes the export to run once immediately, even if the Schedule Start Time is set to a future date and time.

The following is a sample `.export` file.

-Visitor Segment.export	
-Query	SegmentExportQuery
Command	cmd.exe /c wc -l "Exports\Visitor Segment.txt > Exports\Count.txt"
Filter	Visitor_Segment
Level	Visitor
Output File	Exports\Visitor Segment.txt
Output Format	%TrackingID%[TAB]%Visitor Segment%[TAB]%=Sessions%[CR][LF]
Schedule End Time	
Schedule Every	week
Schedule Start Time	2007-01-01 00:00 EST
Time Limit (sec)	1800



Note: The Visitor Segment.export file shown in the sample refers to the Visitor Segment filter. Modifying the definition of this filter changes the definition of the export.

Create dimensions for use with segment export

Navigation title:

Any data that you want to export must be defined as a dimension within the profile.

If the dimension does not already exist in the profile, you must create it. You can create dimensions using any of the following methods:

- Add a dimension to the Transformation.cfg file. See the *Dataset Configuration Guide*.
- Create a new .dim file. See [Work with derived dimensions](#) on page 463.
- Make selections in a visualization such as a process map or a segment and save the selections as a dimension. See [Save a dimension from a process map](#) on page 400 and [Create a segment dimensions](#) on page 413.

Exporting a field of data such as Tracking ID or Email Address (which can have lots of elements) requires that you create a denormal dimension that stores the raw strings of data.

For more information about denormal dimensions, see the *Dataset Configuration Guide*.

Segment Export with Custom Headers

Create custom column export headers for your segment export files to add easily understood descriptions for exported segments. This export feature also lets you output as TSV and CSV files.

New functionality has been added to Segment Export, including the ability to export with a header, or in CSV and TSV formats.

You can create column headers for your export files.

Creating a New Segment Export

1. Open a workspace and right-click **Tools > Detail Table**.
2. Right-click and select **Add Level > Extended** > Choose an item.
3. Right-click title and select **Add Attribute**.

Select a dimension from the menu.

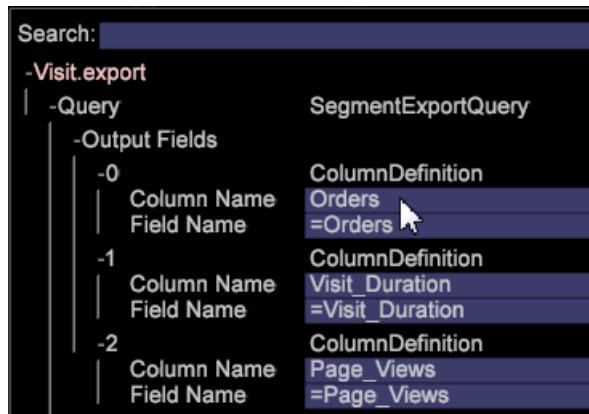
4. Right-click title and select **Add Metric**.

Select a metric from the menu.

-
5. Right-click title and select **New Segment Export**.



New Segment Export with Header automatically populates the Column Name with the name of the metric.
New Segment Export requires you to set a custom name.

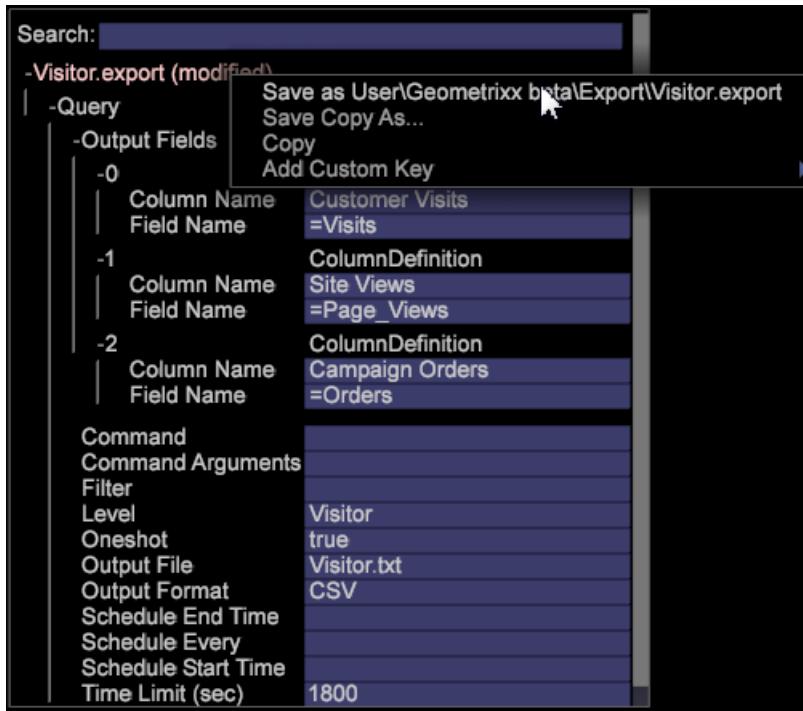


Note: The Column Name field cannot be left empty or the header will not be present.

6. Right-click and name the segment and then click **Save Export File**.

An export window will open.

7. Right-click the export name and click **Save as <export filename>**.



- Right-click Admin > Profile Manager > Expand Export. Find the export file you just created and save it to an existing profile.



Export a segment using S/FTP delivery

You can now use CSV, TSV, Segment Export, and Segment Export with Header using FTP and SFTP protocols to export segment files from the client (workstation) to the server.

Setting up S/FTP Export configuration files

To set the export configuration, two new export configuration files were added to set up an FTP or SFTP connection, allowing the Server details to be picked from the *FTPServerInfo.cfg* file and the credentials will be picked from *FTPUserCredentials* folder (corresponding to the Server Name given in the command arguments).

- Set the **FTPServerInfo.cfg** file.

Enter the FTP server information and set connection retries allowed from the workstation. Edit from the workstation or the server at `Server\Addresses\Export\FTPServerInfo.cfg` file.

```
FTP Servers = vector: 1 items
  0 = ftpServerInfo:
    Address = string:
    Name = string:
    Port = int: 21
Connect Retries = vector: 1 items
  0 = connectServerRetries:
    Retries = int: 0
    Server Name = string:
```

- Set the **FTPUserCredentials.cfg** file.

Enter user credentials to connect to servers using the `Server\Admin\Export\FTPUserCredentials.cfg` file. This file contains user credentials needed to connect to servers and can only be edited from server and not from workstation (client).

```
FTP User Credentials = vector: 1 items
  0 = ftpUserCredInfo:
    User Name = string:
    User Password = EncryptedString:
    Server Name = string:
    Public Key Path = string:
    Private Key Path = string:
    Passphrase = EncryptedString:
```



Note: Ensure that SSH keys you generate for authentication are in the format identical to those that are generated when you use SSH Keygen command.

Example for generating SSH keys using keygen:

```
ssh-keygen -t rsa -b 4096 -C "<label>"
```

There are six parameters in the **FTPUserCredentials.cfg** file required for various FTP or SFTP transfers.

1. *User Name*
2. *User Password*
3. *Server Name*
4. *Public Key Path*
5. *Private Key Path*
6. *Passphrase*

Protocol	Parameters
FTP	Set parameters 1, 2, 3.
SFTP using password authentication	Set parameters 1, 2, 3 when transfer uses password authentication (-p in the command arguments).
SFTP using key authentication	Set parameters 1, 2, 3, 4, 5, 6 when transfer uses key authentication (-k in the command arguments).

Setting the FTP and SFTP Export Commands

1. Open an export table.

From the Workstation, right-click a *Detail Table* and choose one of the export types—CSV , TSV, Segment Export, or Segment Export with Header. Or open the .export file from a command-prompt and edit (see [Configure segments for export](#) on page 234).

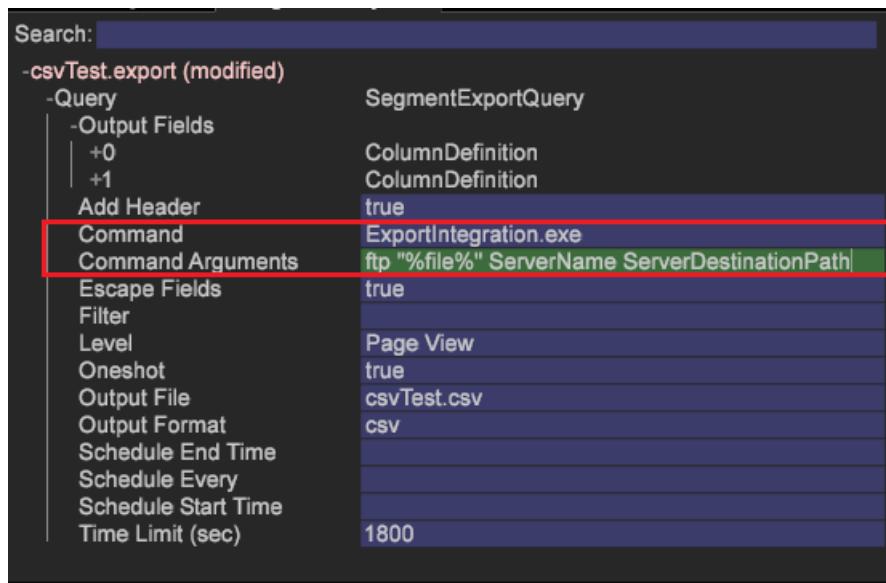
2. In the *Command* field, set it to point to the export executable:

```
ExportIntegration.exe
```

3. Set the *Command Arguments* fields as shown below for the protocol and authentication required:

FTP

```
Command Arguments set to ftp "%file%" ServerName ServerDestinationPath
```

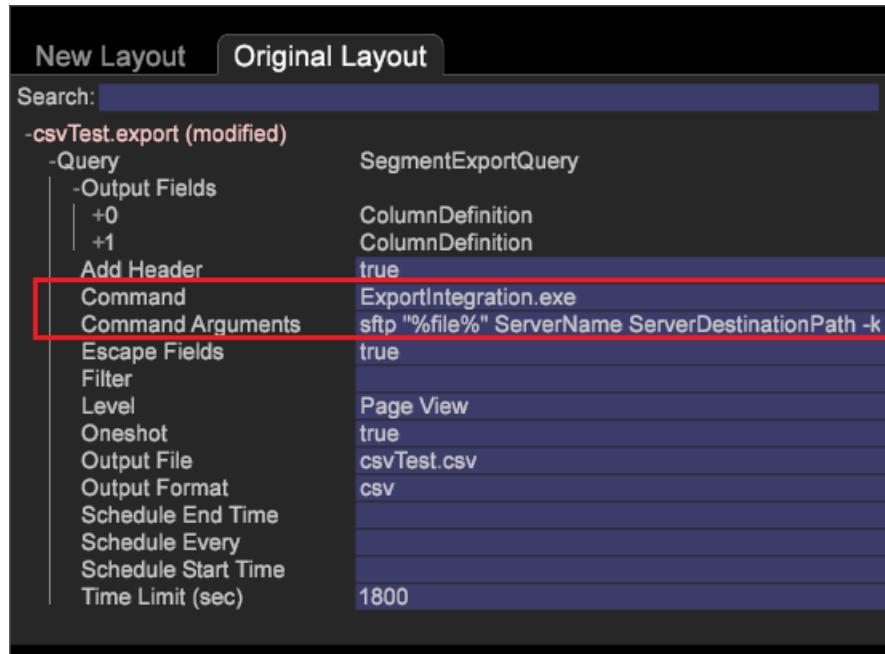


SFTP (if using password for authentication)

```
Command Arguments set to sftp "%file%" ServerName ServerDestinationPath  
-p
```

SFTP (if using keys for authentication)

```
Command Arguments set to sftp "%file%" ServerName ServerDestinationPath  
-k
```



All Command Arguments are mandatory and need to be entered as shown.

S/FTP export using private/public keys

To implement the FTP and SFTP Export using private and public keys, place the configuration files in these folders:

- Place **FTPServerInfo.cfg** in the Server/Addresses/Export/ folder.
- Place **FTPUserCredentials.cfg** in the Server/Admin/Export/folder.

Six parameters are included in the **FTPServerInfo.cfg** file:

1. *User Name*
2. *User Password*
3. *Server Name*
4. *Public Key Path*
5. *Private Key Path* -- Place the private key path in the configuration file without the extension, for example:
 Private Key Path = string: E:\\Server\\campaign\\campaignprivatekey
6. *Passphrase*

FTP uses parameters 1,2, and 3.

SFTP uses parameters 1,2, and 3 when the transfer uses password authentication.

SFTP uses all six parameters when the transfer is done using key authentication. For example, if you are using keys for authentication:

```
'Command Arguments' = sftp "%file%" ServerName ServerDestinationPath -k
```

The configuration files need to be in the correct location.



Note: The public keys need to point to a **.pem** file and not to a folder location. You can create keys using an SSH key generation function from applications like Cygwin. (Putty generates keys in a **.ppk** format that is not supported.)

Profiles and Audiences Export

Data Workbench lets you export files to integrate with Profiles and Audiences as part of an integrated Adobe Experience Cloud.

Profiles and Audiences is part of the *Experience Cloud Identity Service*, a core service of the Adobe Experience Cloud. The Profiles and Audiences export allows audiences to be shared across the Experience Cloud using a unique Experience Cloud ID (ECID) that is assigned to every visitor and then used by Audience Manager. The ExportIntegration.exe application E:\Server\Scripts is employed to generate both Profiles and Audiences, and Adobe Target exports.

Configuring the FSU Server to use Profiles and Audiences

1. Access your FSU server.
2. Open the MMPExport.cfg file. (Server/Admin/Export/MMPExport.cfg).
3. Enter values in the all fields as required. For example:



Note: MMP/AAM integration relies on Amazon's s3 bucket for data transfer.

The s3 information required for MMP (s3) transfer can be obtained from Audience Manager team.

Sample MMPExport.cfg

```
MMP Export Configuration = MMPExportConfiguration:  
s3 Bucket = string: aws_bucket_for_mmp  
s3 Object Directory = string: test/files/  
s3 Region = string: us-east-1  
s3 Access Key = string: ZZKI62005YBA  
s3 Secret Key = string: ioqwa3OpNE5  
data Provider Name = string: 895  
client ID = string: mcprofile2-test  
client Secret = string: saea1287617212987q  
username = string: mmptest  
password = string: pass  
numRecordsPerChunk = int:  
numThreads = int:  
maxRetriesOnSendFailure = unsigned int:
```



Note: The MMPExport.cfg file also lets you take all records, split them into sets, and create chunks of records. The chunks of records are then exported to Amazon S3. Three mandatory parameters are required to create chunks of records: numRecordsPerChunk, numThreads, and maxRetriesOnSendFailure.

Definition of Parameters

Parameter	Definition
s3 Bucket	The AWS S3 bucket where the export is transferred to.
s3 Object Directory	A path to save s3 files. This supports sub-directories.  Important: Space and multibyte characters are not allowed in the path and will create errors in the export. (The hyphen is allowed).

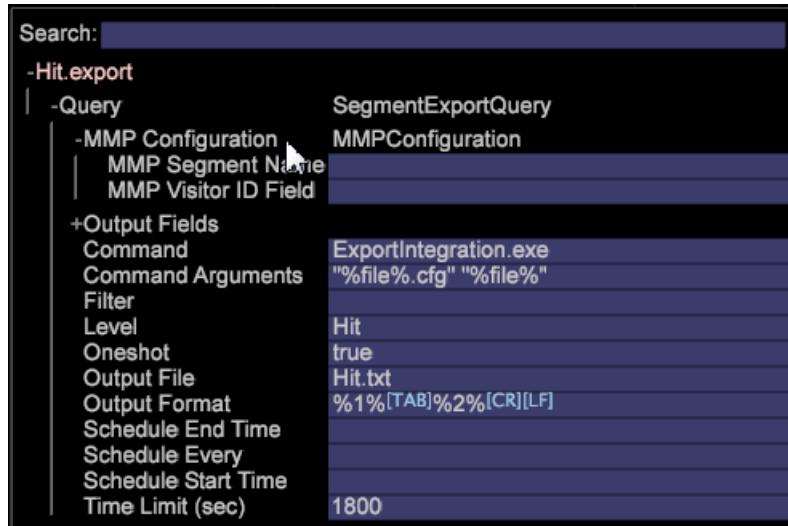
Parameter	Definition
<i>s3 Region</i>	The AWS s3 Region where the export is sent to. Ex. us-east-1
<i>s3 Access Key</i>	AWS s3 Access Key
<i>s3 Secret Key</i>	AWS s3 Secret Key
<i>data Provider Name</i>	This will be the folder name that is used for storing segments and traits in AAM respectively. This should be unique per customer.
<i>client ID</i>	This is a unique client ID provided to a customer when provisioned for MMP.
<i>client Secret</i>	This is a unique client secret provided to a customer when he/she is provisioned for MMP.
<i>username</i>	MMP username
<i>password</i>	MMP password
<i>numRecordsPerChunk</i>	Determines the chunk size in terms of number of records. The implementation clips the user specified value to min = 1000 records (~50 KB chunks) and max = 50000 records (~2.5 MB chunks). A default value of 10000 is used in case the user does not specify this configuration property.
<i>numThreads</i>	Determines the parallelism of the chunk sending part. It accepts a value between 1 to 24 threads, and its default value is 12 threads.
<i>maxRetriesOnSendFailure</i>	Determines the number of retry attempts to be made in case of chunk send failures. Default value is 0 specifying no retries. Sleep interval of 2 seconds is used between retries.

Generating MMP export from the client

1. From the client, open a workspace and right-click **Tools > Detail Table**.
2. Add **Level**.
3. Right-click the header and select **Add Attributes**.
4. Right-click the header and select **New Master Marketing Profile Export**.



5. Expand Query.



6. Expand MMP Configuration.

7. (required) Enter the **MMP Segment Name** and **MMP Visitor ID Field**. These parameters cannot be left empty.
8. The **MMP Segment Name** should match the Segment ID defined in the MMP.
9. The **MMP Visitor ID** is the attribute column defined in step 4 that corresponds to the **Visitor ID**.
10. Once these fields are entered, you can save the export by right-clicking the header for the export and choose **Save as "User\export"**.
11. Open **Admin > Profile Manager** and save the export to the profile.

If all data is entered correctly, this will generate an export file in the FSU (Server/Exports) and it will also transfer the export to the AWS using the information in `MMPExport.cfg`. The log for this is provided in `Server/Trace/`. e.g., `MMP-102014-133651-[Segment Export Name].log`

```
Query = SegmentExportQuery:  
Command = string: ExportIntegration.exe  
Command Arguments = string: \"%file%.cfg\" \"%file%\"  
Filter = string:  
Level = string: Page View  
MMP Configuration = MMPConfiguration:  
MMP Segment Name = string: 12345  
MMP Visitor ID Field = string: Tracking ID  
Oneshot = bool: true  
Output Fields = vector: 3 items  
0 = ColumnDefinition:  
Column Name = string:  
Field Name = string: Tracking ID  
1 = ColumnDefinition:  
Column Name = string:  
Field Name = string: PID  
2 = ColumnDefinition:  
Column Name = string:  
Field Name = string: SID  
Output File = string: MMPTest.txt  
Output Format = string: %1%\t%2%\t%3%\r\n  
Schedule End Time = string:  
Schedule Every = string:  
Schedule Start Time = string:  
Time Limit (sec) = double: 1800
```

Configuration Details	Description
MMP Segment ID	Required. This is an identifier you would define first in Audience Manager.
MMP Visitor ID Field	Map the MCID.

Export to Adobe Target

Export Data Workbench data to Adobe Target using `TargetBulkUpload.exe` from the Detail Table.

Data Workbench lets you export files to integrate with Adobe Target as part of an integrated Adobe Experience Cloud.

The `TargetBulkUpload` file is found in the `Server\Scripts` folder in the server installation files. The executable has retry logic, as well as additional logic to optimize performance.

You can modify the `TargetBulkUpload.cfg` file and move it to `Server/Admin/Export` folder before running the upload script. For example, you can set a Max Timeout Interval to 720 minutes (default) to timeout the upload after the specified period.

How it works

After the data gets successfully sent to Target, the status of the upload is continuously monitored. If the upload succeeds, a success message is logged. If the upload fails or is pending, the monitoring continues. You can configure the timeout interval in the `TargetBulkUpload.cfg` file. If the upload gets stuck at Target, a message is logged and the status can still be monitored.

There are two log files generated in the trace for the triggered export under `/server/Trace/`:

- `targetbulkuploadexportname.log`
- `targetbulkuploadexportname.log.completed`

The targetbulkuploadexportname.log file has the detailed status for all the records of multiple batches, the edge server they are going to, and the status (successful, failed, profile not found, status unknown, and stuck). In case any batch is found to be stuck, the batch is not processed any further. A stuck batch URL is available to track the status. See the following example data from the targetbulkuploadexportname.log.completed file:

```
1205057 total rows
568740 successful
62 failed
28964 profile not found
112169 unknown status
492339 stuck status
```

The stuck status value is incremented with the total stuck batch size regardless of how many uploads are successful or failed. The total rows value is also incremented by the same number of stuck batch size.



Note: Previously, DWB data was exported using the ExportIntegration.exe. Currently only the MMP, CRS, and S/FTP exports are used with this executable. Adobe Target integration now uses the TargetBulkUpload.exe in Data Workbench.

Segment export wizard

Export segments using the segment export wizard

The segment export wizard provides a step-by-step process to configure and export segments rather than *exporting segments from a detail table*.

Export segments using the wizard

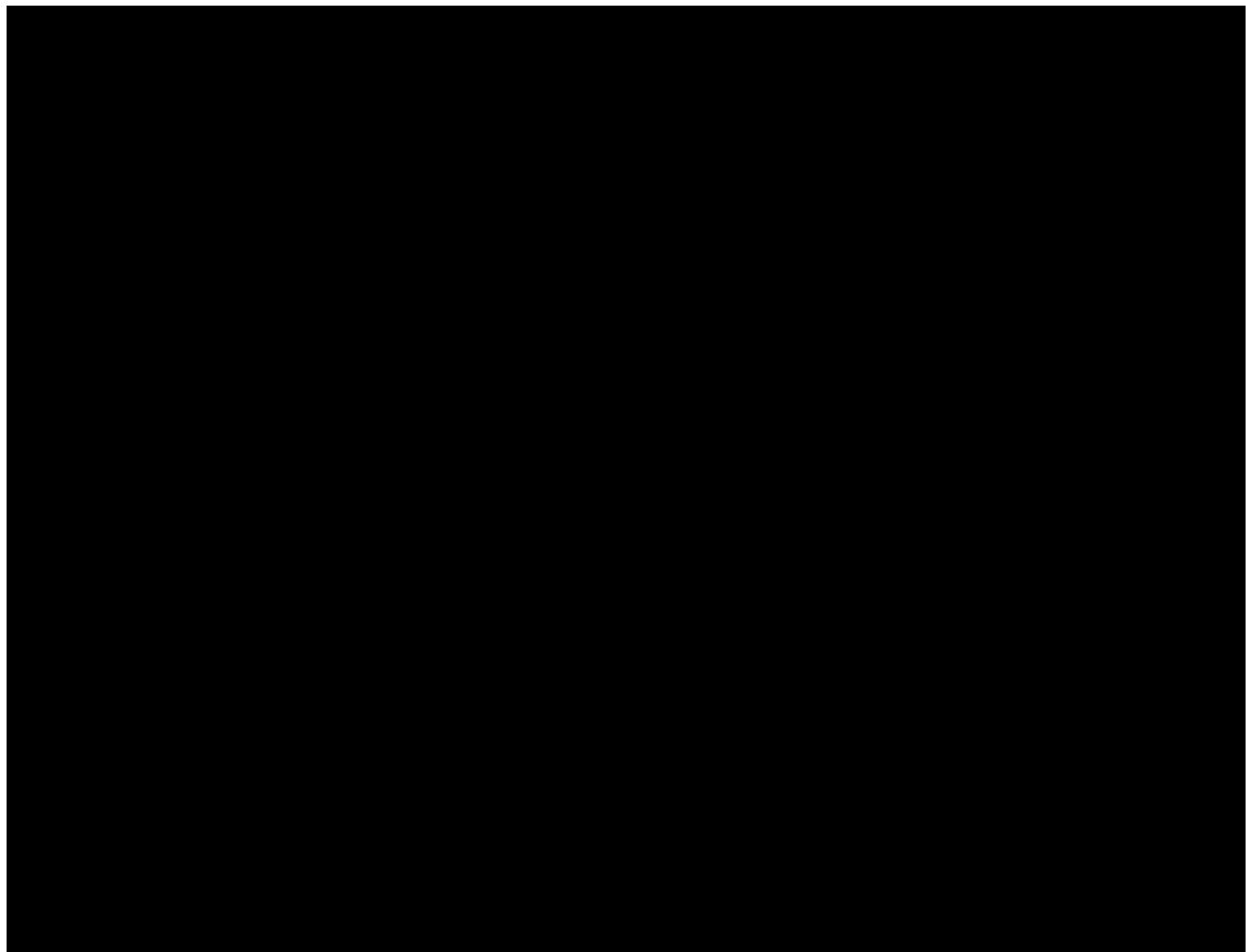
To open the wizard, right-click in a workspace and select **Admin > Wizards > Segment Export Wizard**.



Note: Only the segments applied before opening the wizard will be captured. Also, the segment exports created from the wizard cannot spawn external commands.

1. Select the various parent levels of the dimensions and metrics to be added to your export.

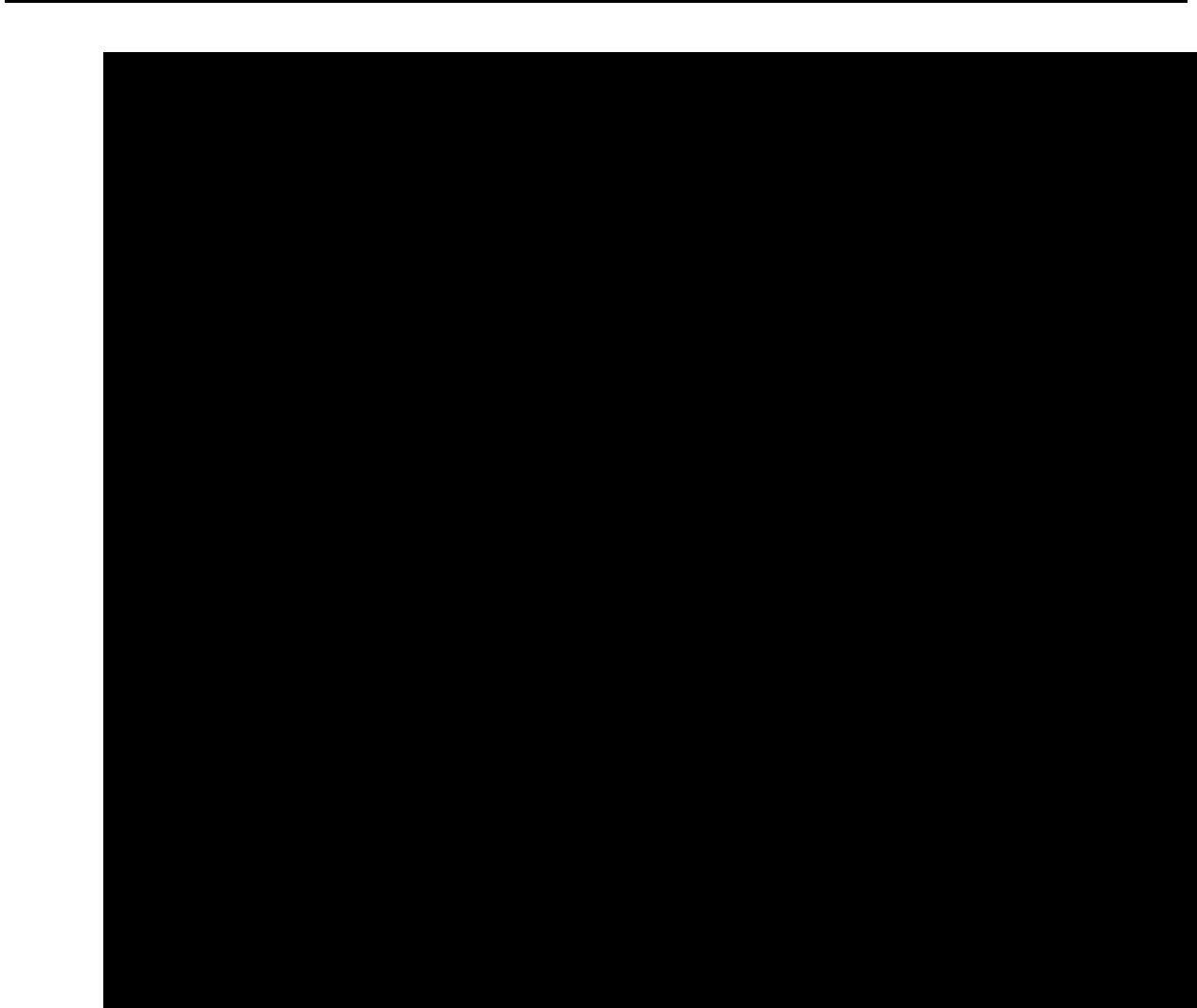
The levels displayed depend on the profile selected. You can select multiple dimension levels based on the profile.

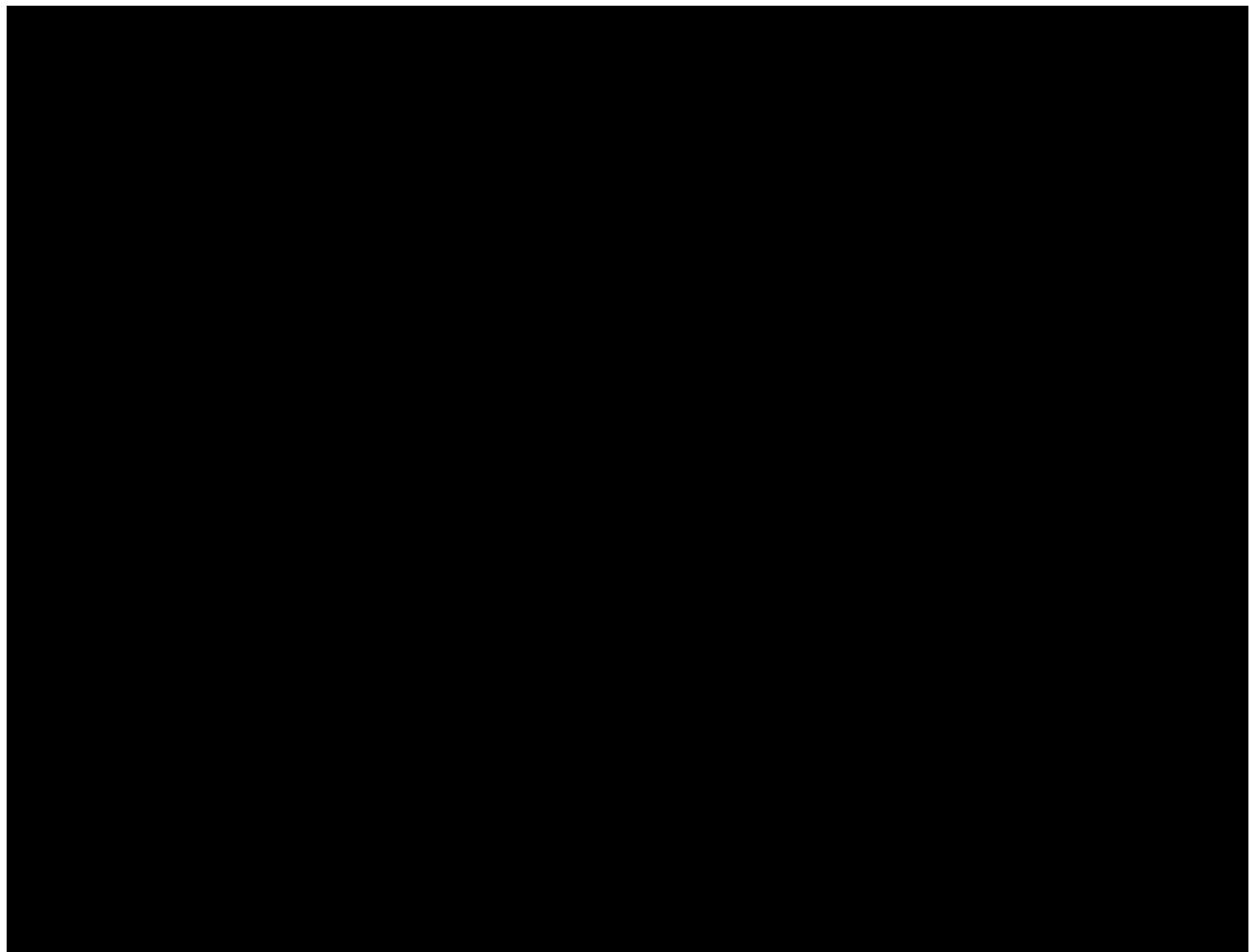


2. Click **Next**.
3. Select the Dimension and Metrics for the selected levels.

For example, after selecting Page View as the parent level, you can select the child dimensions and metrics available to be exported.

4. Click **Next**.

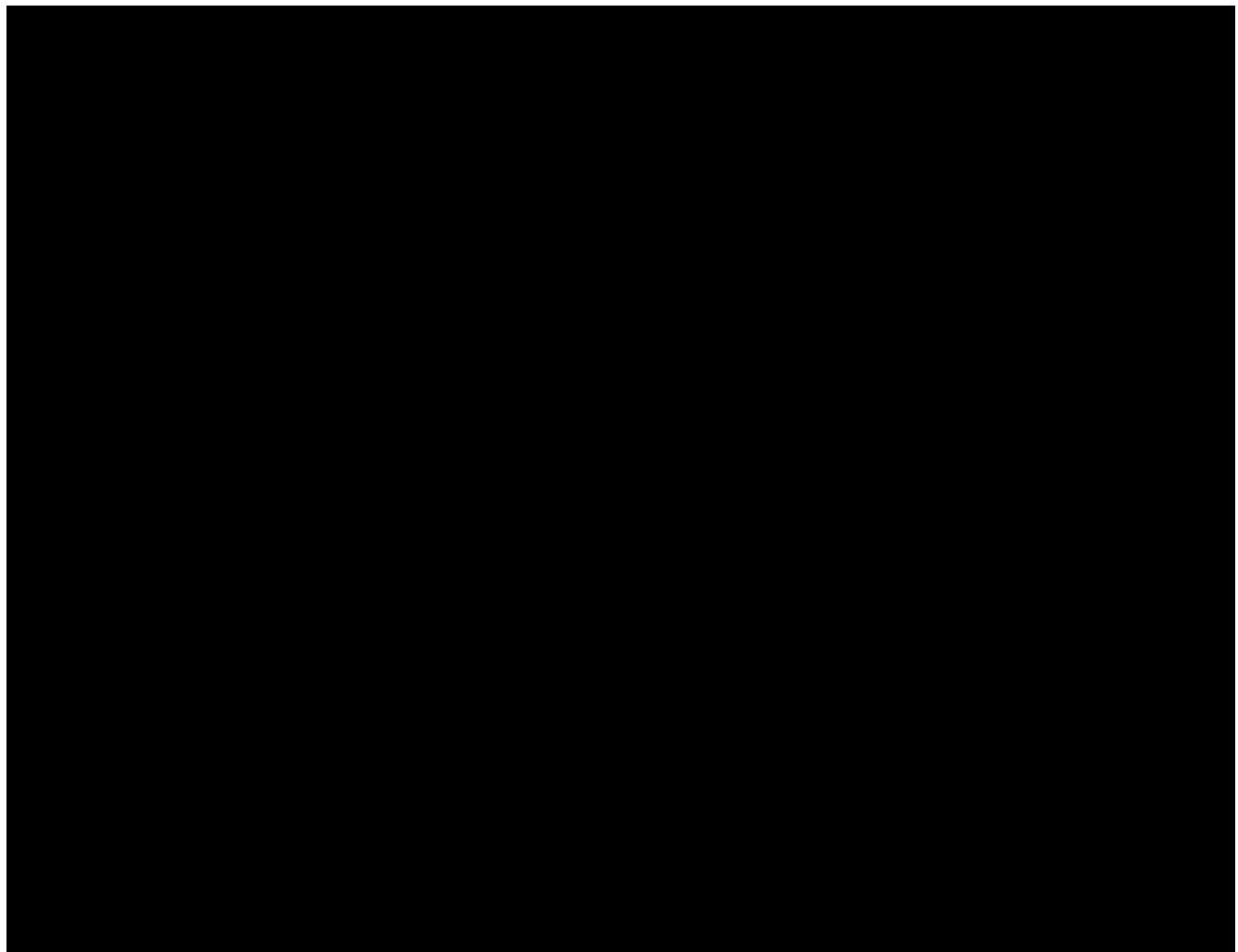


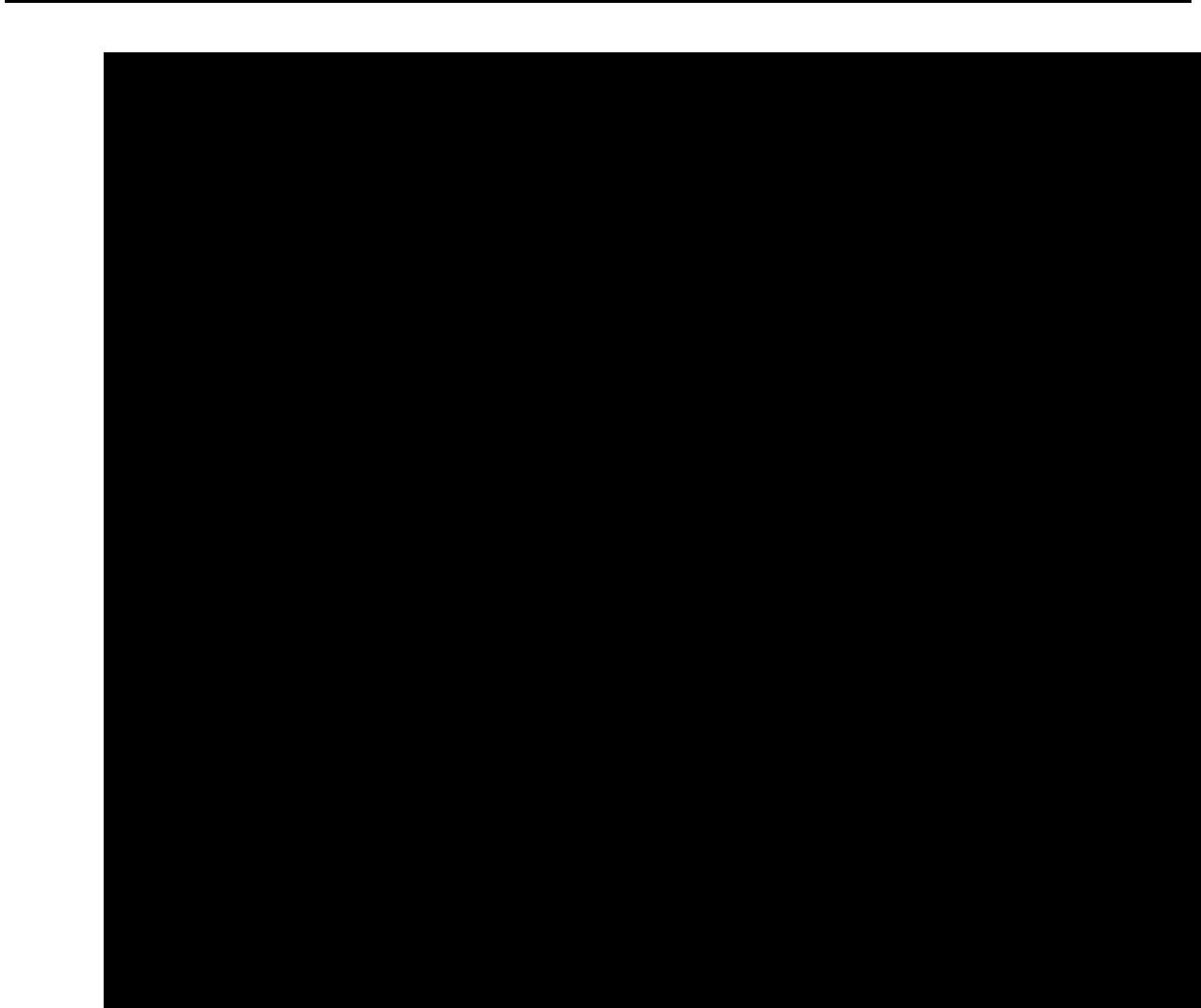


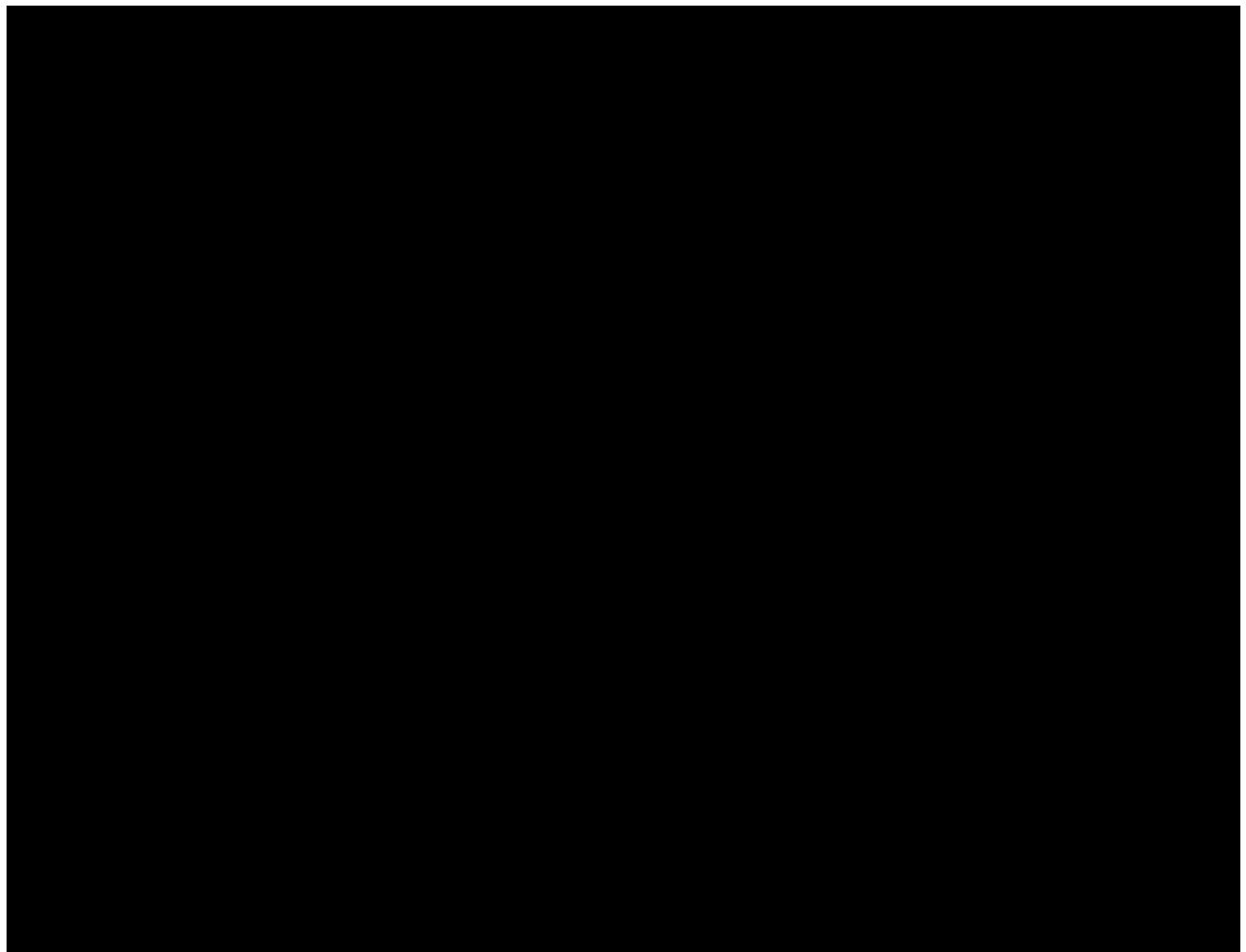
5. Select the export format and enter a name for the export file.



The CSV, TSV , Segment Export, and Segment Export With Header types need no extra configuration. However, the Profiles and Audience Export, Custom Record Service and Adobe Target Export need to be configured in Step 3. For example, see the configuration fields for the Profiles and Audience Export. Configure these export types and click **Next**.





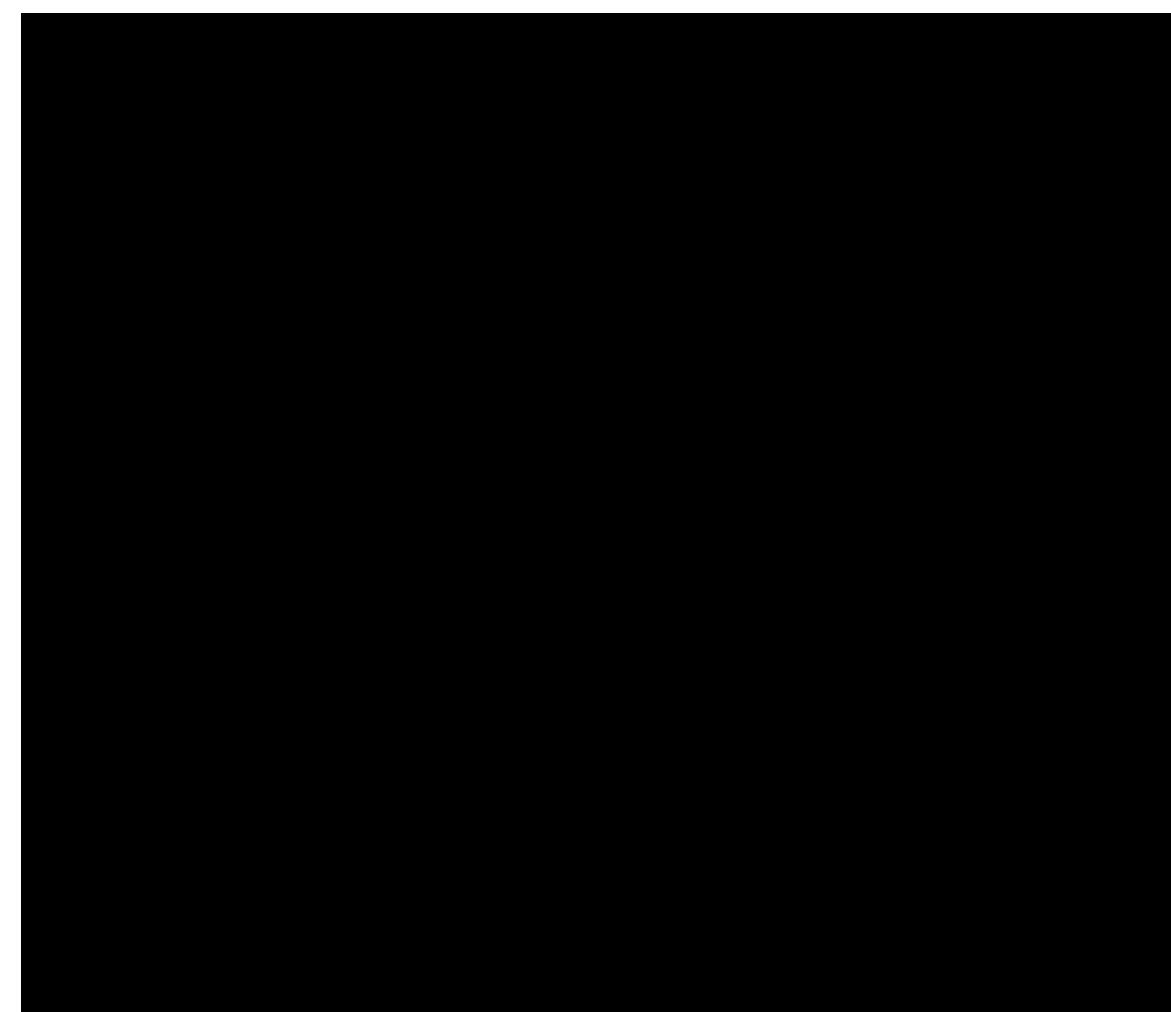


6. Configure the selected export type.

Header—If Header is True, then name the **Output File** field.

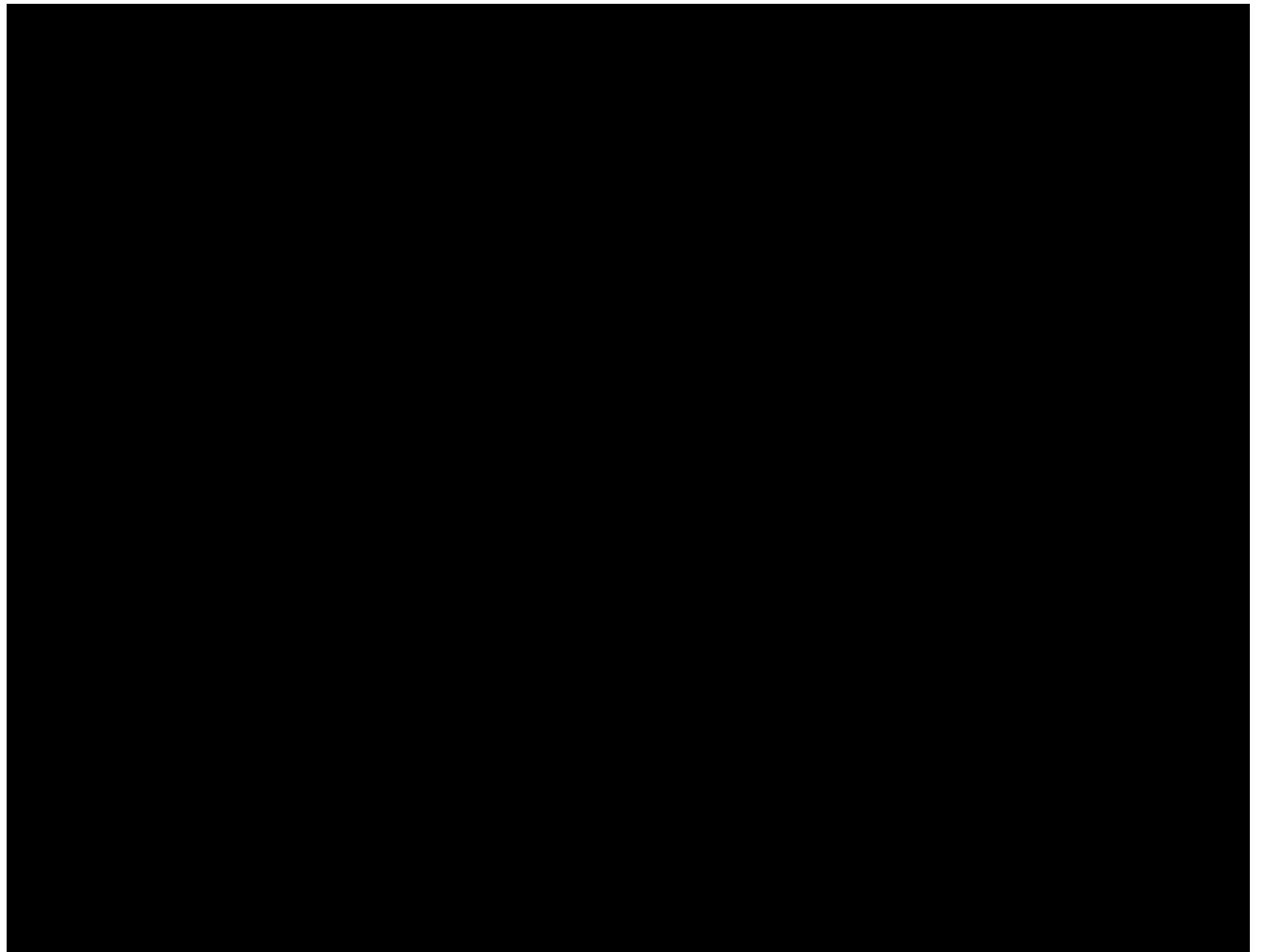
Escape Field—Set as **True** or **False**.

Order of Fields—Select a field and move up or down to set the order in the export file.

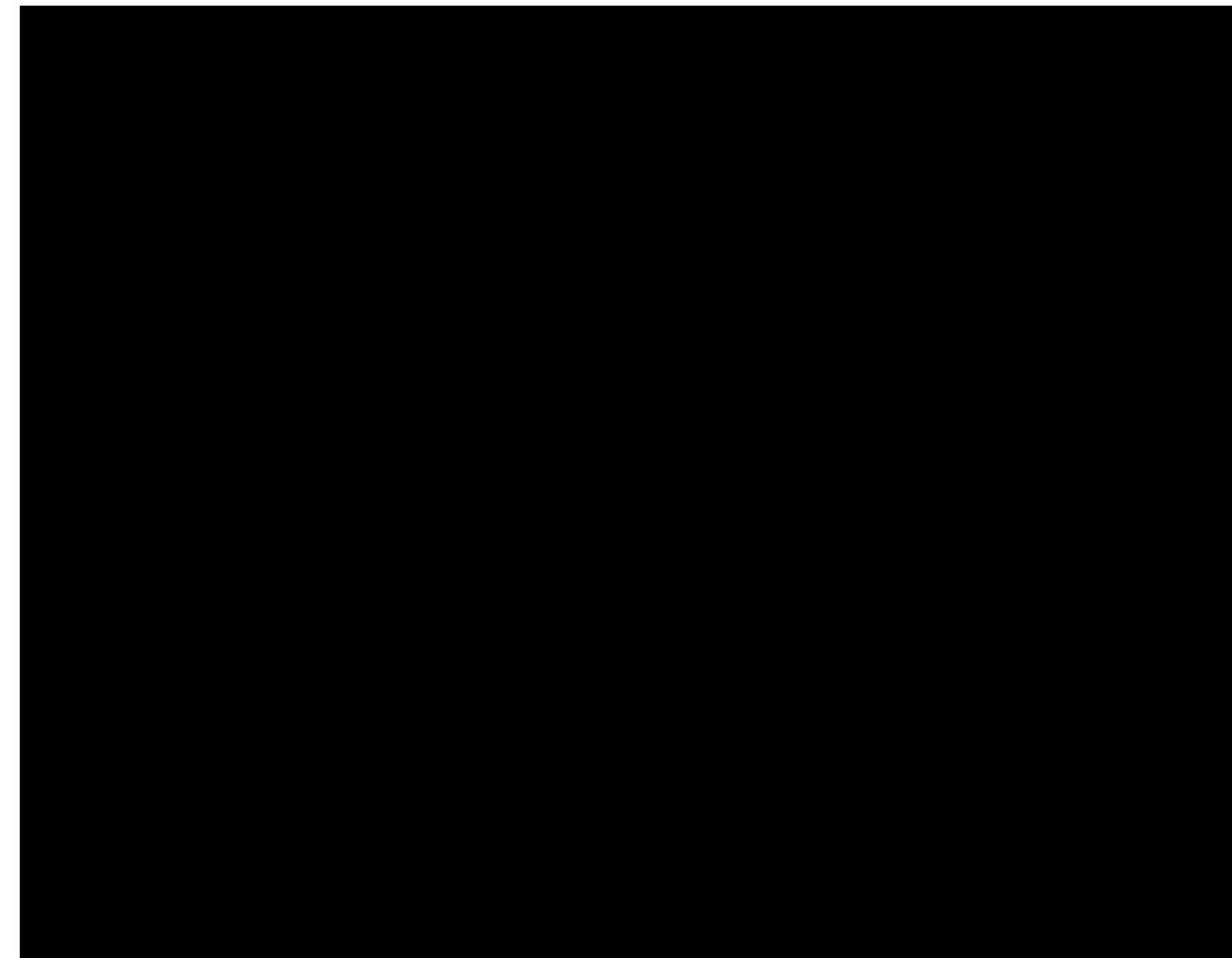


Click **Next**.

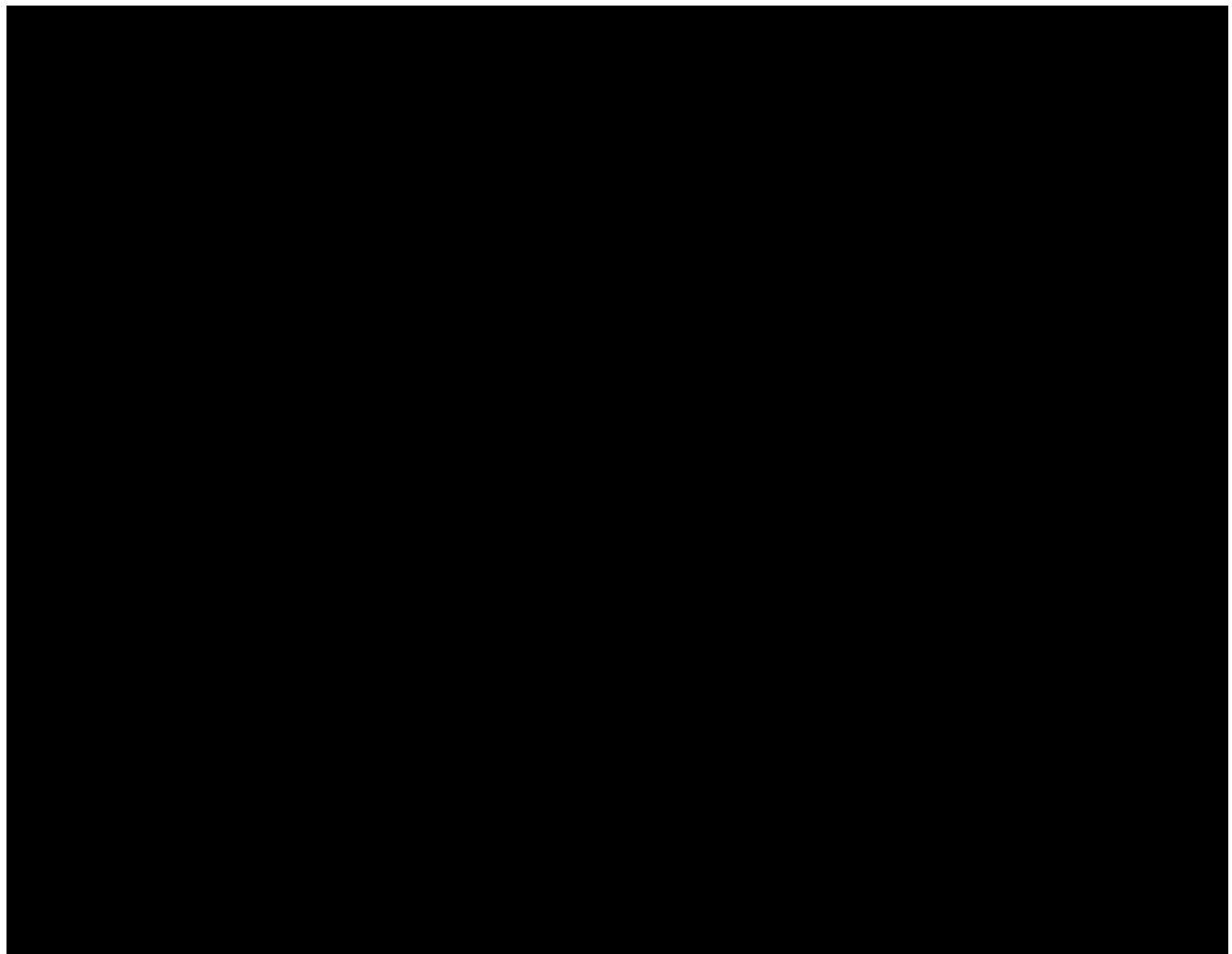
7. View the Level and applied filters in this dialog. Click **Next**.



8. If **CSV** , **TSV** , **Segment Export** or **Segment Export with Header** is selected then there are three options:
Generic Export - The output file will be generated by server in Server/Export folder.



FTP Export - The output file will be transferred to the server selected. (The server's list will be picked from `FTPServerInfo.cfg` file.)



SFTP Export - The output file will be securely transferred to the server selected.

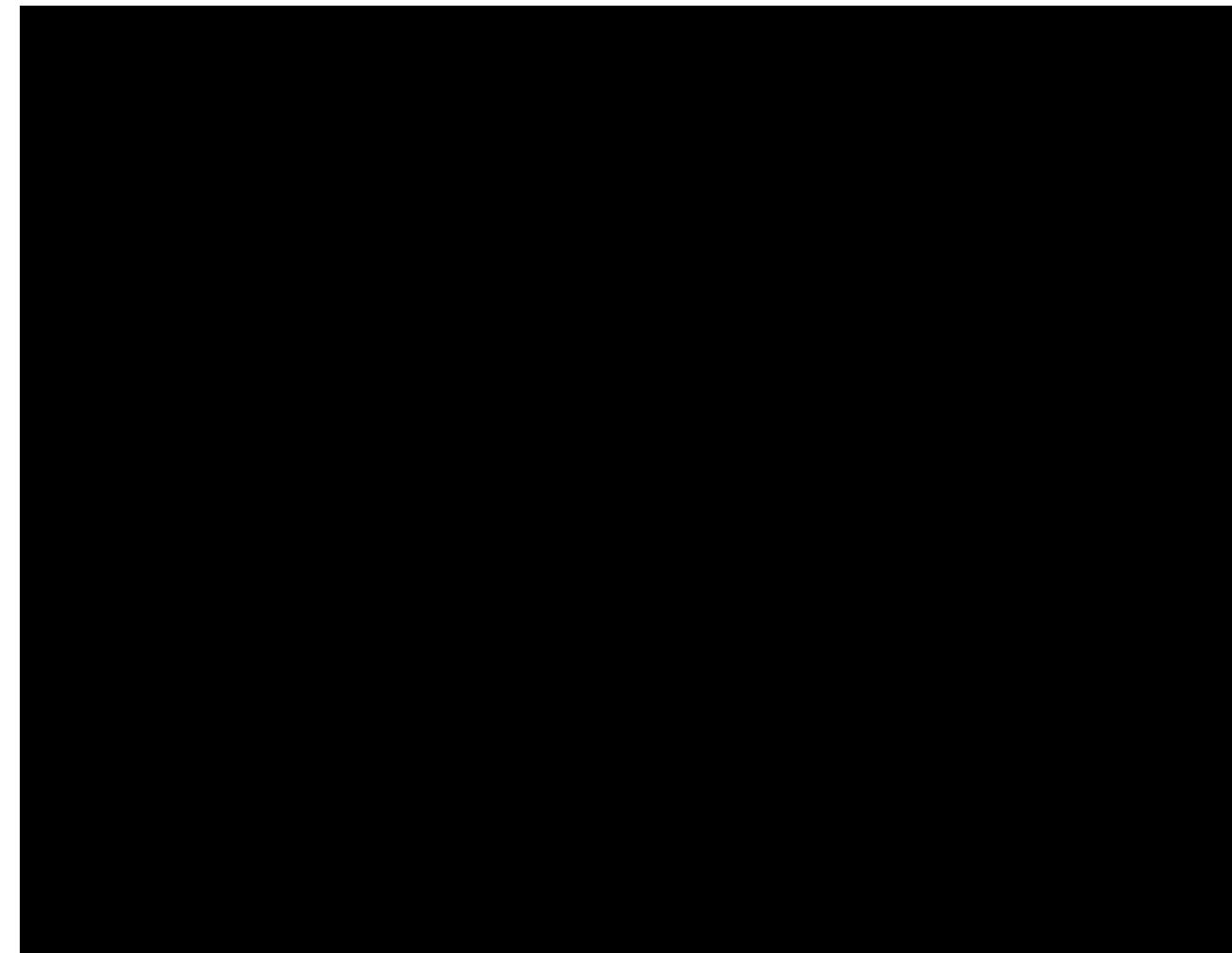
9. Click **Next**

Note: If the export type selected is **Profiles and Audience Export**, **Custom Record Service**, and **Adobe Target Export**, then the text will be static based on the export selected.

10. Configure scheduling parameters.

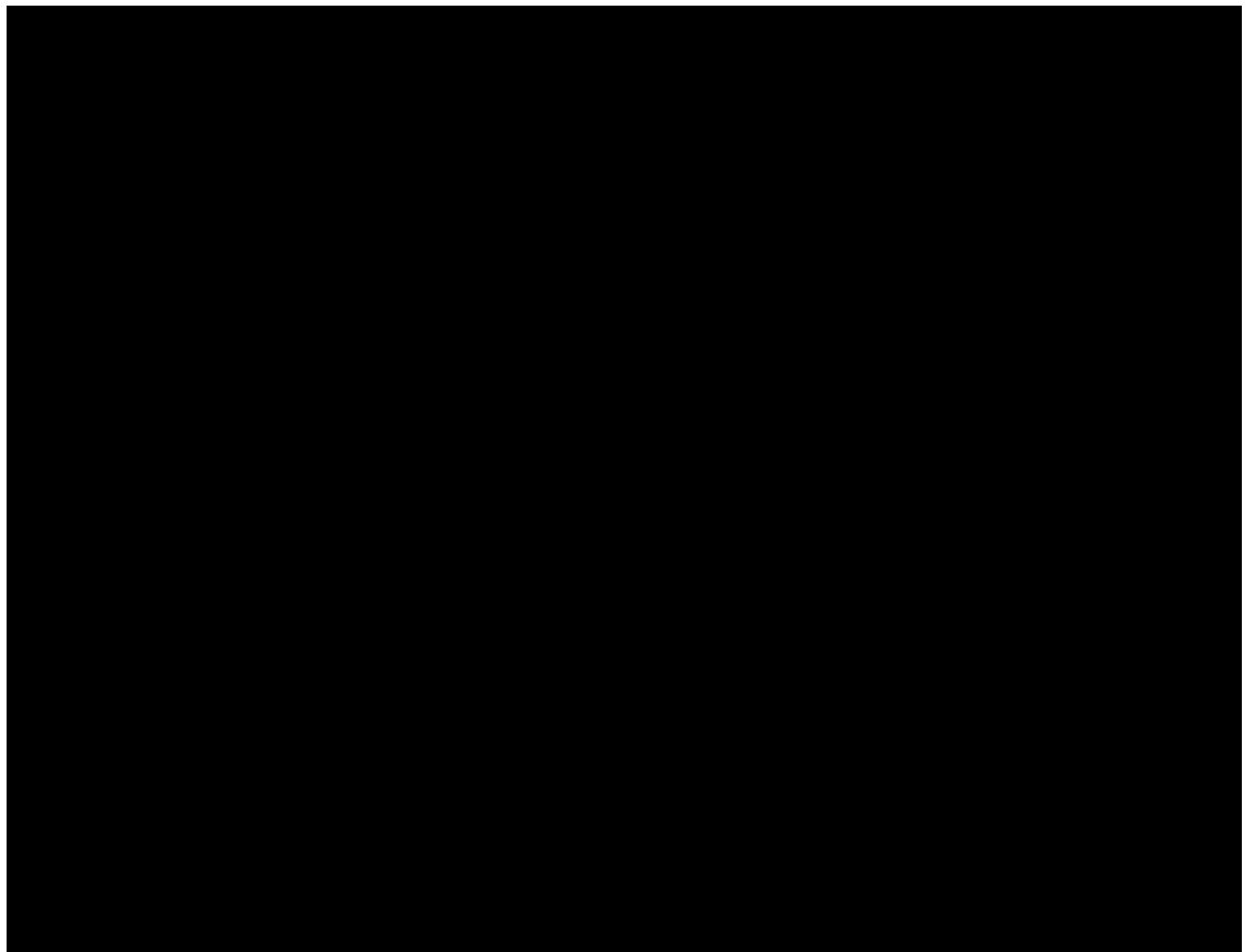
One Shot can be set to True or False.

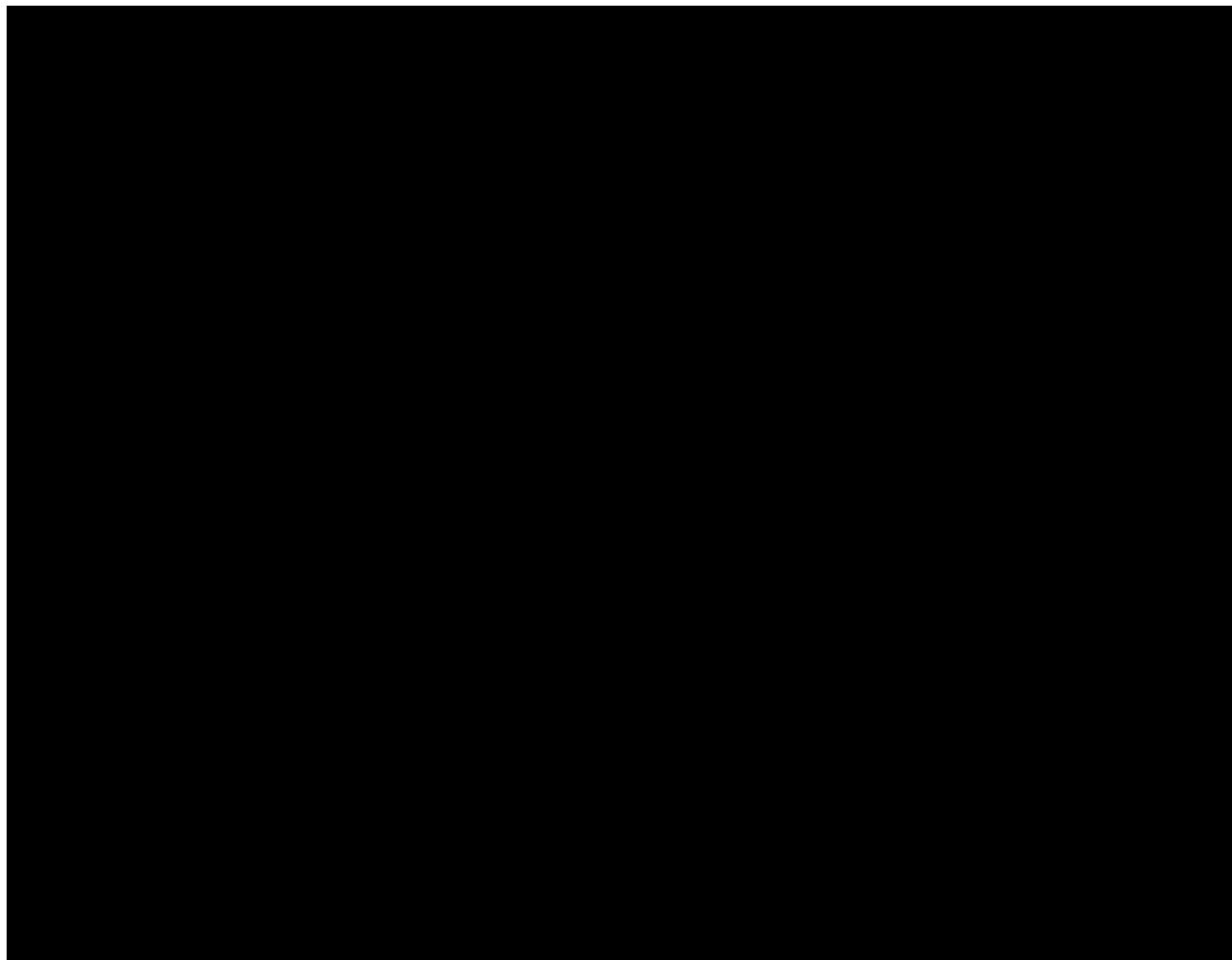
Advanced Scheduling can be turned On or Off by clicking the Advanced Scheduling Configuration button.



Like exporting from the Detail Table, One Shot will go away if Advanced Setting is On. Click **Next**.

11. Preview the export file and then click **Run Export**.





The following exports types are available using the wizard:

Segment Export types

- Generic
- FTP
- SFTP

Segment Export with Header

- Generic
- FTP
- SFTP

CSV Export

- Generic
- FTP
- SFTP

TSV Export

- Generic

- FTP
- SFTP

Analytics Reports

Analytic reports are provided for users of the Adobe SC profile. These data workbench reports—Page Views, Traffic, Unique Visitors, Referring Domain and other significant report types—are standard reports in Adobe Analytics.

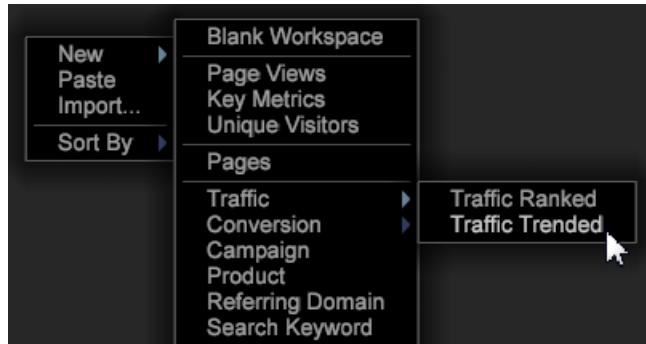
For Adobe Analytics' users, these templates allow users employing the Adobe SC profile (which uses the SC/Insight data feed) to view reports generated from Data Workbench using templates similar to the Adobe *Marketing Reports & Analytics* capability. Users can access these pre-configured reports using data workbench templates.



Note: This menu will only appear for those users who have implemented the Adobe SC profile.

To Open Reports

On the worktop, right-click and select **New** to view and open report templates.



The following are a list of Analytic reports:

Reports	Description
<i>Page Views</i>	A trended report that displays the number of times your website pages were viewed for the selected time period (hour, day, week, month, quarter, or year). This report allows you to track page views for each page on your site, as well as an aggregate of page views for your entire site.
<i>Key Metrics</i>	The Key Metrics Report lets you compare metrics to see whether they trend together. For example, as your page views increase, does your visitor count increase?
<i>Unique Visitors</i>	Shows you the number of unique visitors who accessed your site. Each visitor is counted once, regardless of how many times the person visits your website.
<i>Pages</i>	Ranks the pages on your site based on the pages that receive the most traffic. If your business question deals with quantitative data for pages, you can use this report to answer that question by adding the right metrics.
<i>Traffic</i>	Traffic reports give you in-depth insight into how visitors interact with your website.
<i>Conversion</i>	Provide comprehensive, accurate, and detailed analysis of customer activity. Metrics such as campaign management, sales cycle, customer fallout, and customer conversion let you measure e-commerce transactions, sources of sales, advertising effectiveness, customer loyalty, and more.

Reports	Description
Campaign	Displays information about the effectiveness of your advertising efforts. You can see which types advertising efforts give you the most traffic and which of your employees is responsible for driving those efforts.
Product	Identifies how individual products and groups of products (categories) contribute to your various conversion metrics, such as Revenue or Checkouts.
Referring Domain	Shows the domains that referred the customers that most impacted your site's success metrics. Referrers fall into two main categories: Domains and URLs. Domains refer to the domain name, and appear as the base domain without the query string or subdirectories attached. URLs include the base domain name, as well as any query strings or subdirectories.
Search Keyword	Reports that display a breakdown of search keywords.

Using attribution reports

Navigation title:

Attribution reports and visualizations in Data Workbench include Rules-based Attribution and Best Fit Attribution.

Attribution Profile

Using the new rules-based Attribution profile in Data Workbench, you can quickly analyze attribution events and assign responsibility leading up to a successful conversion defined by you. The Attribution profile comes complete with the information necessary for your data Architect to set up and extend its features, and includes pre-built workspaces for your Analyst to jump right in and start analyzing.

The Attribution profile allows you to gain a new perspective on the relationships between your marketing efforts and a successful customer lead generation or sales conversion. The Attribution profile helps you qualify interactions that should receive allocation of credit for realized revenue or participation downstream in the customer journey. It helps identify the impact of your marketing efforts and costs by allowing you to quickly analyze attribution events, and then assign responsibility for first or last touches or other events leading to a successful sale.

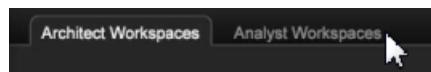


Important: The Attribution profile is configured for immediate use by users who have implemented the Adobe SC profile that uses the Analytics (SC/Insight) data feed. By default, the Marketing and Conversion events are employed as the default types of interactions evaluated in the provided rules-based models.

See [Deploying the Attribution Profile](#) on page 101 and [Attribution Models](#) on page 103 for additional information.

Architecture and Analyst Workspaces

Within the Attribution profile, you have Architect and Analyst workspaces defined on separate tabs in the workbench.



Architecture Workspaces

Within the **Attribution** tab, click the **Architect Workspace** tab to open workspaces specifically designed to set up your configuration files for basic attribution modeling.



The Architecture tab includes workspaces to step through each of the configuration files in the profile dataset folder. For example, **Attribution Configuration - Step 1** lets you identify the Attribution values within the Transformation section of the `profile.cfg` file.

Analyst Workspaces

Click the **Analyst Workspaces** tab to open workspaces prebuilt analysis utilizing the dimensions and metrics provided with the Attribution profile.

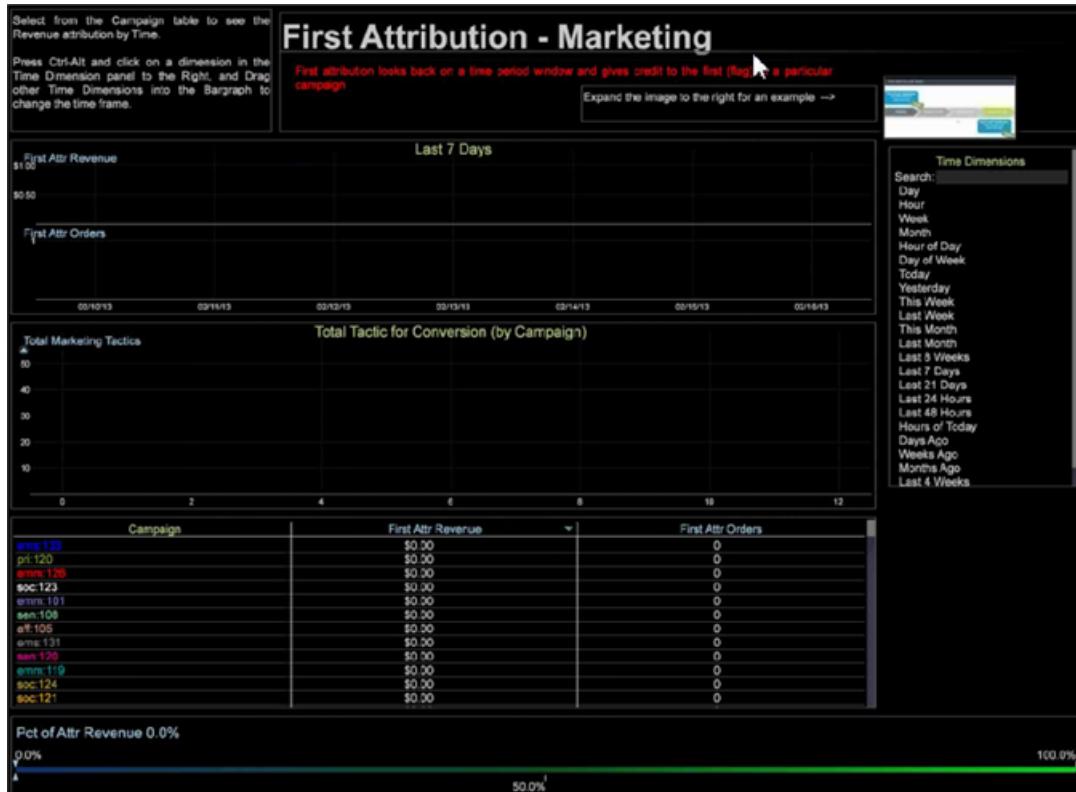
These workspaces are organized into four categories:

1. **Basic Reports** expose a single model within a workspace.
2. **Comparative Reports** extended the analyses by presenting multiple models within a single view.
3. **Investigation Reports** expands the reporting templates to present the attribution models in different formats. This section also introduces and exposes the position-based weighting ratios.
4. **Pathing Reports** provide visibility into the customer's marketing journey with multiple pathing visualizations to fully explore and express the process flows and interaction paths

The screenshot shows the Analyst Workspaces interface with several pre-configured reports:

- Campaign Attribution:** A chart showing revenue attribution across multiple channels.
- Basic Reports:** A section with five reports:
 - Last Attribution - Conversion:** Shows attribution for conversion.
 - Last Attribution - Marketing:** Shows attribution for marketing.
 - Even Attribution:** Shows attribution for even touchpoints.
 - Starter Player Cycler:** Shows attribution for a cycler tactic.
 - Latency Attribution:** Shows attribution based on latency.
- Adjacency Attribution:** A chart comparing adjacency attribution.
- First Last Even Attribution:** A chart comparing first, last, and even attribution.
- First Attribution - Conversion:** Shows attribution for conversion using the first model.
- First Attribution - Marketing:** Shows attribution for marketing using the first model.

The Analyst tab includes workspaces pre-configured with reports. For example, **First Attribution** lets you select from the **Campaign** table to see the **Revenue** attribution based on **Time**.



Deploying the Attribution Profile

The Attribution profile is an inherited, ready-to-drop-in profile. In combination with the Adobe SC profile and Analytics (SC/Insight) data feed, the profile can be deployed to quickly expose new attribution models across digital channels.

After saving the Attribution profile to the primary server, there are two additional steps necessary to integrate it into the current profile within the `Profile` directory: (1) Set up the `Profile.cfg` file, and (2) Declare the Required Fields.

Setting up `Profile.cfg` file

Like all profiles, the Attribution profile needs to be added to the `profile.cfg` file. Because the Attribution profile depends on the Adobe SC profile, the Adobe SC profile needs to be listed first in the configuration file before the Attribution profile.



Note: These steps will require a re-transformation of the dataset.

1. Open the `profile.cfg` file in your custom profile folder. (Open in `server\Profiles\(\custom profile name)\profile.cfg`.)
2. If the Attribution profile is not listed in the configuration file, add it to the list.



3. Make sure the **Attribution** string is listed below the **Adobe SC** profile string.
4. Save the updated `profile.cfg` file and then save it to server from the Profile Manager.

Declaring the Required Fields

The Attribution profile takes predefined fields and with a series of transformations exposes those fields in new and useful ways through extended dimensions. To provide the most immediate value the Attribution profile depends on fields available with the Adobe SC profile.

Default Variables	Field Name and Decoder Position (Adobe SC)
Campaign	x-campaign, #199
Marketing Channels	x-va_closer_detail, #162 x-va_instance_event, #163
Order event	x-order, #206 x-purchaseid, #200
Revenue	x-revenue, #205

Default Variables	Field Name and Decoder Position (Adobe SC)
Units	x-units, #204

1. Verify that these fields are declared in the Decoder Group used to define the Adobe Analytics data source. The default decoder group is provided under Dataset\Log Processing\Decoding Instructions.cfg.
2. Verify that these fields are declared in the **Fields** section of the SC_Fields.cfg file. This file can be located under Dataset\Log Processing\SC_Fields.cfg.

Attribution Additions and Troubleshooting

The Attribution profile added a configuration file, 0a_Marketing_Channels.cfg, which copies the value of the x-va_closer_detail into a new field called x-marketing-channel, when the x-va_instance_event field matches "1". Both x-va_closer_detail and x-va_instance_event are decoded by default, and passed from decoding in the installed packages available when you update to version 6.2.

The x-marketing-channel field is then used in the Simple dimension called Marketing Channel.



Important: If you have altered your profiles by removing previously unused fields that are now being used, you will want to verify that the x-va_closer_detail and x-va_instance_event fields are being decoded and passed through for use.

If fields are missing, then you will get a message in your detailed status:

x-va_closer_detail is not available

or

x-va_instance_event is not available

Attribution Models

Seven different attribution models are provided to use with the Attribution profile to help illustrate and quantify the customer-marketing journey.

First and Last model



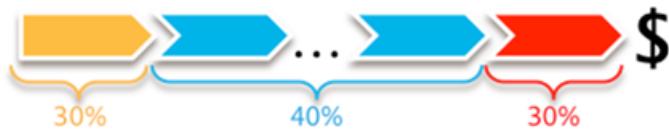
The first attribution models are the traditional **First** and **Last** touches. Understanding the first step into converting a sale or the last touch before checking out and buying a product provides an understanding of campaign types—for awareness, the First touch, and for call-to-action, the Last touch.

Even model



The view into the marketing engagement is expanded with the introduction of **Even**. Every qualifying marketing interaction is given an equal share of the subsequent order and revenue generated from the conversion.

Starter, Player, and Closer model



And moving beyond an equal allocation the **Starter, Player, & Closer** model provides a position based weighting scheme with adjustable weighting. The weights can be explored while performing analysis.

Consider this example: The starter (or first) gets 30 percent of the converted revenue, closer (or last) gets 30 percent, and all the players (those in-between) share an equal portion of the remaining 40 percent.

 **Note:** It's fairly common to assign the allocation back to marketing interactions based on either revenue or occurrence of the order. These next two models attribute different characteristics to the marketing interactions to expose other dynamics of success.

Adjacency model



The Adjacency model provides visibility into the marketing position away from the conversion, answering the question: Is the marketing channel typically the 1st (closest), 5th, or 10th marketing channel interaction away from success?

Latency model

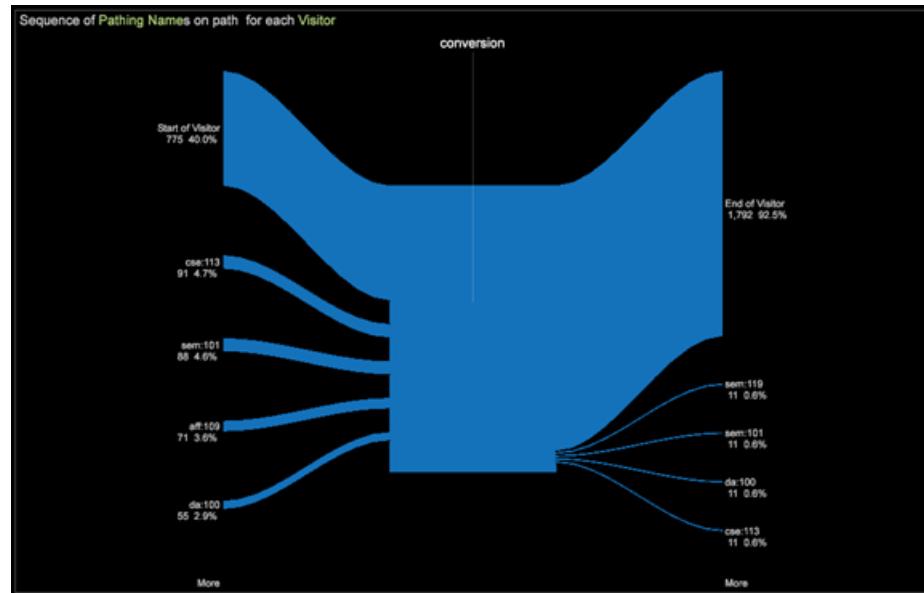


The **Latency** model helps describe the lag time between the interaction with the marketing channel and the occurrence of success. This is useful in presenting awareness versus call-to-action type campaigns and knowing which levers the business can pull to get a more timely response from a target audience. (A value of zero (0) days means the marketing interaction occurred the same day as the success.)

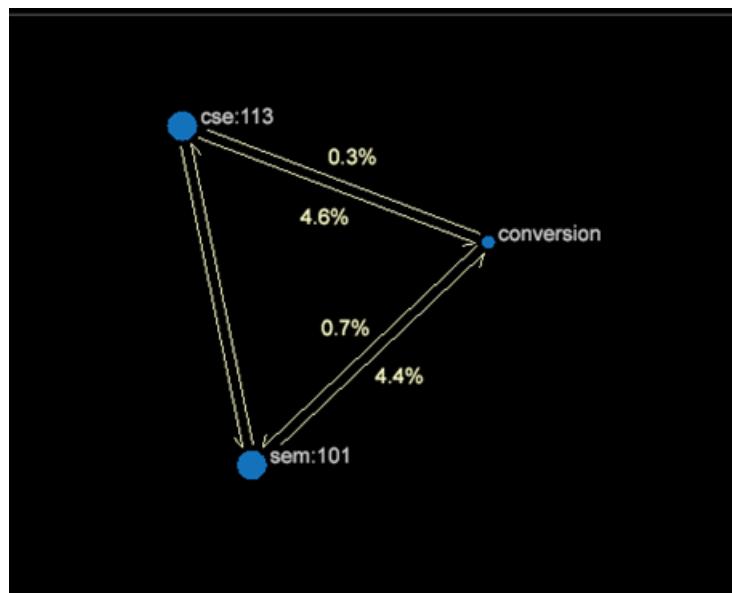
Pathing model

The Pathing model provides different approaches for exploring the customer's engagement with marketing and his or her successful conversion, establishing the relationships between marketing interactions within the customer journey. Explore the process maps to understand high-level flows between supporting marketing channels and success. Evaluate direct sequential marketing interactions with the path browser.

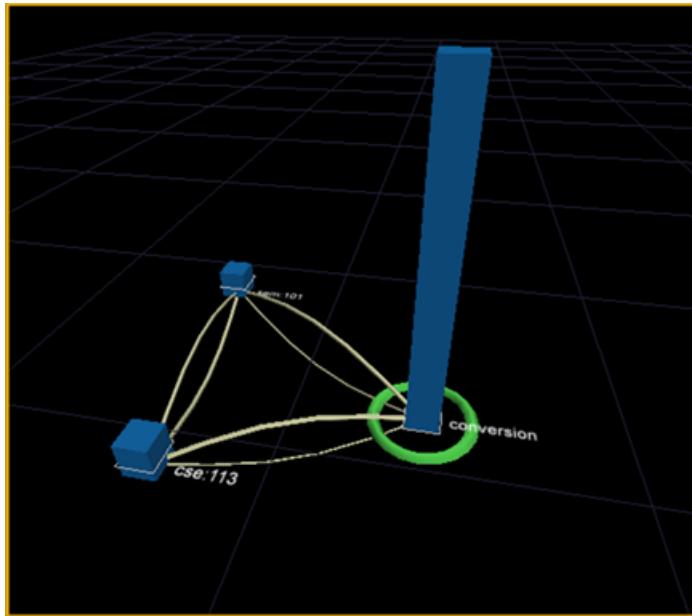
Path Browser visualization



2D Process Map visualization



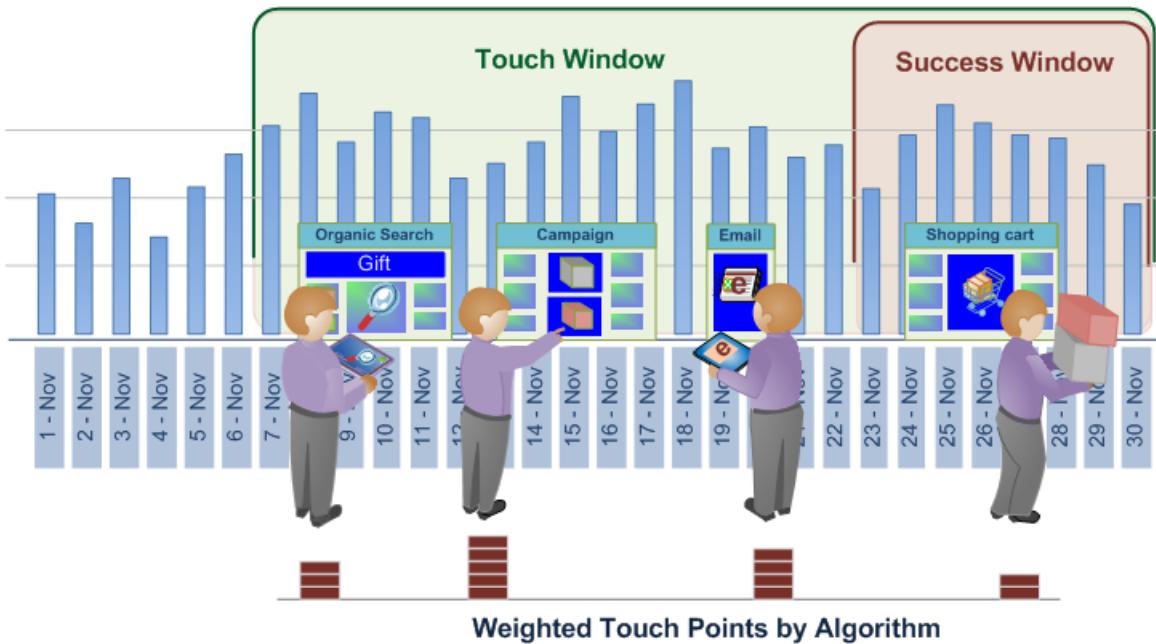
3D Process Map visualization



Best Fit Attribution

Best Fit Attribution is a machine-learning approach to assigning attribution values across the different channels of a successful conversion event. Data Workbench automatically evaluates contributions to success across a window of time per channel, and then builds an attribution model based on your customers' actual interaction patterns.

Best Fit Attribution lets you compare the interactions, or touches, that contributed to a successful sale, email sign-up, or other performance indicators. The attribution analysis automatically assigns weight to the most important touches and provides an attribution model per channel based on your data and responsive to your market and internal protocols.



For example, if a customer visits your site through an organic search, then engages with a campaign, and then signs up for an email, **rules-based Attribution** would identify the first touch or last touch, or evenly distribute success attribution across all touch points using preset attribution models. Where rules-based attribution is defined by the user, the Best Fit attributes sets values through an algorithm by calculating the probability of a conversion as a function of the observed touch points.



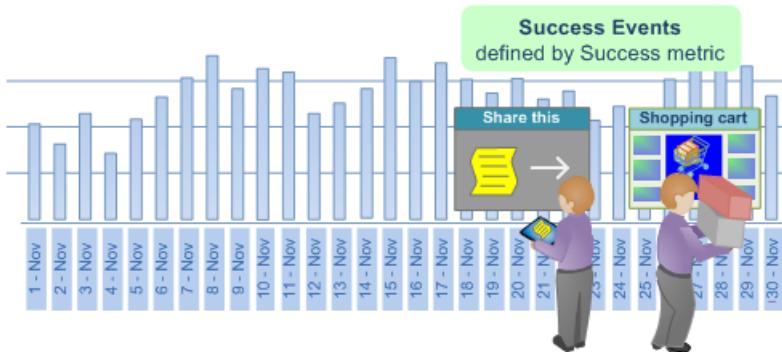
Note: To run **Best Fit Attribution** in Data Workbench, you need to update your server certificate (.pem file) to support Adobe Analytics Premium. You also need to add **Premium** to your custom **Profile.cfg** for the client and receive new certificates from Adobe ClientCare for Server and Report Server. See [Upgrade Requirements and Recommendations](#) on page 44 for Data Workbench 6.3.

Basic Setup

See [Build a Best Fit Attribution Model](#) on page 55 for step-by-step instructions.

Set the Success metric

Define a metric representing a success event.



The Success Metric is often *Orders*, although you can leverage Data Workbench to define a very complicated success metric in conjunction with the Success Window.

Set the Touch metric (optional)

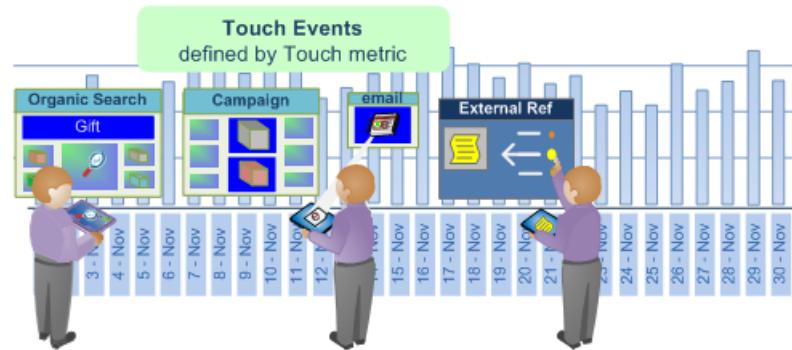
Identify interactions to track that led to a successful conversion, then set the Touch metric over which attribution will be calculated.



Note: Setting a Touch metric is only required if you are using it to derive Channel Metrics from drag and drop Dimension elements instead of using existing Channel metrics.

If you do not have a metric defined for campaigns or channels, but do have dimensions representing channels, the Best Fit Attribution can build them for you automatically based on the Touch metric.

For example, with the Touch Metric set as *Hits*, and given a dimension called *Media Type* with elements that include *Email*, *Press Release*, *Print Ad*, and *Social Media*, the visualization will generate Channel metrics of the form *Hits* where *Media Type* = *Email* when you drag and drop the element(s) onto the visualization.

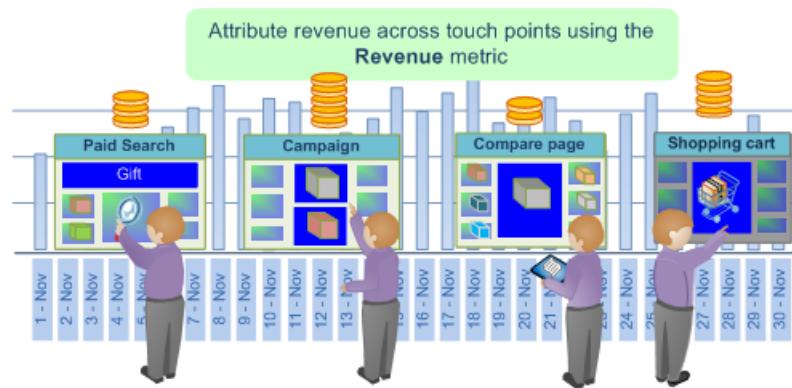


The Touch metric then determines the allocation of attribution scores to identify marketing interactions considered influential for success, allowing you to qualify marketing touches for the population identified in the Success window. You can set metric such as *Page Views* or *Hits*, or use customized touch metrics specific to your needs.

In many cases, the Touch window should include the Success window to evaluate a long lead time in the sales cycle.

Set the Revenue metric.

You can opt to identify revenue across touch points by setting an appropriate revenue metric. If specified, the model will display the distribution of revenue over the input channels.



You can set a revenue metric with currency data types to allocate success across all top touch points defined and analyzed. This metric breaks down the final sales revenue and allocates based on the weighting allocated by the algorithm.

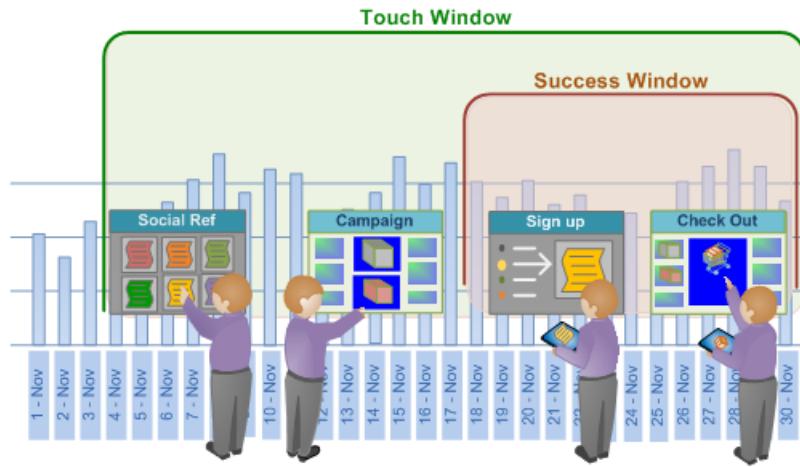
Set the Success and Touch Windows.

The Success window defines the population to examine and the period for successful events, allowing you to indicate the windows of time and breadth of population to consider for the analysis through a workspace selection. The **Success** window defines the period and population to examine for success events. The **Touch** window specifies the historical time period to examine for channel interactions leading up to the success events.



Note: Setting a Touch Metric is only required if you are trying to build Success metrics automatically by dragging dimension elements onto the visualization.

You can set a day, month, year, or any available time frame to constrain your evaluation of success and touch events across the sales cycle or for specific audiences entering your site. Creating windows to limit attribution allows you to focus your analysis on the relevant periods of time for your specific needs.



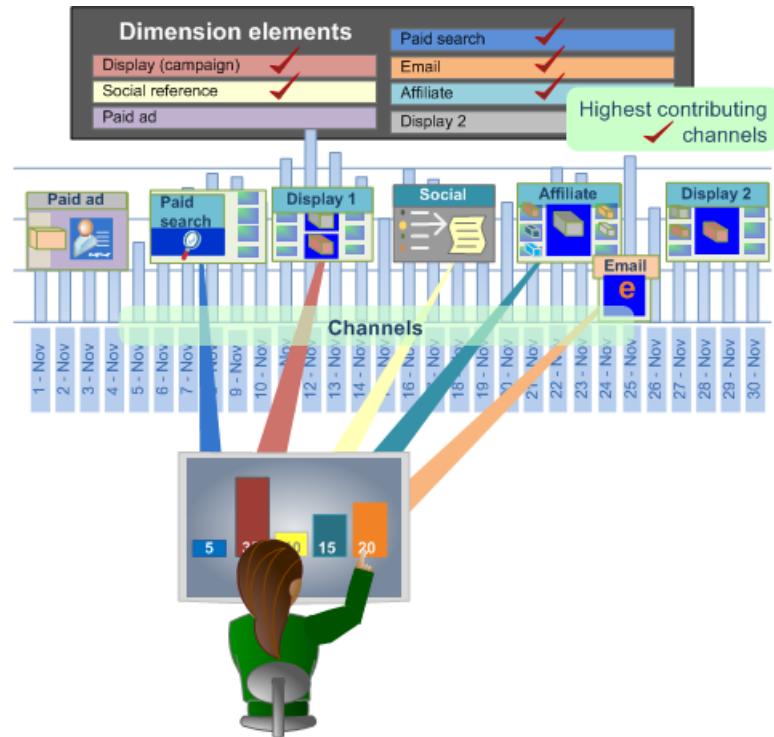
In many cases, you will want the Touch window to include the Success window to let you extend your analysis over a long lead time based on your sales window. Or you can track and analyze touches separate from the success event.

Select the Channels.

When entering channels you have two choices.

Add the Touch Metric and add Dimension Elements to the Channels

In many cases, you will want to break down the top touch points by dimension elements to define specific channels. Based on the element values, Best Fit Attribution will automatically select the top performers and rank them according to percentage and display them in a chart visualization.



An attribution model will be built by drawing on the visitors who interacted during your Success window and examining the channel Touches during the Touch window that did or did not result in a successful event.

Breaking Down by Channels

When entering channels you have two options:

- Add a **Touch Metric** and then add **Dimension Elements** for the Channels.
or
- Create metrics that filter for the channel elements that you want to evaluate.

Option 1: Add a Touch Metric and add Dimension Elements for Channels.

This is the easier approach. Best Fit Attribution creates the metrics automatically to evaluate for attribution. In the example below the Touch Metric is *Hits* and Channels are: *Display Campaigns*, *Email Campaigns*, and *SEM Campaigns*.

Using this method, Best Fit Attribution creates a metric in the background for evaluating the attribution across the channels (but you never see the auto-generated metric and they are not saved). In the example below, three metrics are created where Hits is filtered for each of the three channels (e.g, *Display Campaigns*, *Email Campaigns*, and *SEM Campaigns*). This is the easiest because you let the Best Fit Attribution create the metrics for you.

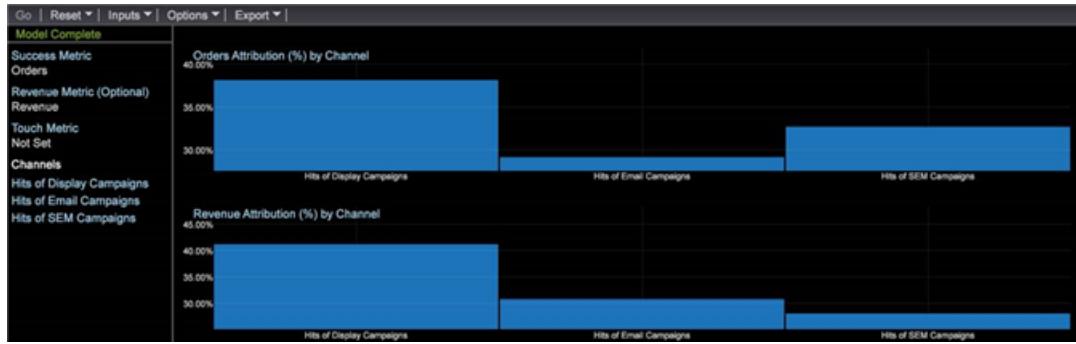


Option 2: Create a Metric.

In the second option, you create and save the metrics for the channels that you want to evaluate by filtering a specific channel. An example of a such a metric is shown below.

Editing	Hits of SEM Campaigns.metric
Name	Hits of SEM Campaigns
Formula	sum(One, Hit)[MarketingChannel = "SEM Campaigns"]
Preview	6,614,522

Then, instead of entering a Touch Metric and Dimension Elements for the Channels you can click on the menu bar in the visualization and select **Inputs** > **Add Channel** and then select the metrics that you created.



See the example of the second method below. You can see that the results of both options are identical.

Build a Best Fit Attribution Model

Open Best Fit Attribution from the **Premium** menu and follow these steps to build a Best Fit Attribution model.

See an overview of [Best Fit Attribution](#) on page 50.

1. Open **Best Fit Attribution**.

Open a workspace and click **Premium > Best Fit Attribution**.



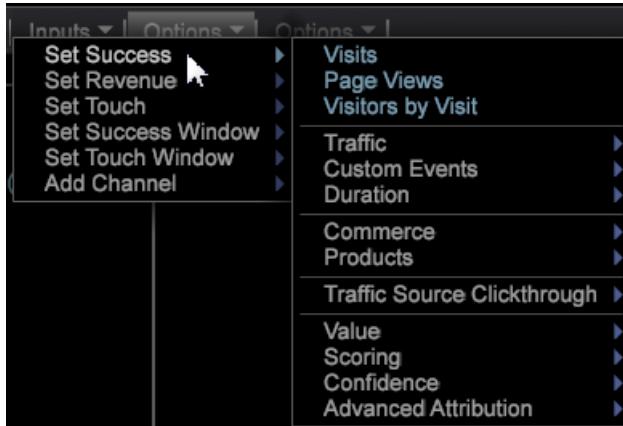
Note: Best Fit Attribution is an Adobe Analytics Premium feature that requires you to enable Premium in your Profile. It requires you to update your certificate and add the Premium profile to your profile.cfg file. See [upgrade instructions](#) for DWB 6.3.

2. Set the **Success** metric.



Note: You can either drag a metric from a **Finder** table to the left pane of the Attribution visualization, or select from the **Inputs** menu.

Click **Inputs > Set Success**. The metric menu will open.



Select a metric that identifies a successful conversion.

3. (optional) Set the **Revenue** metric.

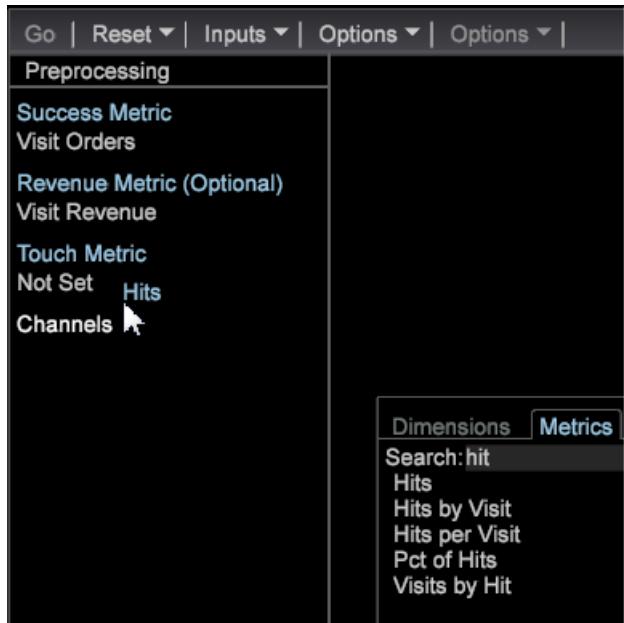
Set a metric to evaluate revenue across the conversion process.

4. Set the **Touch** metric.



Note: Setting a Touch Metric is only required if you are trying to build Success metrics automatically by dragging dimension elements onto the visualization.

Click the **Inputs** menu and select **Set Touch**, or drag a metric from the Finder.

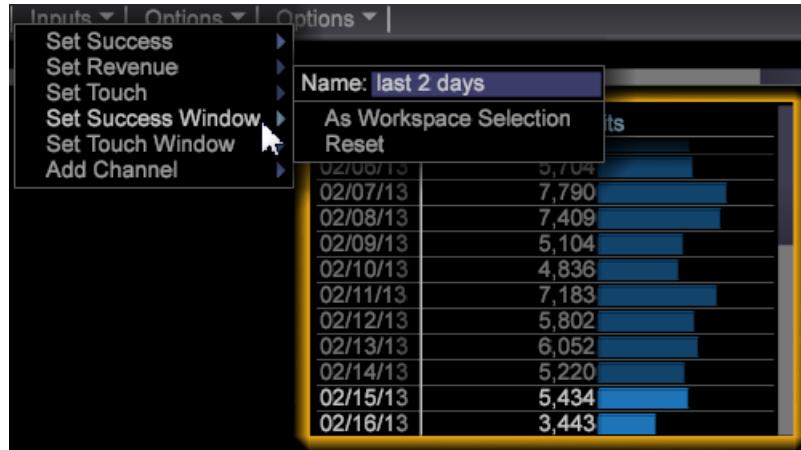


This will be used to derive channel metrics when dimension elements are used as inputs.

5. Set a **Success** window.

Click **Inputs > Success Window**. Select a date range from a table and then name the Success window.

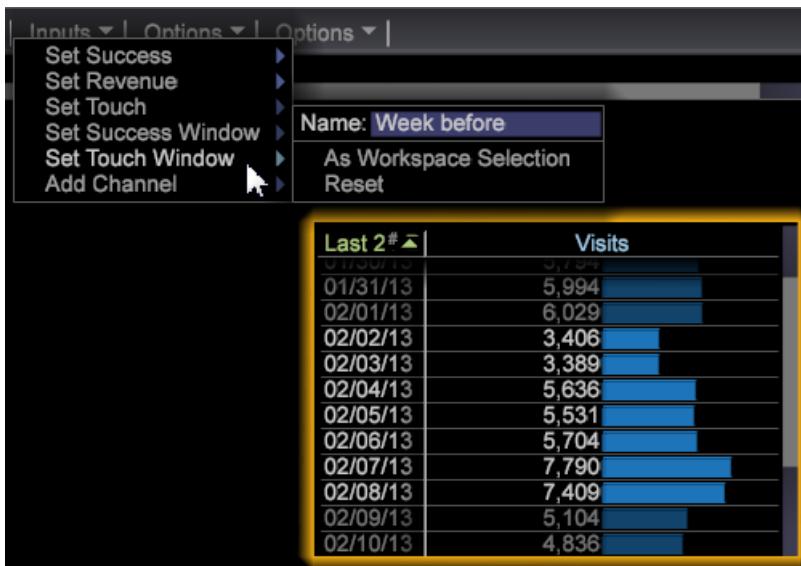
Click **Workspace Selection** and the selected dates will be assigned as the range of time for the Success metric.



Note: Since the Success window is a workstation selection, you can include any dimension(s) to your Success window.

6. Set a **Touch Window**.

Click **Inputs > Touch Window**. Select a date range from a table and then name the Touch window. Click **Workspace Selection** and the selected dates will be assigned as the range of time for the Success metric.



By default, the **Touch** window will be set to the same time period as the **Success** window.

7. (optional) Set a Training Filter.

You can also specify a **Training Filter** in the workspace to filter visitor data.



Note: In setting both the Success and Touch windows, you can apply the Training filter to the current workspace selections to further limit your data.



Note: The training set is always drawn from visitors who satisfy the Success window. By filtering using the Filter Editor, you can create a subset of visitors reported in the Success window.

8. Specify channel metrics that represent touches.

Either drag metrics to the visualization, or choose them from the **Inputs > Add Channel** menu. If you do not already have metrics defined for campaigns or channels, but do have dimensions representing channels, the visualization can build them for you automatically with the specification of a Touch metric.

For example, with the Touch metric set to **Hits**, and given a dimension called **Media Type** with elements that include things like **Email**, **Press Release**, **Print Ad**, and **Social Media**, the visualization will generate Channel metrics of the form **Hits** where **Media Type = Email** when you drag and drop the element(s) onto the visualization.

9. Press **Go**.

The Best Fit Analysis process will run, and a chart will display attributions per channel based on the selected inputs.



Note: Right-click **Model Complete** on the completed analysis to see statistics for the attribution model.



When complete, a graph will display an attribution model calculated per channel, and a distribution of the *Revenue* metric (if set). The model can be saved internally or exported to other systems.



Note: **Streaming**, **Online** and **Offline** modes produce different effects when building an attribution model based on the latency of the data being evaluated. In Streaming mode, the detail **Model Complete** message will display. In Online and Offline modes, the detail **Local Model Complete** will display.

Options menu

The **Options** menu provides advanced features to set up and display Best Fit Attribution analysis.

Options menu	Description
Set Training Filter	<p>The Training Filter is used with the Success Window to filter the population when building the attribution model. This will provide a subset of data that includes only the visitors that you want to analyze.</p> <p> Note: Experienced users can also leverage the flexibility of filters to focus beyond the time line of the Success and Touch Windows. For example, in addition to selecting a time range, you can select a set of <i>Referring Domains</i> to only examine the attribution for users from those domains.</p>
Show Complex Filter Description	Displays the filter code for the Training Filter, Success Window, and Touch Window.
Save Model	Saves the current attribution model for future use.
Load Model	Opens a previously saved attribution model.
Presentation View	Hides the top menu bar for presentation.

Options > Advanced includes features to set the training set size and specify the approach to take in the case of a class imbalance.

Advanced > Training Set Size	<p>Sets the training set size.</p> <p> Note: The default training size is Large for 250,000 visitors.</p> <ul style="list-style-type: none">• Tiny = 50,000• Small = 75,000• Normal = 100,000• Large = 250,000• Huge = 500,000
Advanced > Class Balance	Identifies and defines the number of input records to generate for a class imbalance issue based on dataset size.

Reset and Remove options	Description
Reset Model	From the Reset menu, select Reset Model to clear the visualization but keep input metrics.
Reset All	From the Reset menu, select Reset All to clear the visualization and the input metrics.
Remove	Right-click on any input and select Remove to clear the metric from the selected input.
Remove All	Right-click on <i>Channels</i> and select Remove All to clear all input metrics.

Visualizations

Navigation title:

Overview of the concepts and features that you need to understand when working with visualizations.

Opening visualizations

Navigation title:

Information about opening visualizations.

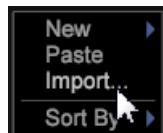
Because your implementation of Data Workbench can be fully customized, it may differ from what is documented in this guide. Exact paths to each visualization are not provided in this guide.

All visualizations can be opened by right-clicking within a workspace and selecting the desired menu option.

After opening up a new workspace, you may need to click **Add > Temporarily Unlock**.



Note: Visualizations cannot be imported like workspaces. When you right-click in the worktop and select **Import**, you can import an existing workspace, but not a visualization residing outside of the workspace.



If you attempt to import a visualization, you will receive a "not a valid workspace" error.

Quick reference

Navigation title:

View a list of Data Workbench client user interface actions and shortcut keys.

See also [Keyboard shortcuts for the Data Workbench client](#) on page 280.

Actions for the Data Workbench client

Use these actions when performing tasks within the client user interface.

Selecting Data

Action	Result
Click	Selects item
Click + drag	Selects range
Ctrl + select	Adds to selection
Shift + select	Remove from selection

Closing workspaces

Action	Result
Click the Data Workbench logo in the upper left corner of the workspace	Saves your changes
Ctrl + Delete	Closes without saving your changes

Moving graphs

Action	Result
Click + drag	Moves graphics
Right-click + drag	Rotates graphics
Click both mouse buttons + drag	Zooms into graphics

Adding a page to a map

Action	Result
Click + drag	Add to page hierarchy
Ctrl + Alt + click + drag	Add page to a bar graph
Ctrl + Alt + click + drag	Add to page table

Keyboard actions for text

Action	Result
Click	Selects an item
Click + drag	Selects a range of text
Ctrl + select	Adds text to selection
Shift + select	Removes text from selection

Keyboard shortcuts for the Data Workbench client

Use these shortcut keys for workspaces and configuration files.

Keyboard shortcuts from the workbench thumbnail view

Action	Result
Ctrl + N	Opens a new workspace. If you are limited to creating a single type of blank workspace, then the new workspace will open immediately. If you have rights to open multiple types of workspaces based on your implementation, then Ctrl + N displays a menu allowing you to choose the type of workspace to be created.
Ctrl + O	Imports a workspace

Resize and refit workspaces between window and full page views

Commands	Quick Keys	Combined menu commands
Full screen view. Workspace fills the screen and refits to the new size.	Ctrl plus Ctrl + (on keypad) <i>or</i> Ctrl Shift + (on keyboard)	• File > Page Size > Fill Screen <i>followed by</i> • File > Refit Workspace
Window view. Workspace displays in a standard window view and refits to the new size.	Ctrl minus Ctrl -	• File > Page Size > Standard <i>followed by</i>

Commands	Quick Keys	Combined menu commands
		• File > Refit Workspace

General keyboard shortcuts for text

Action	Result
Ctrl + X	Cuts text
Ctrl + Z	Undo action for previously typed characters
Ctrl + Shift + Z	Redo action for typed characters
Ctrl + A	Selects all

Keyboard shortcuts within a workspace

Action	Result
Ctrl + P	Prints a workspace.
Ctrl + W	Closes a workspace without saving.
Ctrl + S	Saves a workspace.
Ctrl + O	Opens a visualization file. It is the same as selecting Add > Open from the workspace menu.
Ctrl + E	Open a configuration (by default) or other files. It is the same as selecting Add > Edit from the workspace menu. Select All to open other files.
Commonly-used files menu	Click Add > Open from a workspace to open recently opened files (saves the last ten) or files that you "pin" to the new menu. From the menu, you can click the pin icon in the list to save the visualization.

Keyboard shortcuts within an open configuration (.cfg) file

Action	Result
Right-click a node and click Paste	Pastes a node into a .cfg file
Right-click the vector and click Paste as child	Pastes a vector into a .cfg file at a vector level

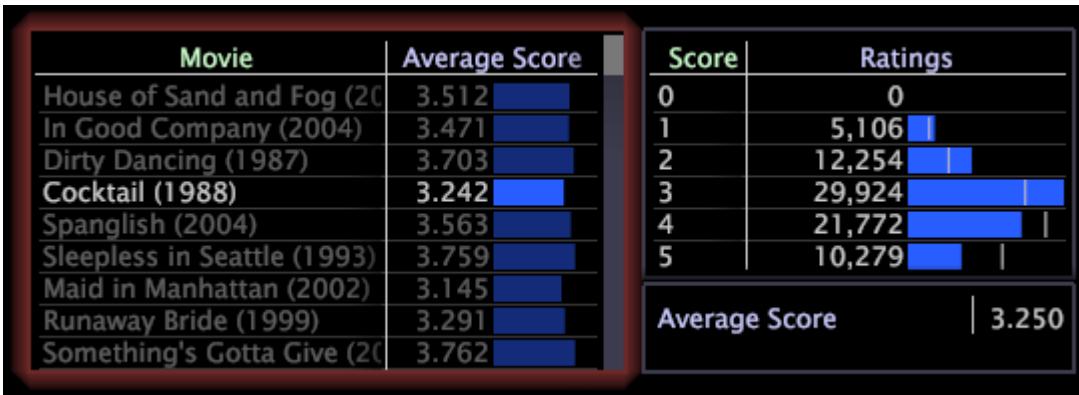
Make selections in visualizations

Navigation title:

Selecting elements within a visualization dynamically filters the dataset.

When you make a selection in a visualization, all of the other visualizations in the workspace automatically update to reflect the data associated with only those elements that you have selected.

The following workspace shows a Movie table in which the movie *Cocktail* is selected. In the workspace, the Score table and the metric legend automatically filter their displays for the selected element (that is, their displays reflect the data for the movie *Cocktail*).



As shown in the example above, when a selection is made, a glow appears around the visualization and the non-selected elements within that visualization dim. To facilitate comparisons with the complete dataset, narrow white lines referred to as benchmarks appear in the bar graph to mark the shape of the original, unfiltered data. For more information about benchmarks, see [Benchmarks](#) on page 285.

To make a selection

You can make selections in any visualization that displays at least one dimension except scatter plots and legends.

Use the following mouse and key sequences to select the desired element(s):

To...	Use this sequence...
Select a single element	Click
Select a range of elements	Click+drag
Add an element to current selection	Ctrl+click
Clear a single selection	Shift+click
Clear all selections (that is, re-select all elements)	Right-click any value in the visualization

To clear a selection

Use the following mouse and key sequences to clear a selection:

To	Use this sequence...
Clear a single selection	Shift+click
Clear all selections (that is, re-select all elements)	Right-click any value in the visualization

Understanding How a Selection Affects Other Visualizations

Navigation title: Understanding How a Selection Affects Other Visualizations

Within a workspace, a visualization represents a set of query results.

When you make a selection, Data Workbench filters the results of the queries that it uses to produce the visualizations in the workspace. The specific filter varies by visualization.

The following examples illustrate how Data Workbench applies a selection to three different types of visualizations. Reviewing these examples helps you understand the filtering effect that selections have on visualizations. They also help you understand how to interpret the results that you see in a filtered visualization.

- [Filtering a Visualization with a Sessions Metric](#) on page 283
- [Filtering a Visualization with a Visitors Metric](#) on page 283
- [Filtering a Visualization with a Visitors-by-Session Metric](#) on page 283

Filtering a Visualization with a Sessions Metric

In this example, the /direct.asp/?ldPage=hme URI in the visualization on the left is filtering the metric for Sessions displayed in the visualization on the right.

URI	Sessions
/direct.asp?ldPage=hme	4,661
/v3/nav/ebootnavbranch220.htm	4,296
/direct.asp?ldPage=hme_tst1	3,854
/loading.asp?ldPage=hme_s1_r	3,354
/pops/disclosure_pop.asp	2,367
/direct.asp?ldPage=pro_evr_t1	1,949

URI	Sessions
/direct.asp?ldPage=pro_evr_t1	578
/direct.asp?ldPage=pro_evr_t8	73
/direct.asp?ldPage=pro_wc	593
/direct.asp?ldPage=pro_wc_aa	219
/pops/disclosure_pop.asp	1,113
/direct.asp?ldPage=pro_wc_aa	126

- **Effect of Selection on the Query:** Data Workbench filters the Sessions for the selected URI. In this example, the query that generates the value for the /pops/disclosure_pop.asp element is filtered as follows:

```
Sessions[ URI="/pops/disclosure_pop.asp" AND URI="/direct.asp/?ldPage=hme" ] by Page View by Session
```

- **Interpreting the Visualization:** The filtered visualization represents the number of Sessions that include the URIs listed in the visualization and /direct.asp/?ldPage=hme. This example shows that there were 1,113 sessions during which visitors viewed both /pops/disclosure_pop.asp page and /direct.asp/?ldPage=hme in the same session.

Filtering a Visualization with a Visitors Metric

In this example, the /direct.asp/?ldPage=home URI in the visualization on the left is filtering the metric for Visitors in the visualization on the right.

URI	Sessions
/direct.asp?ldPage=hme	4,661
/v3/nav/ebootnavbranch220.htm	4,296
/direct.asp?ldPage=hme_tst1	3,854
/loading.asp?ldPage=hme_s1_r	3,354
/pops/disclosure_pop.asp	2,367
/direct.asp?ldPage=pro_evr_t1	1,949

URI	Visitors
/direct.asp?ldPage=hme	4,870
/loading.asp?ldPage=hme_s1_r	2,207
/v3/nav/ebootnavbranch220.htm	2,046
/pops/disclosure_pop.asp	2,041
/direct.asp?ldPage=hme_tst1	1,633
/direct.asp?ldPage=pro_wc	1,050

- **Effect of Selection on the Query:** Data Workbench filters the Visitors for the selected URI. In this example, the query that generates the value for the /pops/disclosure_pop.asp URI is filtered as follows:

```
Visitors[ URI="/pops/disclosure_pop.asp" by Page View by Visitor  
AND URI="/direct.asp/?ldPage=hme" by Page View by Visitor ]
```

- **Interpreting the Visualization:** The filtered visualization depicts the Visitors who have viewed the URIs listed in the visualization and /direct.asp/?ldPage=hme (although not necessarily during the same session). The example above shows that 2,041 visitors have viewed both /pops/disclosure_pop.asp and /direct.asp/?ldPage=hme.

Filtering a Visualization with a Visitors-by-Session Metric

In this example, the /direct.asp/?ldPage=hme URI in the visualization on the left is filtering the metric for visitor-by-session in the visualization on the right.

URI	Sessions	URI	Visitors by Session
/direct.asp?idPage=hme	4,661	/direct.asp?idPage=hme	3,961
/v3/nav/eftopnavbranch220.htm	4,296	/loading.asp?idPage=hme_s1_r	1,482
/direct.asp?idPage=hme_tst1	3,854	/v3/nav/eftopnavbranch220.htm	1,269
/loading.asp?idPage=hme_s1_r	3,354	/pops/disclosure_pop.asp	1,069
/pops/disclosure_pop.asp	2,367	/direct.asp?idPage=hme_tst1	292
/direct.asp?idPage=pro_evr_t1	1,949	/direct.asp?idPage=pro_wc	569

- **Effect of Selection on the Query:** Data Workbench filters the Visitors by Session for the selected URI. For example, the query that generates the value for the /pops/disclosure_pop.asp URI is filtered as follows:

```
Visitors[ ( URI="/pops/disclosure_pop.asp" by Page View
    AND URI="/direct.asp/?idPage=hme" by Page View ) by Session ]
```

- **Interpreting the Visualization:** The filtered visualization depicts the Visitors who have viewed both URIs listed in the visualization and /direct.asp/?idPage=hme during the same session. This example shows that 1,069 visitors saw both /pops/disclosure_pop.asp and /direct.asp/?idPage=hme during a single session.

Selecting Data Using Other Methods

Navigation title:Selecting Data Using Other Methods

In addition to making simple yet powerful point-and-click selections in visualizations, your Adobe application provides three methods that you can use to select data and work with that selected data.

Each of these methods is appropriate for particular situations:

- **Segments:** Segments enable you to build dimensions on-the-fly from your selections. Segments are most useful when working with disjointed, non-overlapping groups that you want to use as a dimension.

Example: You want to compare two types of shoppers based on their previous purchases: Window Shoppers versus Paying Customers.

Using a segment affects only the workspace in which you are using the segment dimension. See [Segments](#) on page 412.

- **Filters:** The filter editor enables you to make complex data selections and see the selection criteria clearly specified in the editor. You also can save any filter to use again.

Example: You want to analyze the data related to customers from California who placed more than five orders in one month.

Using a filter affects only the workspace in which you have applied the filter. See [Filter editors](#) on page 348.

- **Subsets:** Subsets enable you to identify and work with smaller subsets of your data. Subsets give you a smaller dataset to analyze, which provides you with faster query results. When using a subset, all of your benchmarks relate to your subset, not to the entire dataset, which is much more useful when analyzing a specific subset.

Example: You want to perform an in-depth analysis of the data related to only those customers from California.

Using a subset affects all of your workspaces because the subset is applied globally to Data Workbench. See [Subsets](#) on page 289.

Zoom to View Visualizations

Use the Zoom feature to see metric labels in a visualization with a high disparity of values.

Use the zoom feature to better view metric labels when values reach a higher disparity. For example, when you set a higher metric regression value against previous values. Previously the label would disappear with the change in the contrast of values. You can now zoom in to the visualization by clicking <Ctrl> and moving the mouse wheel while hovering over the graph.

Benchmarks

Navigation title:

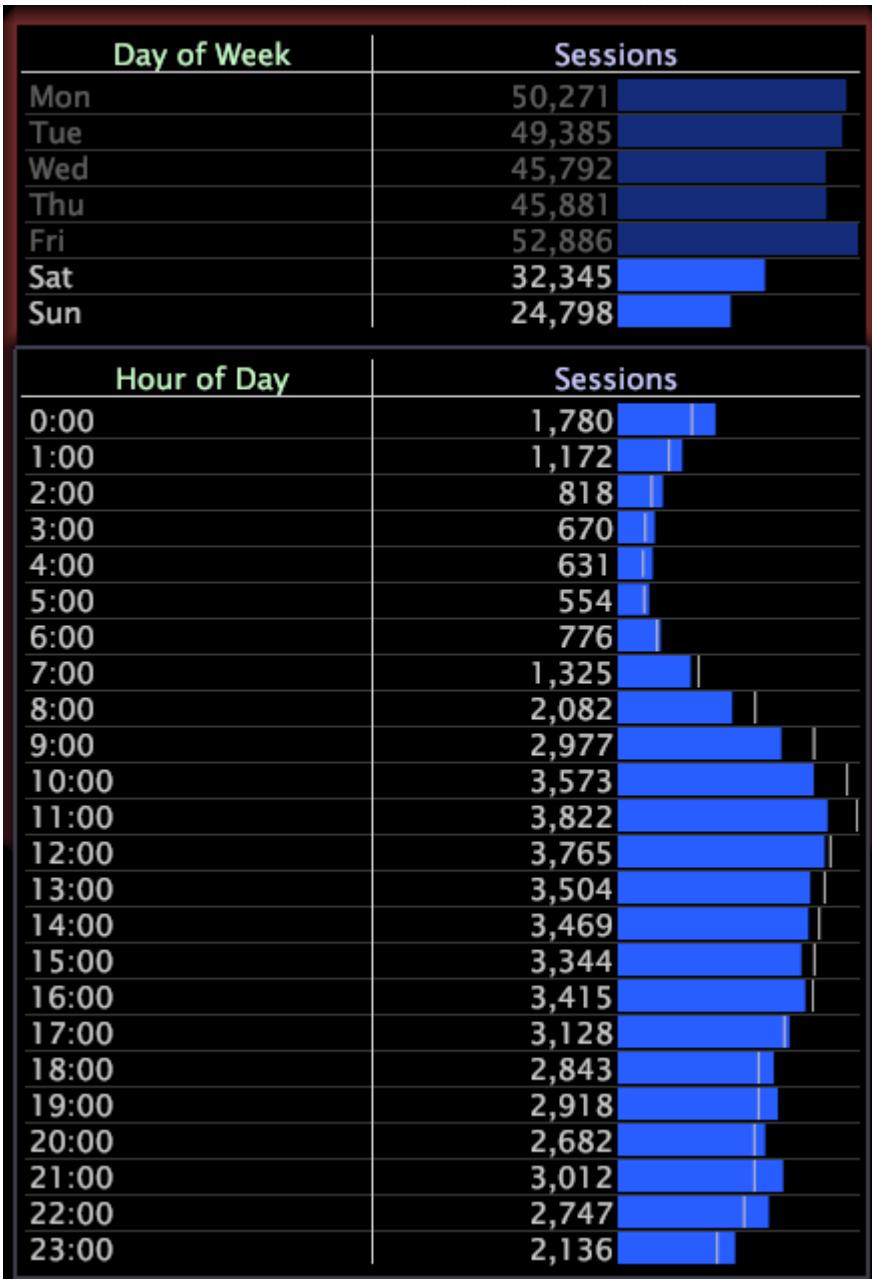
Benchmarks display qualitative differences in data by comparing the distribution of a metric based on all of the data with the distribution of a metric based on a selection in the data.

They provide a quick reference of normal behavior by illustrating what a visualization would look like if no selections were made. The differences in the data are based on relative values, not absolute values.



Note: Benchmarks can represent differences in data for the dataset or for any data subsets that you create. For more information about subsets, see [Subsets](#) on page 289.

The following graphs show the sessions on a website by day of the week and hour of the day. The visualization on the left provides a good representation of the pattern that the site traffic takes over time. To compare the overall site traffic pattern to the traffic pattern on weekends, you would select Saturday and Sunday in the Sessions by Day of Week graph as shown in the visualization on the right.



After a selection is made, benchmarks (the thin white vertical lines highlighted above) appear in the Sessions by Hour of Day graph on the right. From the benchmarks, we can see that the traffic pattern on weekends when compared to the overall traffic pattern generally includes fewer sessions earlier in the day (08:00 to 11:00).

The Radar Visualization is also useful for analyzing benchmarks. See [Radar visualization](#) on page 411.

Metric and dimension menus

Navigation title:

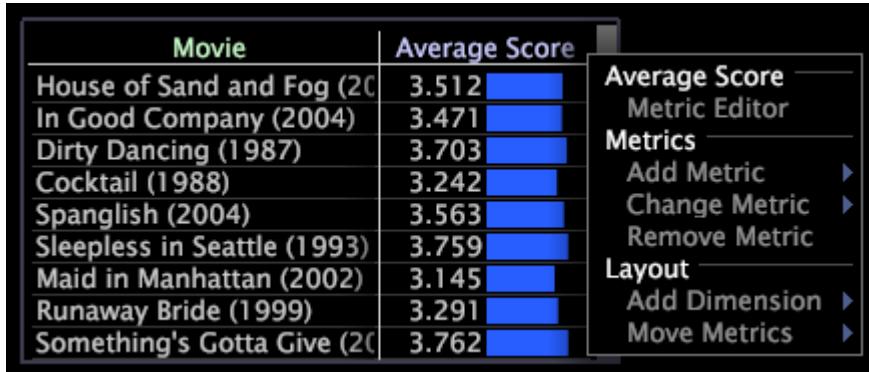
The menus accessed by right-clicking the metric or dimension area within a visualization provide access to actions related to the metric or dimension in that visualization.

For more information about using these menu options with a particular type of visualization, see the appropriate visualization section within [Analysis visualizations](#) on page 304. For information about the default menu options for all other menus in Data Workbench, see [Data Workbench Client](#) on page 198.

The Metric menu

You access the metric menu by right-clicking any metric-related area within a visualization.

This menu provides you with different options depending on the type of visualization in which you are working. The following example shows some of the available menu options, while the following table lists all of the available menu options.



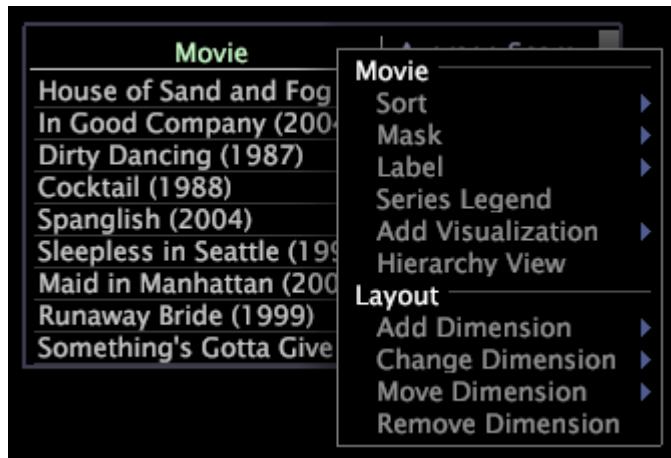
Metric menu

Menu Option	Description
<i>metric name</i>	Heading for the following metric-related menu options.
Zoom to All	Returns to a display of all of the data within the graph. See Zoom in or out in a visualization on page 292.
Zoom to Selection	Zooms in on a selected subset of data in a line graph. See Zoom in or out in a visualization on page 292. For more information about selections, see Make selections in visualizations on page 281.
Metric Editor	Enables you to edit the definition of the metric identified by the metric name field. See Edit a derived metrics on page 462.
Metrics	Heading for the following metric-related menu options.
Add Metric	Adds a new metric to the visualization to the left of the selected metric.
Change Metric	Changes the selected metric in the visualization.
Remove Metric	Removes the selected metric from the visualization.
Layout	Heading for the following layout-related menu options.
Add Dimension	Adds another dimension to the visualization.
Move Metrics	Moves the selected metric to another axis in the visualization.

Use the dimension menu

You access the dimension menu by right-clicking any dimension-related area within a visualization.

This menu provides you with different options depending on the type of visualization in which you are working. The following example shows some of the available menu options, while the following table lists all of the available menu options.



Dimension menu

Menu Option	Description
Add Callout	Adds a callout for the selected element, which draws attention to that element. See Adding callouts to a workspace on page 293.
Rename Element	Renames the element. Click Revert to return to the original name.
<i>dimension name</i>	Heading for the following dimension-related menu options.
Sort	Sorts the elements of the selected dimension as desired.
Mask	Selects a subset of the elements in the selected dimension.
Label	Enables you to edit the label for the selected dimension.
Highlight Selected	Highlights the selected element of the dimension.
Clear Highlighted	Clears the highlighting from any element in the visualization.
Series Legend	Color-codes the elements within the selected dimension as a series. See Enable a series legend on page 421.
Add Visualization	Adds a visualization that uses the selected dimension.
Save Dimension	Saves the selected dimension from the visualization.
Hierarchy View	Changes the current page visualization into a page hierarchy visualization, which displays the pages in a site organized hierarchically by file name and sorted alphabetically. See Apply hierarchy views on page 422.

Menu Option	Description
Layout	Heading for the following layout-related menu options.
Add Dimension	Adds another dimension to the visualization.
Change Dimension	Changes the dimension in the visualization.
Move Dimension	Moves the selected dimension to another axis in the visualization.
Remove Dimension	Removes the selected dimension from the visualization.

Subsets

Navigation title:

A data subset enables you to select and query from only those elements of interest to you in your dataset. This results in faster Data Workbench calculations and, therefore, faster query results.

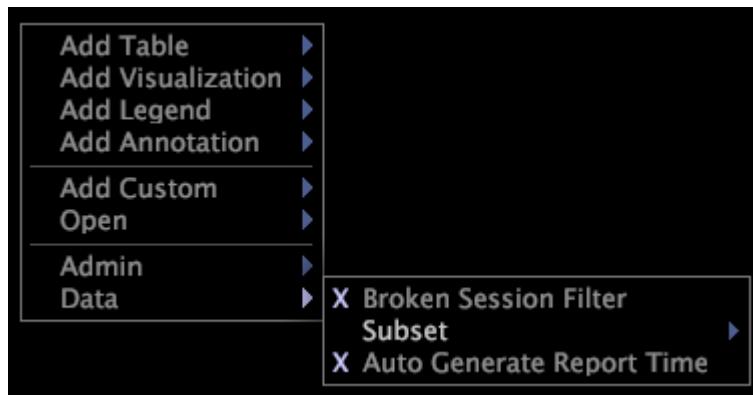
Understanding subsets

Navigation title:

Conceptual information about subsets.

When using a subset, please keep the following things in mind:

- All of your benchmarks now relate to your subset, not to the entire dataset, which is much more useful when analyzing a specific subset. See [Benchmarks](#) on page 285.
- Using a subset affects all of your workspaces because the subset is applied globally to Data Workbench.
- Subsets affect only metrics and denormal dimensions, not normal dimensions.
- When using Report, subsets do not affect the data in reports published for others to view.
- Once applied, your subset is in effect for all subsequent work in the profile, including the next time that you open this instance of Data Workbench, until you remove it.
- The only place that indicates that a subset has been applied is the context menu that you reach by right-clicking within a workspace.



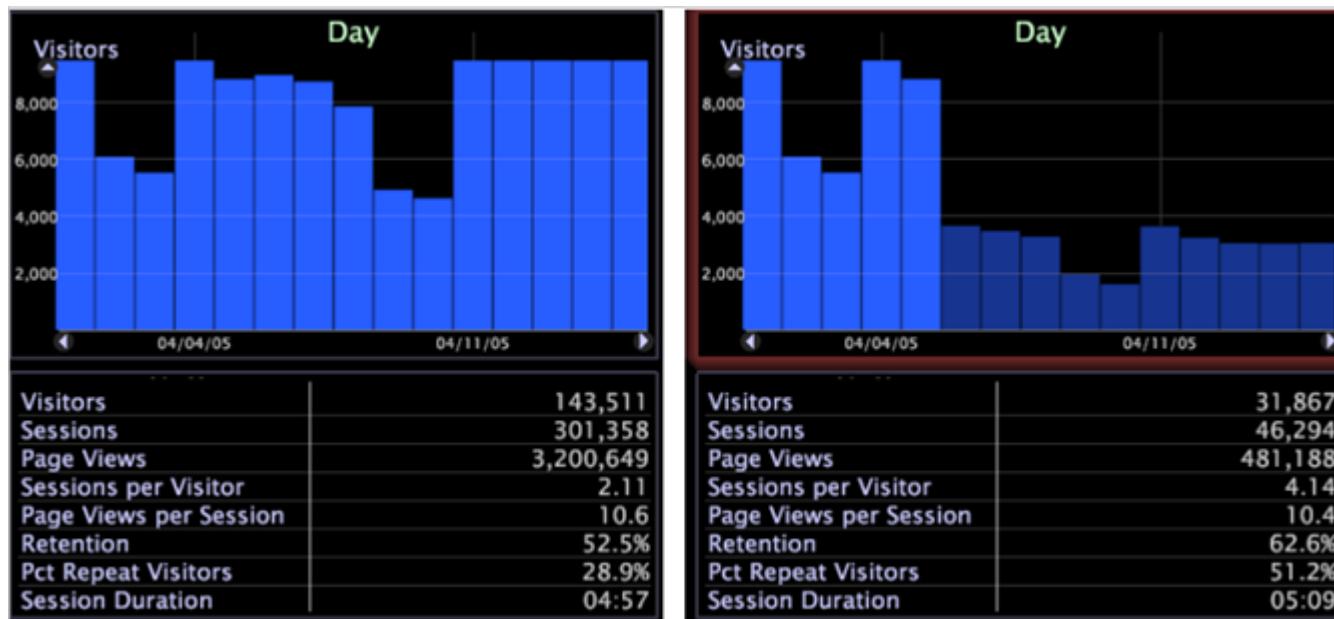
- You must be working online to change or remove a subset. If you are working offline and have applied a subset, you cannot view results from the entire dataset. See [Working offline and online](#) on page 201.



Note: The size of your subset is limited to the amount of data in your filter that resides on a single Data Workbench server. Therefore, if your dataset spans a Data Workbench server cluster, data for your subset comes from only one Data Workbench server in the cluster.

A user at a large retailer wants to create a subset (local cache) of one particular work week of data and then run queries only on that week of data. To do this, the user creates a subset for the days of interest.

The following example shows you a bar graph of Visitors over time and a Traffic metric legend. The first figure contains no selections: all of the data in the dataset is represented. The second figure shows data for a subset of Days = {...} by Visitors, in which Days is based on a selection of the elements April 1st through April 5th in the Day dimension.



Select a level

Navigation title:

When you create a subset, you must specify a level.

A level is any countable dimension. For example, if you are working with website data, if you select the element Tue from the Day of Week dimension and create a subset, you must select the level that you want to view: Page View, Session, or Visitor.

- **Day of Week=“Tue” by Page View:** The page view level shows you only those page views that occurred on a Tuesday.

Day of Week	Sessions
Mon	0
Tue	7,673
Wed	0
Thu	0
Fri	0
Sat	0
Sun	0

- **Day of Week=“Tue” by Session:** The session level shows you only those sessions that occurred on a Tuesday.

Day of Week	Sessions
Mon	0
Tue	10,135
Wed	0
Thu	0
Fri	0
Sat	0
Sun	0

- **Day of Week=“Tue” by Visitor:** The visitor level shows you all of the visitors who came to the site on Tuesdays, but it also shows other days that those same visitors came to the site.

Day of Week	Sessions
Mon	574
Tue	10,109
Wed	3,300
Thu	2,747
Fri	1,938
Sat	0
Sun	0

Create subsets

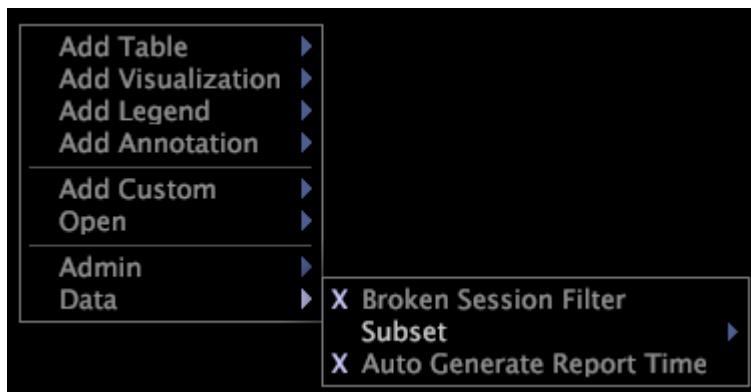
Navigation title:

Steps to create a subset.



Note: To create or change a subset, you must be working online.

1. While working online, make the desired selections within one or more visualizations in your workspace.
For more information about interpreting selections, see [Make selections in visualizations](#) on page 281.
2. Right-click within the workspace and click **Data > Subset > <level name>**



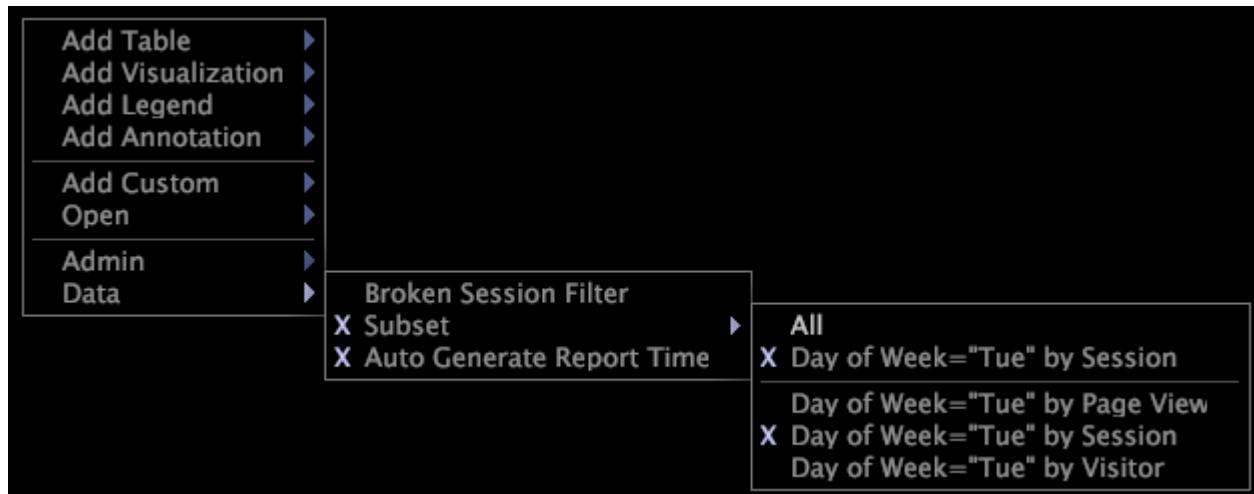
An X displays to the left of the **Subset menu** option and to the left of the subset's level.

Remove a subset

Navigation title:

Steps to remove a data subset.

- While working online, right-click within the workspace and click **Data > Subset > All**.



Zoom in or out in a visualization

Navigation title:

You can zoom in or out on data in graphs, globes, process maps, and scatter plots.

Zooming on objects in process maps and globes

To zoom in on an area

- Point to the desired area, press and hold both mouse buttons, then move the mouse forward (away from you).

To zoom out on an area

- Point to the desired area, press and hold both mouse buttons, then move the mouse backward (toward you).

Zooming on axes in graphs and scatter plots

Both graphs and scatter plots have axes that you can zoom in or out on separately (that is, you can zoom in on one axis, then zoom out on the other).



Note: For the y-axis in a line graph and the x- or y-axis in a scatter plot, the zoom resets every time the underlying data changes. Because this happens quite often when working online, Adobe recommends that you work offline when you need to zoom these axes. See [Working offline and online](#) on page 201.

To zoom in on the x- or y-axis of a graph

When you zoom vertically (along the y-axis), the x-axis remains fixed and the data stretches upward.

1. Place your mouse pointer on the x- or y-axis metric or dimension label area of the graph.
2. Press both mouse buttons and move the mouse forward (away from you). Arrows appear on the axis indicating that more data is available beyond the portion of the visualization currently displayed.

To zoom out on the x- or y-axis of a graph

1. Place your mouse pointer on the x- or y-axis metric or dimension label area of the graph.

-
2. Press both mouse buttons and move the mouse backward (toward you).

To zoom in on all data on the x- or y-axis of a graph

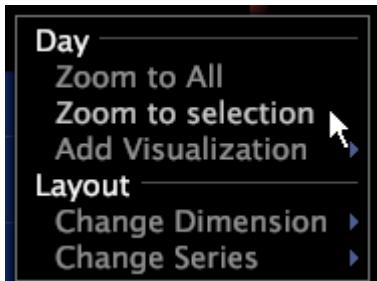
The **Zoom to All** menu command returns to a display of all of the data on the selected axis.

- Right-click the x- or y-axis border areas (near the data labels) and click **Zoom to All**.



To zoom in on data selected in graphs

1. Click and drag the mouse pointer across the range of data you want to select.
2. Right-click in the dimension label area and click **Zoom to Selection**.



Arrows appear on the axis indicating that more data is available beyond the portion of the visualization currently displayed.

Adding callouts to a workspace

Navigation title:

Callouts are windows that you add to a workspace to bring attention to a particular dimension element by creating a new visualization with a virtual selection of that element.

Data Workbench is delivered with a standard set of callout types. Because your implementation can be fully customized, the available callout types that appear in your implementation may differ from what is documented in this guide.

By default, Data Workbench provides the following callouts:

- [To add an annotation callout to a visualization](#) on page 294
- [To add a blank table, line graph, or scatter plot callout to a visualization](#) on page 294
- [To add a blank table, line graph, or scatter plot callout to a visualization](#) on page 294
- [To add a blank table, line graph, or scatter plot callout to a visualization](#) on page 294
- [To add a confidence legend callout to a visualization](#) on page 295
- [To add a metric legend callout to a visualization](#) on page 295

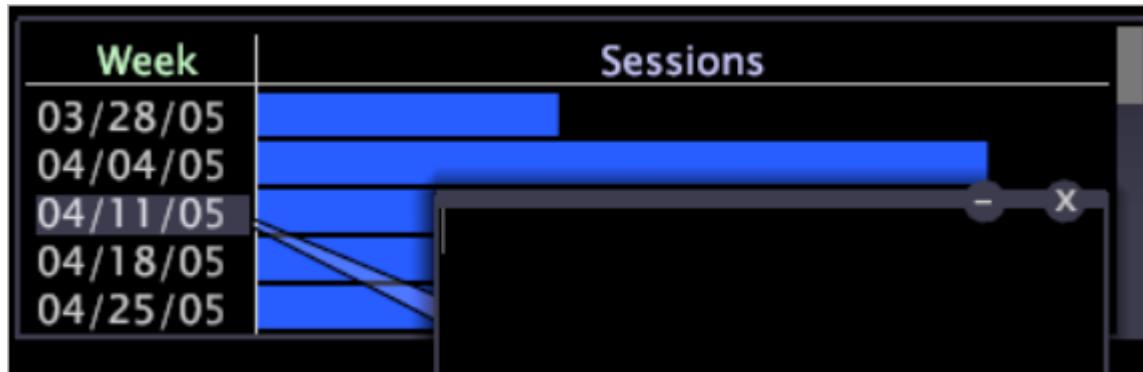


Note: Callouts do not function as selections (that is, they do not affect other visualizations within the workspace) unless you make a selection within the callout.

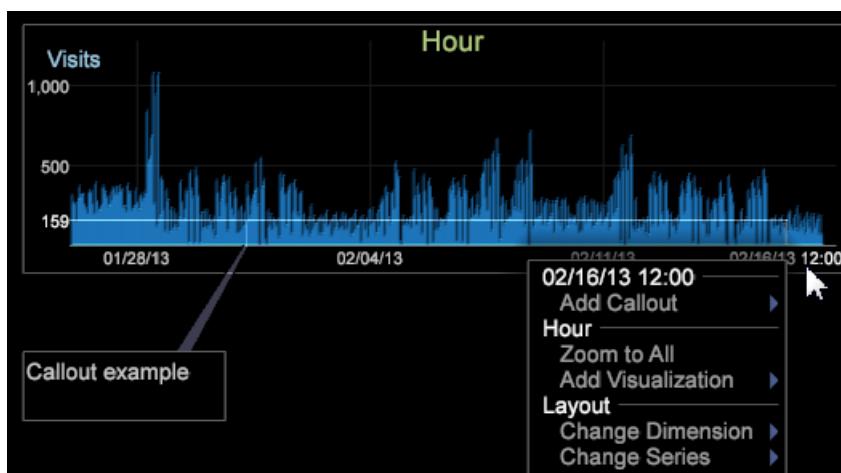
You can add or edit callout definitions by configuring the callout files stored in the *profile name\Context\Callout* folder of the Server installation folder. See [Configure a callout](#) on page 500.

To add an annotation callout to a visualization

1. Right-click the element for which you want to create a callout, then click **Add Callout > Annotation > Image** or **Add Callout > Annotation > Text**. A blank window displays with a visible connection to that element.



To add callouts to Graph visualizations, you need to right-click at the bottom of the visualization (the base axis) to open a menu.

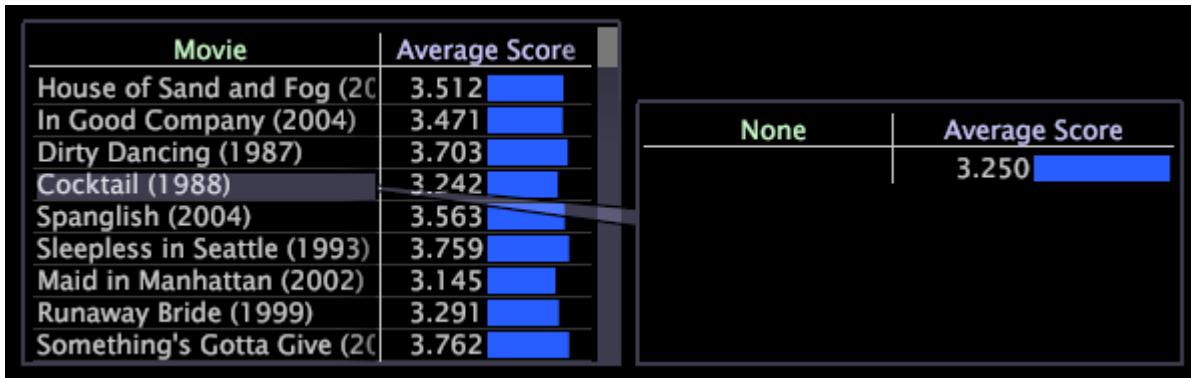


2. Depending on your selection, complete the appropriate step:
 - For a text annotation, type or paste the desired text into the callout, then format the text as appropriate. See [Text annotations](#) on page 308.
 - For an image annotation, paste the desired image into the callout by copying the image, then right-clicking within the callout. Click **Paste image**. See [Image annotations](#) on page 310.

To add a blank table, line graph, or scatter plot callout to a visualization

1. Right-click the element for which you want to create a callout and click **Add Callout > <callout type>**.

The following example shows a Blank Table callout.

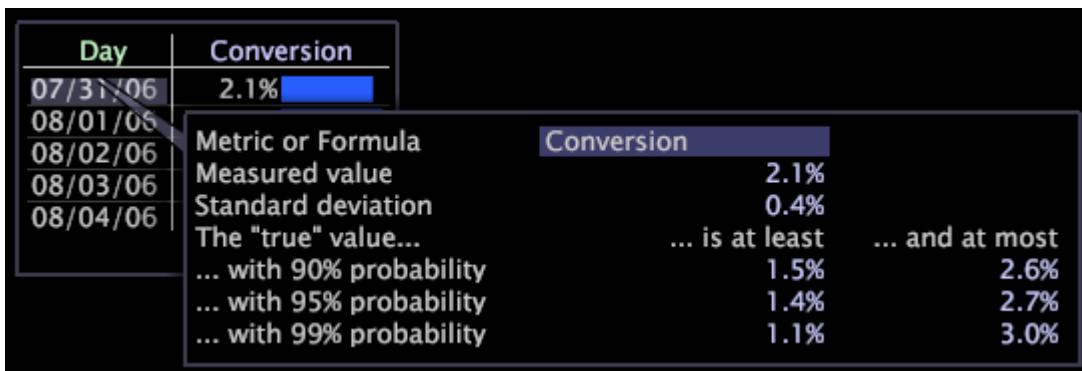


- To select a dimension, right-click **None** and click **Change Dimension > <dimension name>**.

Note: If you change the dimension within a visualization that has a callout, the callout changes from being connected to the element of the original dimension to being connected to the entire visualization.

To add a confidence legend callout to a visualization

- Right-click the element for which you want to create the callout and click **Add Callout > Confidence Legend**.

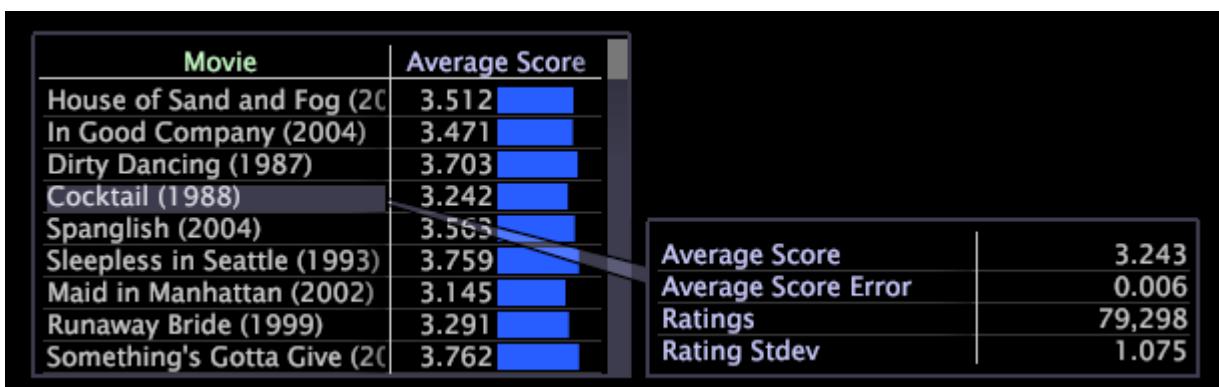


- If desired, change the **Metric or Formula** field.

For expression syntax rules, see [Query language syntax](#) on page 515. See [Confidence legends](#) on page 379.

To add a metric legend callout to a visualization

- Right-click the element for which you want to create the callout and click **Add Callout > Metric Legend**.



-
2. If desired, add metrics to or remove metrics from the metric legend.

See [Metric legends](#) on page 380.

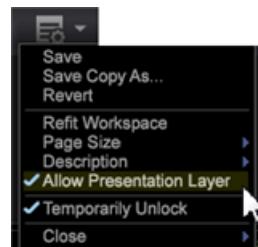
Annotating visualizations

The Presentation Layer lets you mark up and annotate your workspace visualizations and then publish with your call-outs and comments. Add text descriptions, graphic objects, callout arrows, color coding, images, and other features in an overlay to add annotations and clarify important data points, and then share with stakeholders.

Add annotations to your visualizations:

1. Open a workspace.
2. Go to **File > Presentation Layer**.
3. Click **Allow Presentation Layer**.

Toggle the command in this menu to allow or disallow generating a presentation layer. A checkmark will appear when the presentation layer is allowed.



When you allow the presentation layer, a clickable icon will appear in the right side of the toolbar.

 This clickable icon shows that you are in the **Hide Presentation Layer** mode but not able to see annotations.

4. Click icon to toggle to **Show Presentation Layer**.

View annotations in the presentation layer.

 This clickable icon shows that you are in the **Show Presentation Layer** mode but not able to edit.

 **Note:** You can also right-click in the workspace to change modes using menu commands.



5. **Edit and add callout features.**

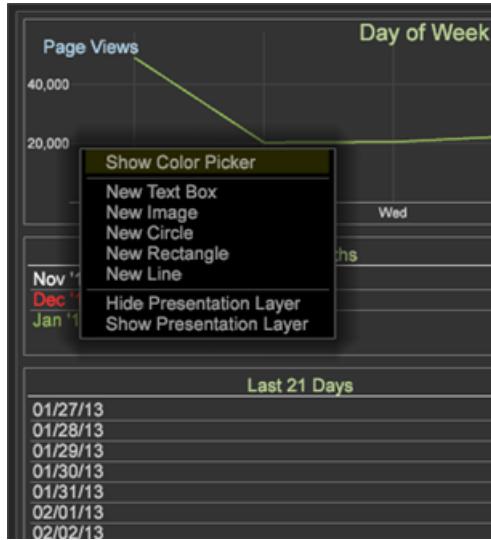
Click icon again to capture a static representation of the workspace visualization and open an overlay to add or edit annotations.

 This clickable icon shows that you are in the **Edit Presentation Layer** mode to add and edit annotations.

A gray overlay appears while in edit mode.

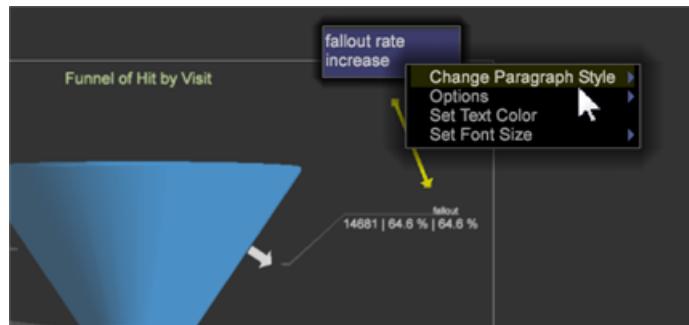
6. **Add callout features.**

Right-click the visualization to add a new text box, arrows, and other callout features to highlight and annotate for presentation.



7. Set graphic features.

Select a graphic object and right-click to set options. You can set colors for a box or other graphic object, adjust text settings in a text box, and add arrows or set layering for lines.

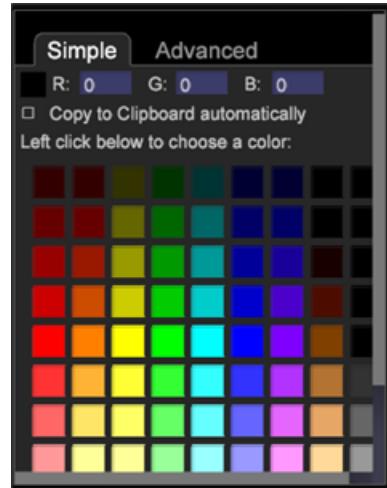


8. Add images to the workspace.

Right-click to add your .png and .jpg images to the overlay and resize.

9. Assign colors to graphic objects.

You can select objects in the presentation layer and assign colors using the color picker.



10. Export and share presentation layer.

After annotating your workspace visualizations, you can export the presentation layer with visualizations and share as a .png graphic file.

Click **Export > Export PNG**.

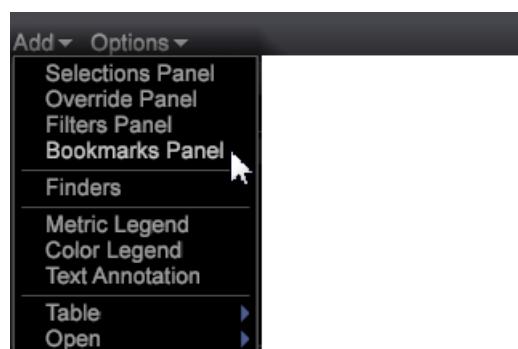
Add a bookmarks

You can bookmark significant workspaces to quickly move between visualizations and reports employed in your workflow.

1. Bookmark a workspace by clicking the Bookmark icon  in the upper right corner of the toolbar.

Before you have a workspace bookmarked, this icon appears as just a book. The bookmark icon changes to a book with a marked page once a workspace has been bookmarked.

2. Click **Add > Bookmarks Panel** in the left pane to open a list of bookmarks.



3. To open a bookmarked workspace, click a workspace name in the **Bookmark Panel**.

The screenshot shows the Adobe Analytics Data Workbench interface. On the left, the Bookmark Panel is open, displaying a list of bookmarked workspaces: Score Visitors, Clustering, Statistics Callout, Correlation Drag & Drop Dimension Elements, Correlation, Binomial Correlation, and Lift Gain Chart. The main workspace is titled "Correlation Drag & Drop Dimension Elements" under "Predictive Analytics". It displays a "Correlation Matrix" table with data for Day, Visits, Revenue, and Page Views. Below it is another smaller table for "Day" with columns for Visits, Revenue, and Page Views.

The selected workspace will open. When you click another bookmarked workspace, the previous workspace will close and the newly selected workspace will open, allowing you to quickly navigate through your workflow.

To delete a bookmark:

- In the Bookmark Panel, right-click and select **Remove <bookmark title>** to delete a selected bookmark, or select **Clear All Bookmarks** to delete all bookmarks.
- You can also right-click on the workspace in the thumbnail view within the worktop and select **Clear Bookmark**.



Important:

- 25 bookmarks can be saved.
- If you add a bookmark and then move the location of the workspace, the bookmark will be invalid and must be deleted from the Bookmark Panel and reset.

Metric Dim Wizard

Use the Metric Dim wizard to create a new Metric Dimension.

A Metric Dim converts a metric into a new dimension. For example, a Metric Dim based on a metric of Page Views and level of Visitor will display dimension elements based on the total Page Views for each Visitor. It lets you extend a currently defined metric based on dimension elements to create and save as a new dimension.

Step 1: select dimension and metric

1. Open the Metric Dim Wizard.

In a workspace, right-click and select **Tools > Create Metric Dim**.

2. Name the Metric Dim.

As a default, the Name field will auto-populate based on Level and Metric selections.

3. Select a Dimension Level.

The dimension level is the parent dimension containing all constituent element values to filter input and define a dimension type.

Dimension levels include:

- Clickthrough
- Hit
- Product
- Visit
- Visitor

4. Select a Metric.

Select a pre-built metric to extend and save as a metric dim.

Create a Metric Dim Step 1 of 3

Select a Metric and a Level, and name your new Metric Dim.

Name: Visits per Product

Select Level: Product ▾

Select Metric: visits

Create Metric Form

Visits
Page Views
Visitors by Visit
Traffic
Custom Events
Duration
Commerce
Products
Traffic Source Clickthrough
Confidence

Help Preview | Cancel | < Back | Next > | Finish

5. (optional) Create a Metric Formula.

Click the box to enter a custom metric formula. The calculated Preview value will appear validating the expression.

Create a Metric Dim Step 1 of 3

Select a Metric and a Level, and name your new Metric Dim.

Name: Visits per Product

Select Level: Product ▾

Select Metric: sum(one, Product) ▾

Create Metric Formula

Formula: sum(one, Product)

Preview: 319,408.00

Help Preview | Cancel | < Back | Next > | Finish

You can add your own *metric expression* or cut and paste from another metric editor or visualization. Syntax errors, formula errors, undefined filters, and other errors are reported in the wizard.

6. Click **Next**.

Step 2: format and set buckets

You can select the metric format and set the bucket values for a dimension expression.

1. Select a **Format** for the new metric dim.



The format defines how the metric will be presented when opened in a visualization. These formats are selected *printf standards*, defined below:

```
%[flags][width][.precision][length][specifier]  
%0.2lf = % _ [flags] 0 [width] .2 [.precision] 1 [length] f[ specifier]
```

In the **Preview** field, a value will appear based on the metric and format selected.

2. Add **Bucket Count** expression.

You can define a metric dim with various ranges, or buckets. This returns subsets of elements based on size, such as [0-4], [5-10],...). Elements of the Dimension Level relate to the elements whose range contains the value of metric. See the bucket expression description at *Syntax for Dimension Expressions*.

3. “Click **Preview** to open table of Metric Dim values before saving.”



The table details metric values per metric dim.

4. Click **Show in Dimension Menu** to add the newly created dimension to the **Dimension** tab in the **Finder**.
5. Click **Next**.

Step 3: finish and save

1. Select to launch the Metric Dim Editor, graph visualization, or table after saving.

Field	Description
Launch Metric Dim Editor	Open the Metric Dim Editor.
Launch Graph	Launch a PNG graphic of the table.
Launch Table	Launch a table in the workspace with values in columns listing values of the new metric dim compared with values of the selected metric.

2. Click **Finish** and save.

A save dialog will open allowing you to save the file. The selected options to view values will open in the workspace.

Editing the Metric Dim

Edit a Metric Dim in the Profiles Manager.

To edit an existing Metric Dim:

1. In a workspace, right-click **Admin > Profile Manager**. Open **Dimension** and select the metric dim.

Search:

File	Base	Predictive	Geograph	Adobe SC	Advanced	MKY Geom	Premium	MKY Geom	User
ChangeList.txt	✓								
Profile.cfg								✓	✓
profile.cfg.off	✓								
Color Legends\	✓					✓			
Colors\	✓								
Context\	✓				✓				
Dataset\	✓	✓	✓	✓	✓	✓			
Dimensions\									
Compatibility.cfg	✓								
LastNDIM Example	✓								
MetricDim Example	✓								
Mikes Metric Dim.dim								✓	
None.dim	✓								
RenameDim.example	✓								
order.txt					✓				
Acquisition\						✓			
Advanced Attribution\							✓		
Campaign\					✓				
Custom Conversion\							✓		
Custom Traffic\								✓	
Geography\									✓
Mobile\					✓				
Paths\					✓		✓		
Products\					✓				
Scoring\									
Site Content\					✓				
Site Metrics\					✓				
Time\					✓				
Traffic Sources\					✓				
Value\						✓			
Visitor Profile\					✓				
Export\	✓							✓	
Filters\					✓				✓
Help\	✓								
Images\	✓								
Localization\	✓								
Maps\									✓
Menu\	✓	✓	✓	✓	✓	✓	✓	✓	✓
Metrics\	✓				✓	✓			
Models\	✓								
OLD\	✓								
Reports\	✓							✓	✓
Sidebar\	✓								
Wizards\	✓								
Work\									
Workspaces\	✓	✓	✓	✓	✓	✓	✓	✓	✓

Editing Cart V
Name Cart V
Formula Hits
Preview 141,72

- Right-click the checkmark in the user column, select **Open** and then select metric dim.

The screenshot shows a file browser window with a sidebar containing navigation links like 'Create a Metric Dim', 'First, select a Metric at', 'Select Level: Visitor', 'Select Metric: Visits', 'Create Metric Formul...', 'Help', and 'Editing Cart V'. The main area displays a grid of files with columns for 'File', 'Base', 'Predicti', 'Geograph', 'Adobe SC', 'Advanced', 'MKY Geo', 'Premium', 'MKY Geo', and 'User'. A context menu is open over a file named 'Mikes Metric Dim.dim', showing options like 'Open', 'To', 'Rename', 'Copy', and 'Path Dimensions'. The status bar at the bottom shows 'Preview 141,72'.

File	Base	Predicti	Geograph	Adobe SC	Advanced	MKY Geo	Premium	MKY Geo	User
ChangeList.txt	✓								
Profile.cfg	✓								
profile.cfg.off	✓								
Color Legends	✓								
Colors\	✓								
Context\	✓								
Dataset\	✓	✓	✓	✓	✓	✓			
Dimensions\	✓								
Compatibility.cfg	✓								
LastNDim.example	✓								
MetricDim.example	✓								
Mikes Metric Dim.dim									
None.dim	✓								
RenameDim.example	✓								
order.txt									
Acquisition\									
Advanced Attribution\									
Campaign\									
Custom Conversion\									
Custom Traffic\									
Geography\									
Mobile\									
Paths\									
Products\									
Scoring\									
Site Content\									
Site Metrics\									
Time\									
Traffic Sources\									
Value\									
Visitor Profile\									
Export\	✓								
Filters\									
Help\	✓								
Images\	✓								
Localization\	✓								
Maps\									
Menu\	✓	✓	✓	✓	✓	✓	✓	✓	✓
Metrics\	✓								
Models\	✓								
OLD\	✓								
Reports\	✓								
Sidebar\	✓								
Wizards\	✓								
Work\									
Workspaces\	✓	✓	✓	✓	✓	✓	✓	✓	✓

3. In the Edit Metric Dim box, make changes as required.

The screenshot shows the properties of a metric dimension named 'Mikes Metric Dim.dim'. The properties listed are: entity (MetricDim), +Format (printf_format), Bucket Count (1024), Bucket Size (1), Bucket Start (1), Level (wdata/model/dim/Product), Metric (wdata/model/metric/Default Metric), and show (true). The status bar at the bottom shows 'Preview 141,72'.

```

Search: -Mikes Metric Dim.dim
entity MetricDim
+Format printf_format
Bucket Count 1024
Bucket Size 1
Bucket Start 1
Level wdata/model/dim/Product
Metric wdata/model/metric/Default Metric
show true

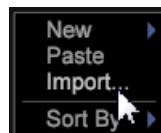
```

Analysis visualizations

Navigation title:

Information about the numerous analysis visualizations available in Data Workbench.

- You can open any visualization by right-clicking within a workspace and selecting the desired menu option, or by clicking **Ctrl + n**.
- Visualizations cannot be imported like workspaces. When you right-click in the worktop and select **Import**, you can import an existing workspace, but not a visualization residing outside of the workspace.



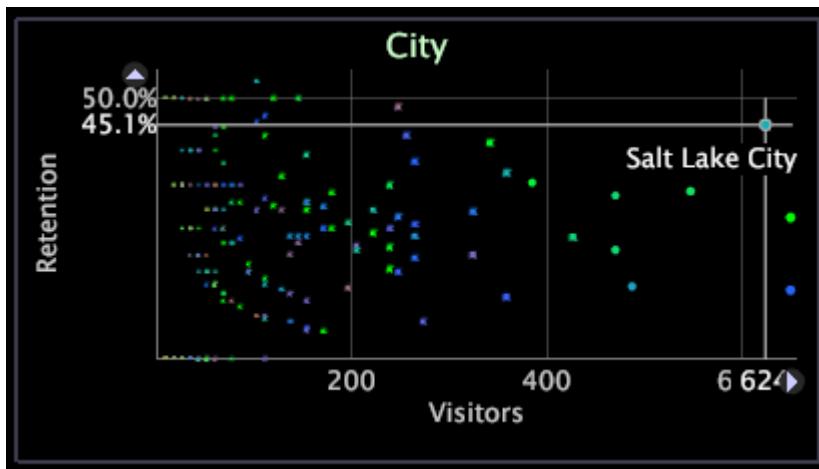
If you attempt to import a visualization, you will receive a "not a valid workspace" error.

2D scatter plots

Navigation title:

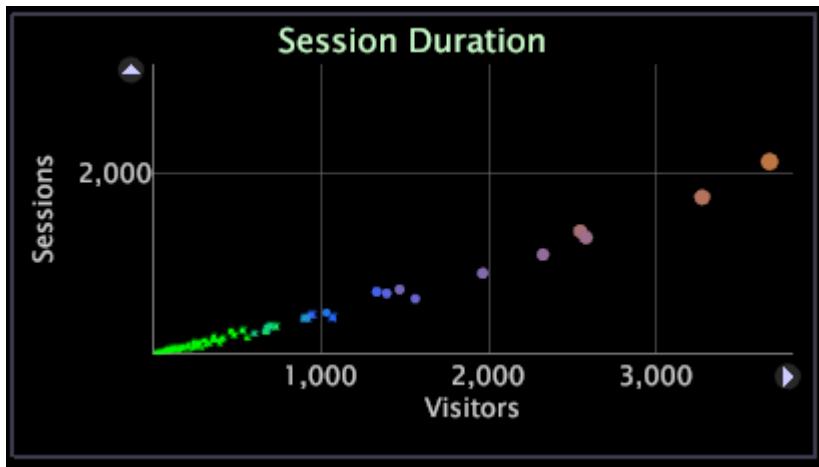
Scatter plots graph the elements of a data dimension (such as Page or City) on a grid in which the x- and y-axes represent different metrics.

Scatter plots can be useful when trying to understand the relationship between large numbers of disparate items, by two different metrics. In the following example, the scatter plot is showing each city by the number of visitors and the respective retention rate.



The scatter plot enables you to quickly see the outliers. Salt Lake City, for example, has a higher than average retention rate per visitor.

Scatter plots can also be used to show the consistency of data. In the following example, the scatter plot shows the number of visitors with sessions of a particular length.



The size of each point on the scatter plot is determined by the radius metric. The default radius metric differs for each Adobe application. For example, in **Site**, the radius metric is based on Sessions by default. You can change the radius metric to have the points in your scatter plots represent any available metric. For steps to do so, see [Changing radius metrics](#) on page 306. The color of the points is based on the color legend that is open within the workspace. For more information about color legends, see [Color legends](#) on page 378.

Select points

To select a single point

- Click the point.

To add another point or group of points to your selection

- Ctrl+click a point or Ctrl+drag across multiple points.

To remove a point or group of points from your selection

- Shift+click a point or Shift+drag across several points.

Changing dimensions

- Right-click the label of the dimension at the top of the graph and click **Change Dimension > <dimension name>**.

Changing metrics

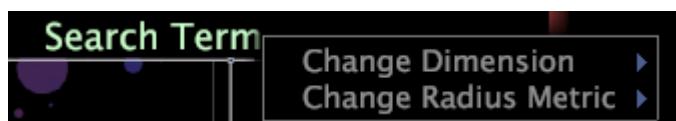
To change the metric shown on the x- or y-axis of a scatter plot

- Right-click the label of the metric that you want to change and click **Change Metric > <metric name>**.

Changing radius metrics

To change the radius metric of a scatter plot

Right-click the label of the dimension at the top of the graph and click **Change Radius Metric > <metric name>**.



3D Scatter Plots

Navigation title:

A 3D Scatter Plot graphs the elements of a data dimension (such as Days or Referral Site) on a three-dimensional grid where the x, y, and z axes represent various metrics.

Like the [Scatter Plot 2D](#), this visualization is useful when trying to understand the relationship between large numbers of disparate items employing different metrics.

To employ the 3D Scatter Plot visualization:

1. Open a new workspace.

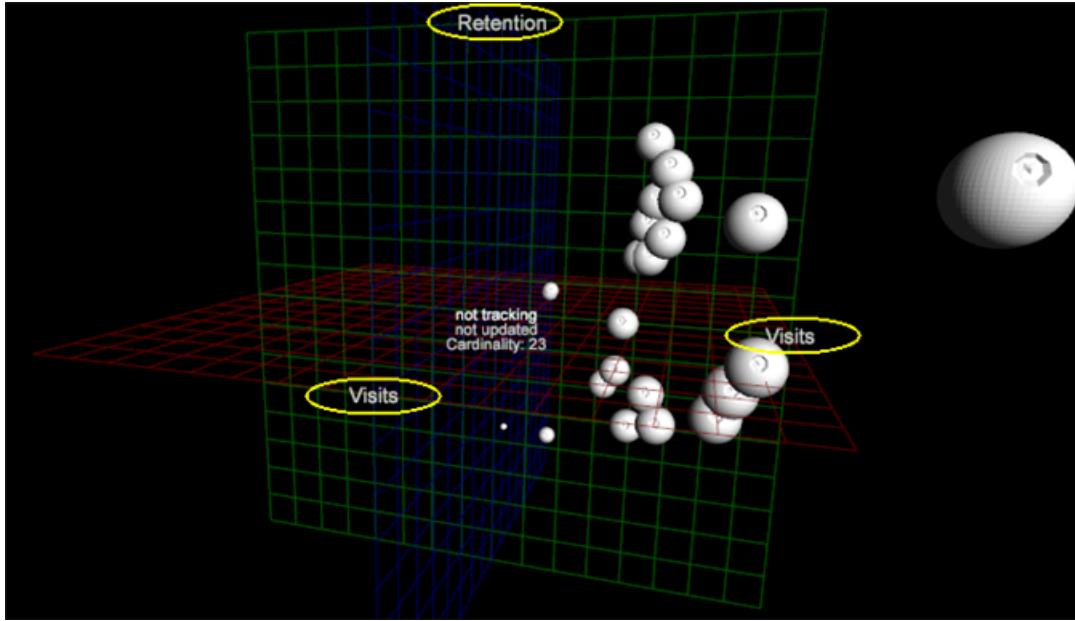
After opening a new workspace, you may need to click **Add > Temporarily Unlock**.

2. Right-click and select **Visualization > 3D Scatter Plot**.

A menu listing **Dimensions** will open.

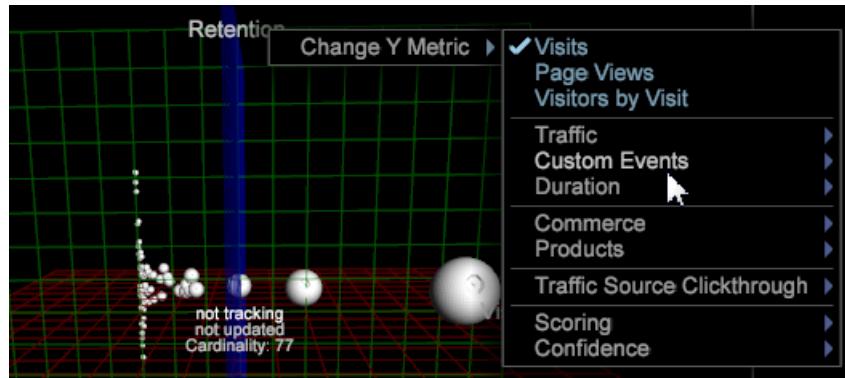
3. Select a dimension for the query.

The 3D Scatter Plot will open the default metrics for that dimension.



Selecting the **Days** menu displays the following 3D Scatter Plot with these default metrics on the following axes: **x=Visits**, **y=Retention**, and **z=Visits**.

4. Change metrics. Right-click on the metric label in the x, y, or z axis and select **Change Metric**. Then select a different metric for the selected axis.

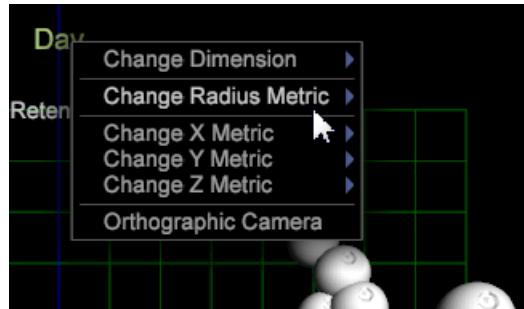


Important:

- Drag a metric to one of the three axis labels and drop it to change the selected axis to the dropped metric.
- Drag a metric anywhere else on the visualization and drop it to change the radius metric for that axis.
- Drag a dimension to anywhere on the visualization and drop it to change the dimension for the visualization.

5. Change the Radius Metric. Right-click the title at the top of the page (titled after the selected dimension) and select **Change Radius Metric**.

The radius metric defines the size of the plotted point based on the metric selection. The relative position of points does not change in the scatter plot, but the plotted point sizes within the visualization increase based on the metric value.



6. Employ the **Orthographic Camera**. This option lets you identify the plotted points in relation to their true perspective based on the radius metric to avoid three-dimensional distortion.

When the 3D Scatter Plot first appears, it displays in a three-dimensional rotating projection, which causes some distortion for points plotted nearer to the perspective, or virtual "camera." (The plots nearer to the camera show up much larger than the points rotating further away from the camera.)

To avoid this perspective distortion, you can select the **Orthographic Camera** option by right-clicking on the title and selecting from the menu. This allows you to represent the three-dimensional objects in two-dimensions. This renders the plotted points as flat and displays the points as relative to the radius metric, lessening the 3-dimensional offsets.

7. Select points from the scatter plot.

- **To remove a point or group of points:** Click the point.
- **To add another point or group of points to your selection:** **Ctrl + click** a point or **Ctrl + drag** across multiple points.
- **To remove a point or group of points from your selection:** **Shift + click** a point or **Shift + drag** across several points.

Annotations

Navigation title: Annotations

Annotations are windows that users can add to a workspace to contain descriptive comments, graphics, or dynamic titles.

Annotations also can be added to workspaces as callouts. For information about working with annotation callouts, see [Adding callouts to a workspace](#) on page 293.

Text annotations

Navigation title:

Text annotations are windows into which you can enter arbitrary text to add descriptive information or comments to a workspace.

You move or manipulate text annotations within a workspace in the same manner as other types of visualizations. You also can format the appearance of the text as well as name, edit, or save any text annotation that you add to a workspace. Text annotations also can be exported to Microsoft Excel and generated and distributed by Report.

For more information about Reports, see the *Insight Report Guide*.

Add or edit text

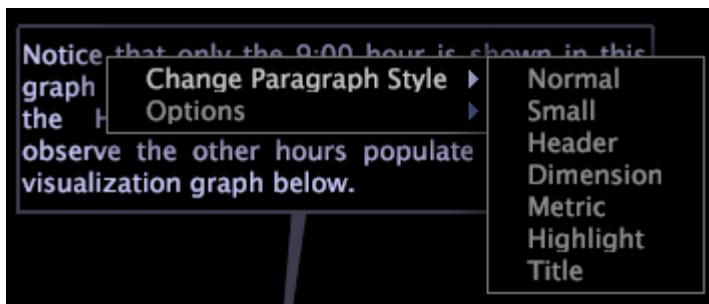
- Click within the body of the text to enter edit mode and begin typing.



Note: Text annotations support basic editing features, including cut (Ctrl+x), copy (Ctrl+c), paste (Ctrl+v), undo (Ctrl+z), redo (Ctrl+Shift+z), select section (click+drag), and select all (Ctrl+a). See [Work with text in windows](#) on page 227.

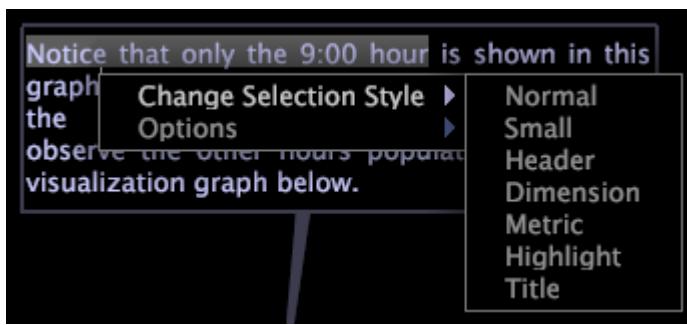
Apply formatting to paragraphs

1. Within the body of the text, right-click the paragraph that you want to format.
2. Click **Change Paragraph Style** > <*text style*>.



Apply formatting to selections

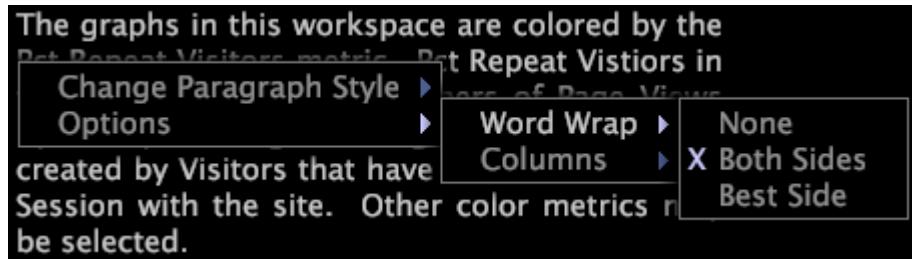
1. Within the body of the text, click+drag to select the text that you want to format.
2. Right-click the selection and click **Change Selection Style** > <*text style*>.



Change word wrap settings

This option controls whether and how the text within the annotation wraps around the visualizations in the workspace.

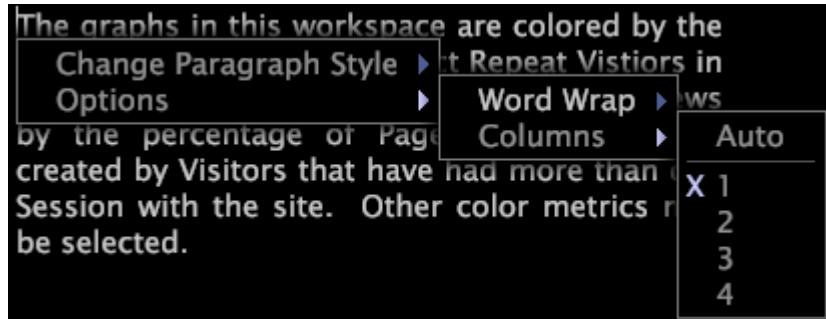
- Right-click within the body of the text, click **Options** > **Word Wrap**, then click the desired option:
 - **None:** This option removes word wrapping from this annotation.
 - **Both Sides:** This option allows the text within the annotation to wrap around both sides of visualizations in the workspace.
 - **Best Side:** This option allows the text within the annotation to wrap around only one side of visualizations in the workspace.



Change column settings

This option controls the number of columns in which to display the text in the annotation.

- Right-click within the body of the text, click **Options > Columns**, then click the desired option:
 - **Auto:** This option selects the most appropriate column layout for the width and length of the text.
 - **1-4:** These options enable you to display the text in one, two, three, or four columns.



Export to Microsoft Excel

For information about exporting windows, see [Export window data](#) on page 228.

Image annotations

Navigation title:

Using an image annotation, you can copy an image from any source and paste it directly into a workspace.

You can then resize, invert the color scheme, or dim the image.

Add an image

1. Copy any image to your clipboard by right-clicking the image and clicking **Copy**.
2. Right-click within the image annotation and click **Paste Image**.



Note: You can add only one image to an image annotation. Pasting a new image into the annotation pastes over the existing image. To add multiple images to a workspace, you must add another image annotation.

Resize an image

- Click+drag any side or corner of the annotation.

Invert an image

- Right-click within the annotation and click **Display Mode > Invert (white to black)** or **Display Mode > Invert (black to white)**.

Dim an image

- Right-click within the annotation and click **Display Mode > Dim** or **Display Mode > Very Dim**.

Return to original display mode

- Right-click within the annotation and click **Display Mode > Original**.

Fit to window

To enable or disable the Fit to Window option

The **Fit to Window** option, when selected, lets the annotation automatically expand or shrink to fit the image size. The **Fit to Window** option is enabled by default.

- Right-click within the annotation click **Fit to Window**. When enabled, an X appears to the left of **Fit to Window**.

Dynamic title annotations

Navigation title:

Using Report, you can generate reports dynamically for the dimension elements that you specify in a lookup file or for a particular number of the dimension elements, such as for the users with the 10 highest order counts.

Using a dynamic title annotation, you can display a title in a dynamic report that identifies either the dimension element name currently being processed by Report or the dimension element name that you have selected to view. Dynamic title annotations enable you to identify which element's data you are viewing on the report without altering any of the other visualizations or accidentally sorting a selection by a particular element.



Note: This functionality is available only when you have licensed Report. For more information about Report, see the *Data Workbench Report Guide*.

For more information about dynamically generating reports, see the *Insight Report Guide*.

To set up a dynamic title annotation

- Right-click within the dynamic title annotation and click **Change Dimension > <dimension name>**.



- To view the information for one particular element, click **Change Element** and click the desired element to preview the report for that particular element.



Exporting to Microsoft Excel

For information about exporting windows, see [Export window data](#) on page 228.

Association Chord visualization

The Association Chord visualization allows you to show both the proportion and association between metrics, dimensions, and elements, displaying larger chords as an indication of a stronger association.

The Associations Table compares values with Cramer's V calculation rather than using Pearson's correlation coefficient as employed in the [Correlation Matrix](#) and [Correlation Chord](#) visualizations (these can only compare metrics, while the Association Table and Chord can compare metrics, dimensions, and elements). The Associations Chord also provides another view into a previously built [Association Table visualization](#) on page 314.

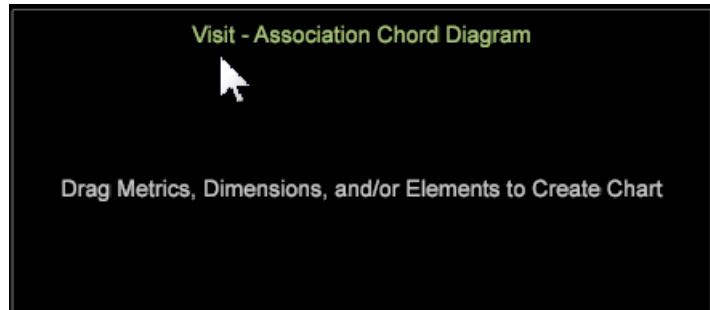
To build an Association Chord

1. In a workspace, right-click **Visualization > Predictive Analytics > Association Chord**.

A menu will open allowing you to select an extended dimension from the list.



Once selected, the blank Association Table will open with the selected Dimension identified in the title.



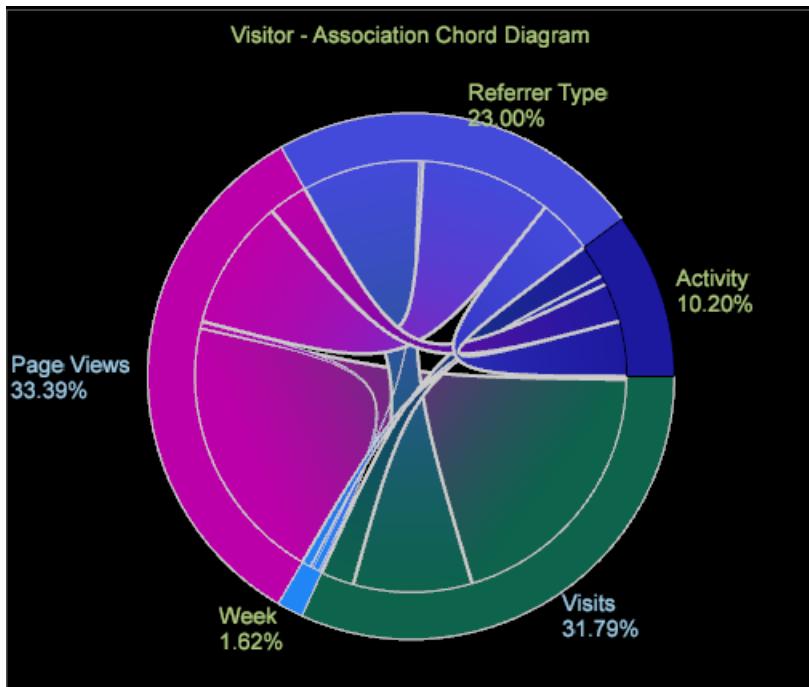
2. **Select a metric, dimension, or dimension element.**

Right-click the chord visualization and select **Add Metric** or **Add Dimension**. Select items from the menu to add to the chord.

You can also drag metrics and dimensions from the **Finder** by clicking **Ctrl-Alt** and dragging metrics and dimensions to the chord. Or drag dimension elements directly from an open table to the chord visualization.

3. **Choose additional metrics, dimensions, and elements to associate.**

After two or more values are selected, the chart will automatically refresh and begin displaying association data. Continue adding metrics as needed to associate data points.



The Chord visualization displays the proportion of the whole represented by the area of each segment. Continue to add metrics/dimensions/elements as need to identify and investigate significant relationships.

4. **View the Chord visualization.**

Hover over each value in the visualization to see relationships.

5. **Change Settings.**

Right-click the chord visualization to open a menu to change the metric, dimension, or elements display the dimensions as absolute numbers or as percentages, remove the selected metric or all metrics, edit colors and details, and export values to an Associations Table.

To build an Association Chord from an Association Table:

1. Open an **Association Table** visualization.
2. Right-click and select **Export Chord Visualization**. An Association Chord diagram will open with values selected in the Association Table.



		Association Table	
		Week	Page Views
Visits	1.000	0.007	0.591
Week	0.007	1.000	0.014
Page Views	0.591	0.014	1.000



Important: Exporting an Association Table from an Association Chord Diagram that contains at least one metric will result in duplicated elements in the rows/columns of the Association Table. To avoid duplicated elements, create a new Association Table and add the desired elements rather than exporting the elements from an Association Chord Diagram.

Association Table visualization

The Association Table visualization let you associate metrics with metrics, dimensions, and dimension elements using Cramer's V algorithm.

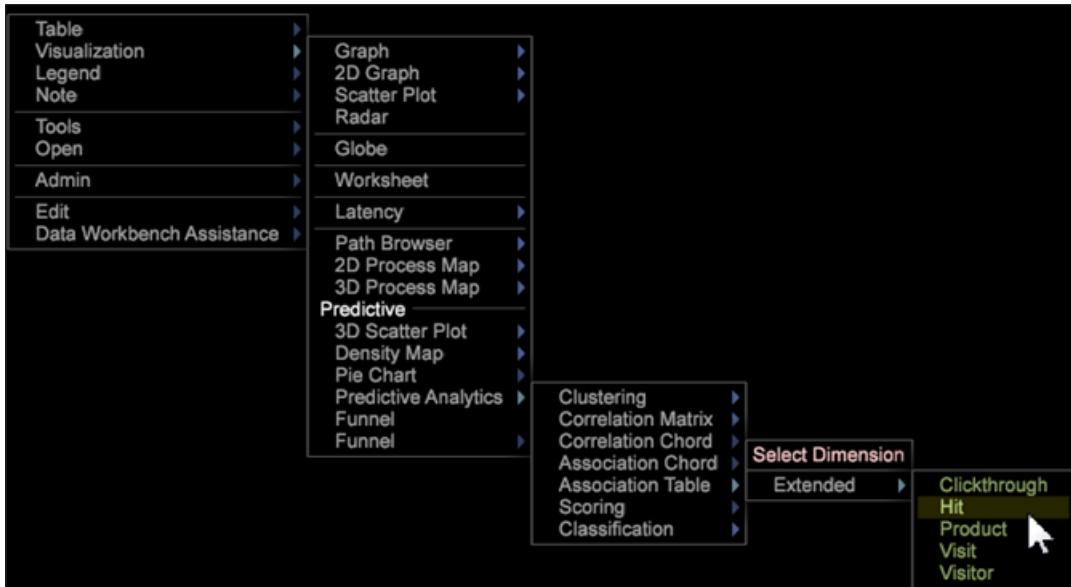
The Association Table compares values with Cramer's V calculation rather than using Pearson's correlation coefficient as used in the [Correlation Matrix](#) and [Correlation Chord](#) visualizations (these can only compare metrics, while the Association Table and [Association Chord visualization](#) on page 312 can compare metrics, dimensions, and elements).

Build an Association Table

The Association Table compares metrics over a countable or non-countable dimension. The table can be modified to highlight associations within the visualization through color picking or to render it as a text map, heat map, or both.

1. Open an Association Table.

Right-click **Visualization > Predictive Analytics > Association Table**.



2. Select an extended dimension—a Clickthrough, Hit, Product, Visit, or Visitor dimension. An Association Table will open with the extended dimension identified in the corner and its associated metric placed in both the row and column.

Association Table	
Product	Visits
Visits	1.000

The Associations Table uses Cramer's V as a symmetrical correlation, resulting in selected metrics, dimensions, and elements values reflected in both the columns and rows of an Association Table. For example, selecting the **Product** extended dimension uses the **Visits** metric as the associated metric in both the row and column of the table, resulting in a perfect yet useless comparison (1.00) because the compared values are identical.

3. Add more values to the Association Table.

Right-click in a column or row and select **Add Metric** or **Add Dimension**. You can also drag metrics and dimensions from a **Finder** panel. Dimension elements can also be dragged and dropped from an open table to the table visualization.

Association Table			
Visitor	Visits	Week	Page Views
Visits	1.000	0.007	0.591
Week	0.007	1.000	0.014
Page Views	0.591	0.014	1.000



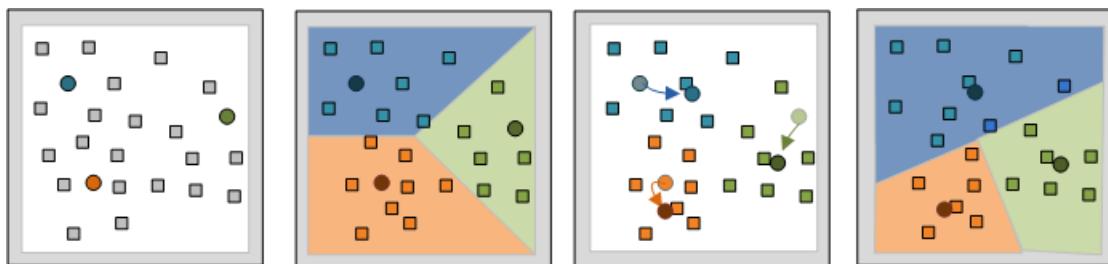
Note: There is a limit of ten rows and columns allowed in the Association Table.

Visitor Clustering

Visitor clustering lets you leverage customer characteristics to dynamically categorize visitors and generate cluster sets based on selected data inputs, thus identifying groups that have similar interests and behaviors for customer analysis and targeting.

Clustering process

The clustering process requires you to identify metrics and dimension elements to use as inputs, and allows you to choose a specific target population to apply these elements to create specified clusters. When you run the clustering process, the system uses the metric and dimension inputs to determine appropriate initial centers for the specified number of clusters. These centers are then used as a starting point to apply the K-Means algorithm.



The initial centers are intelligently chosen via a Canopy Clustering pass.

Data clusters are created by associating every data point to the nearest center.

The mean of each of the K clusters becomes the new center.

The algorithm is repeated in steps 2 and 3 until convergence is reached. This can take multiples passes.

The **Maximum Iterations** in the **Options** menu allows the analyst to specify the maximum number of iterations to be performed by the clustering algorithm. Setting this option may result in faster completion of the clustering process based on the maximum iterations cap at the expense of exact convergence of the cluster centers.



Note: Once the clusters have been defined, the Cluster Dimension can be saved for use just like any other dimension. It can also be loaded into the Cluster Explorer to examine the separation of cluster centers.

In the *Cluster Builder*, you can select **Options > Algorithm** to select algorithms when defining clusters. Currently, there are 3 supported algorithms:

- KMeans
- Kmeans++
- Expectation Maximization

There are 2 ways to run the clustering process:

- Method 1 - Click **Go** in the cluster visualization window.
- Method 2 - Click **Submit** in the cluster visualization window, which directly sends the clustering job to the Server. You can track the progress through the "Detailed Status for Query" option.

The screenshot shows a software interface for managing a Predictive Analytics Server. At the top, there are two tabs: "New Layout" and "Original Layout". Below them is a search bar labeled "Search:" followed by a lock icon. The main area displays a hierarchical tree view of server status under "Detailed Status for Query Server". A red box highlights the "Predictive Analytics Server" node, which has a sub-node "0 Tasks". Other nodes include "Component Status", "Disk Space", "Execution Engine", "Hardware", "Memory Status", "Network Connections", "Performance", and "Poll Latency". Below the tree, there is a list of server details:

- + Processing Status
- + Profiles 1 items
- Location E:\server\
- Product Insight Server
- Query Memory Usage 20,971,548/500,000,000 bytes
- Server Common Name [REDACTED]
- Service Status running
- Status Generated At 04/21/17 00:57:24
- Status Generation Took (s) 0.075087826699018478
- Version 6.61 (16101405)

The algorithm has the following restrictions:

1. If you are using Method 1, you can select any of the supported clustering algorithms.
2. If you are using Method 2, you can select kmeans or kmeans++. The Expectation Maximization option will not available.



Note: In the DPU.cfg file, the value for 'Query, Memory Limit' is set to 500 MB by default. This value must be increased while running multiple clustering jobs. For instance, if you are running 5 clustering jobs in parallel, increase this value to 1 GB. There is no way to cancel the clustering job without restarting the Server.

Recommendations

The number of iterations (number of times the data is scanned) and the convergence threshold that you configure, grossly affects the clustering performance. The following table provides a broader guideline that you can follow:

Number of Clusters	Algorithm	Iterations	Convergence Threshold	Normalization
6	Kmeans	25,50	1e-3	Min-Max

Number of Clusters	Algorithm	Iterations	Convergence Threshold	Normalization
6	Kmeans	25,50	1e-6	Min-Max
6	Kmeans++	50	1e-6	Min-Max

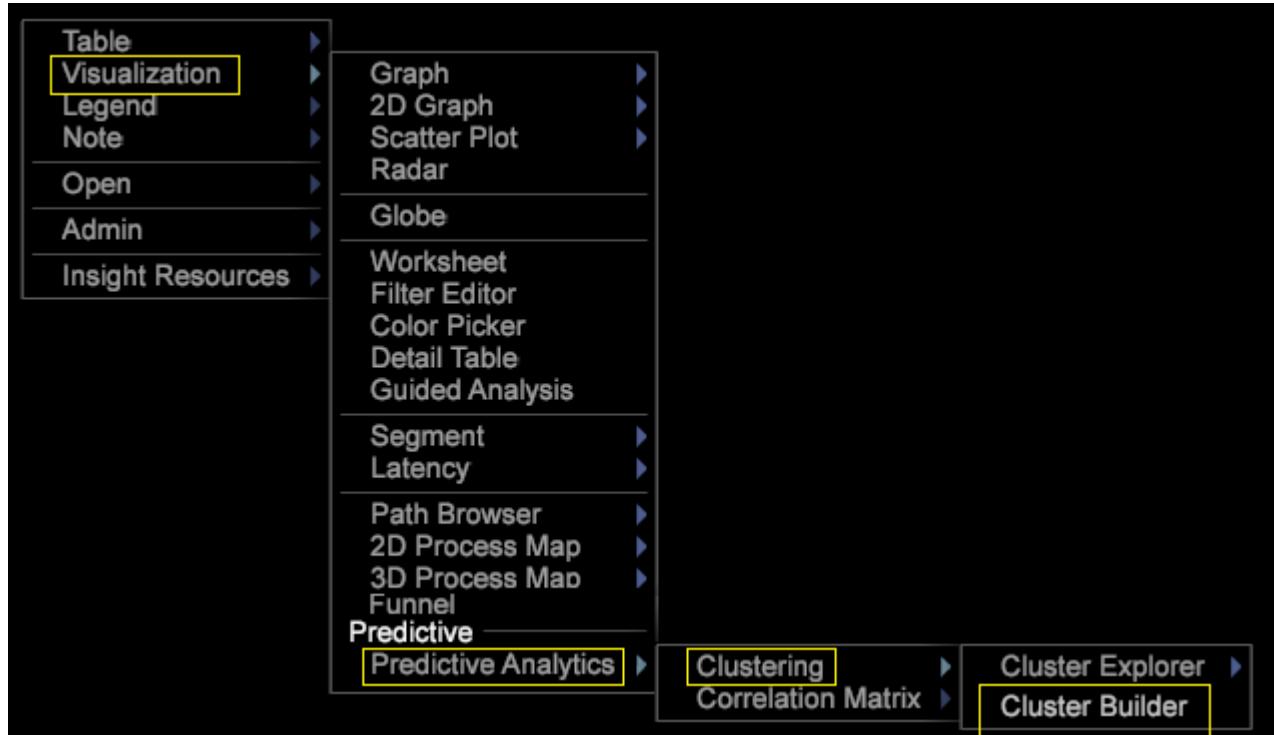
Building Clusters

Select input variables, the number of clusters, and a target population (if desired) to define clusters in your dataset.

Building Clusters

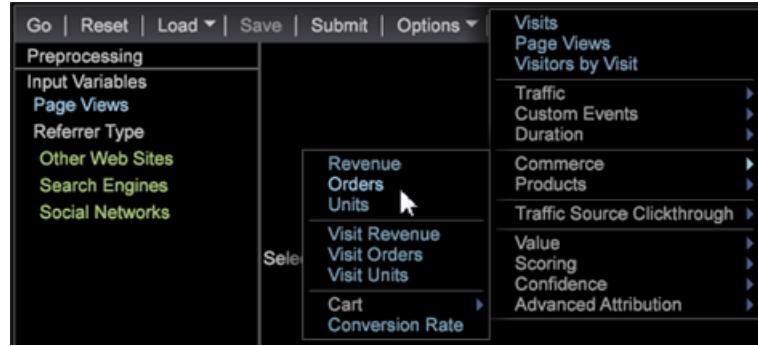
1. Open the Cluster Builder.

Click **Visualization > Predictive Analytics > Clustering > Cluster Builder**.



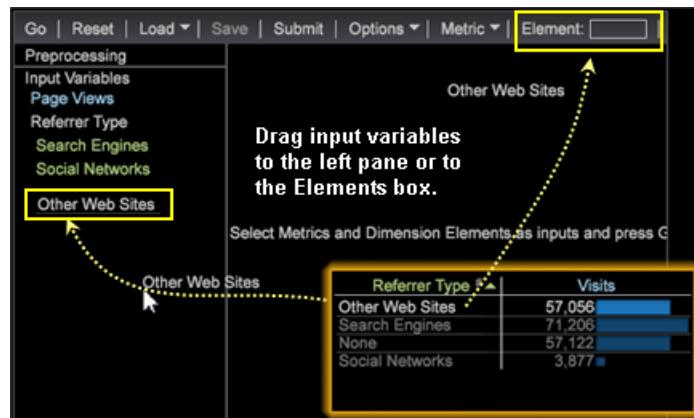
2. Select input variables.

- Add metrics to the **Input Variables** list by selecting from the **Metric** menu in the toolbar.



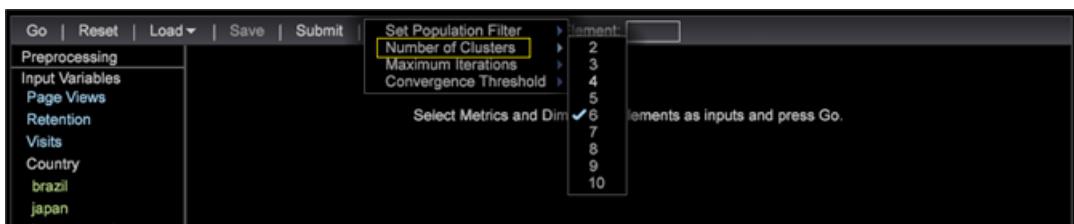
- Add dimension elements to the **Input Variables** list by dragging them from a Dimension's table.

Press **Ctrl + Alt** and drag selected dimension elements to the **Input Variables** list or to the **Element** box in the toolbar.



By default, clustering is performed on the entire dataset. You can see all input variables in the left **Preprocessing** pane.

3. Use the **Options** menu to select the desired number of clusters.



4. If you want to cluster a subset of the Visitors in your dataset, you can define a Population Filter.

The screenshot shows the SAS Data Studio Cluster Builder interface. The top menu bar includes 'Go', 'Reset', 'Load', 'Save', 'Submit', and 'Options'. The 'Options' menu is open, with 'Set Population Filter' selected. A sub-menu for 'Set Population Filter' is shown with options: 'Number of Clusters', 'Maximum Iterations', and 'Convergence Threshold'. A message below the menu says 'Select Metrics and Dimension Elements as inputs and press Go.' In the bottom right corner, a context menu is open over a bar chart showing 'Visitors for morePurchase.filter'. The context menu includes 'Apply Filter / Design Filter' and 'Include data with:' followed by a checkbox for 'Visitors with at least 1 Purchased Products'.

Start by defining the desired subset using selections in your Workspace or by using the **Filter Editor**. Once you have the desired subset selected, set the Target Population in the **Options** menu. It is recommended that you give the targeted group an identifying name.

The **Options** menu also has settings to control the maximum number of passes and the acceptable threshold for center convergence.

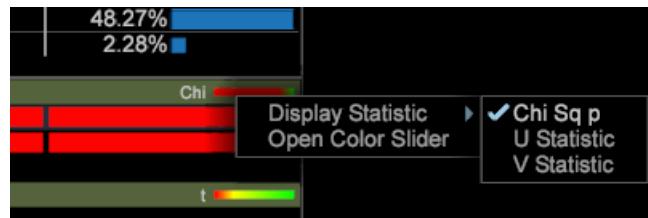
- After inputs and options have been configured, click the **Go** button to run the clustering locally or press **Submit** to send the task to the Predictive Analytics Server. Submissions to the server will save the resulting dimension to the dataset when convergence is complete.

When running locally, you will see the Cluster Builder move through four canopy clustering stages as it defines intelligent centers based on the inputs.

Once the centers of the clusters stop changing more than the specified convergence threshold, the Cluster Dimension is converged and the Cluster Builder displays additional information about how relevant an input was to each cluster.

- Customize the clusters.

Right-clicking on the statistics' color bar opens a context menu allowing you to customize the relevance thresholds, and in the case of the dimension element distributions, to choose which test is displayed.

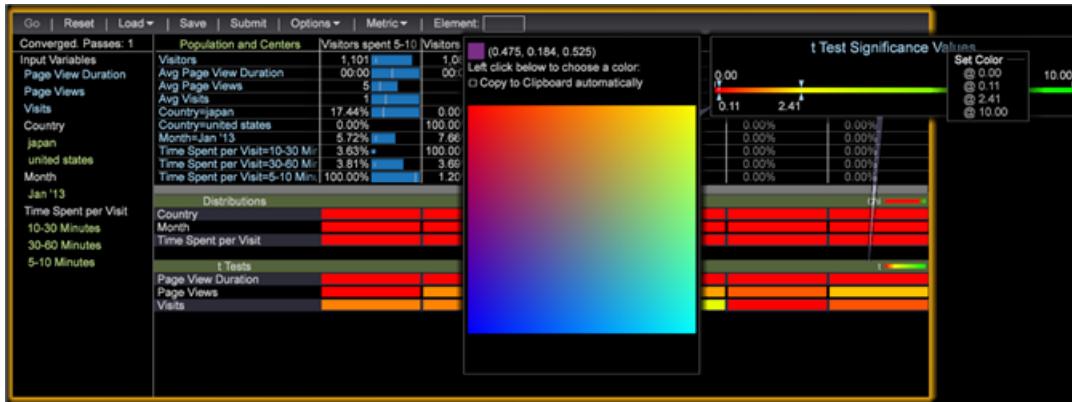


Metric inputs provide a t-test for each cluster, while dimension element inputs provide three distribution tests (Chi squared, an entropy U statistic, and Cramer's V statistic) for each cluster.



Note: If you add or remove inputs during convergence, the process will pause until you press **Go** again.

After building clusters, you can open the color picker to assign colors for different distribution results.



- With the Cluster Dimension converged, you can add metrics to the table and make selections as normal. You can also right click on the element names (Cluster 1, Cluster 2, etc) to open the context menu to rename them to something more meaningful.



- If you wish to use this cluster dimension in other visualizations, you can **Save** it locally or **Submit** it to the server.

If you wish to run convergence again or see the relevance of inputs, Cluster Builder can also load existing cluster dimensions.

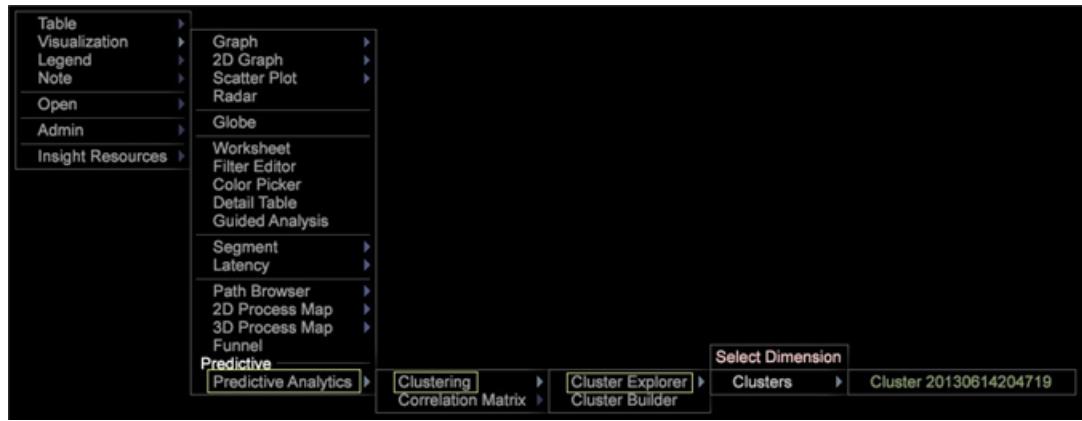


Tip: When selected, **Reset** will completely release all the input variables and give you a blank cluster builder visualization to define new clusters.

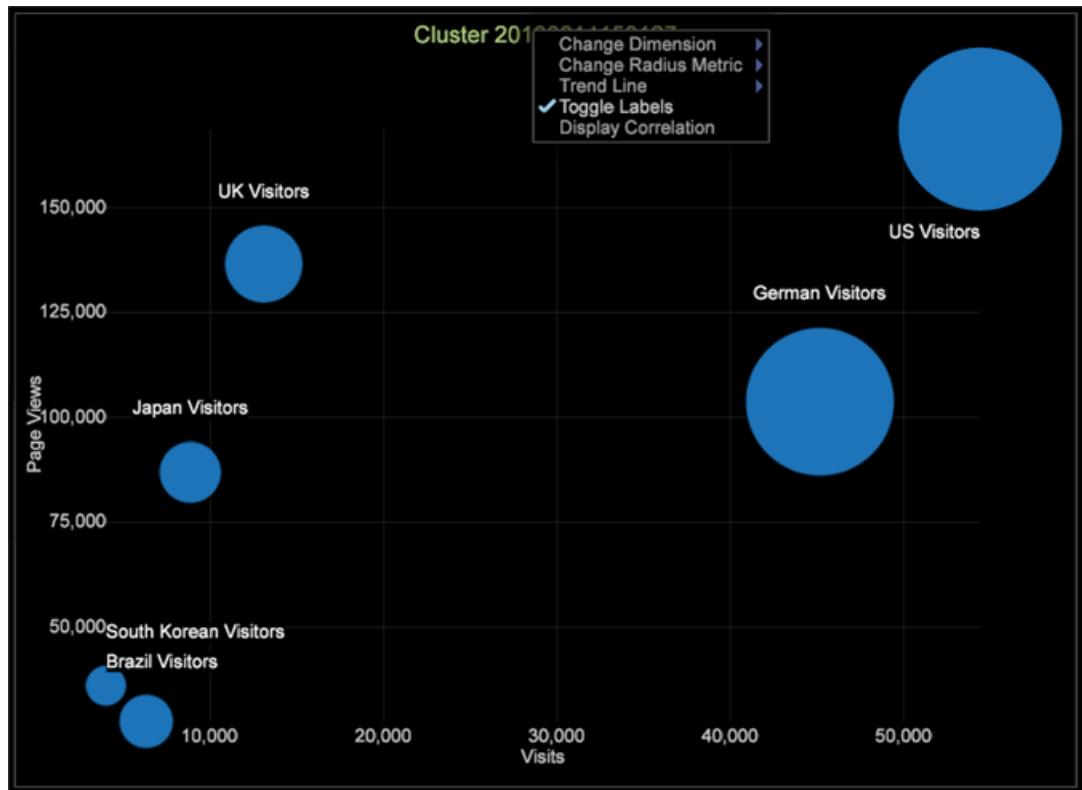
Exploring Clusters

After a Cluster Dimension is built, you can save it and examine it using the Cluster Explorer.

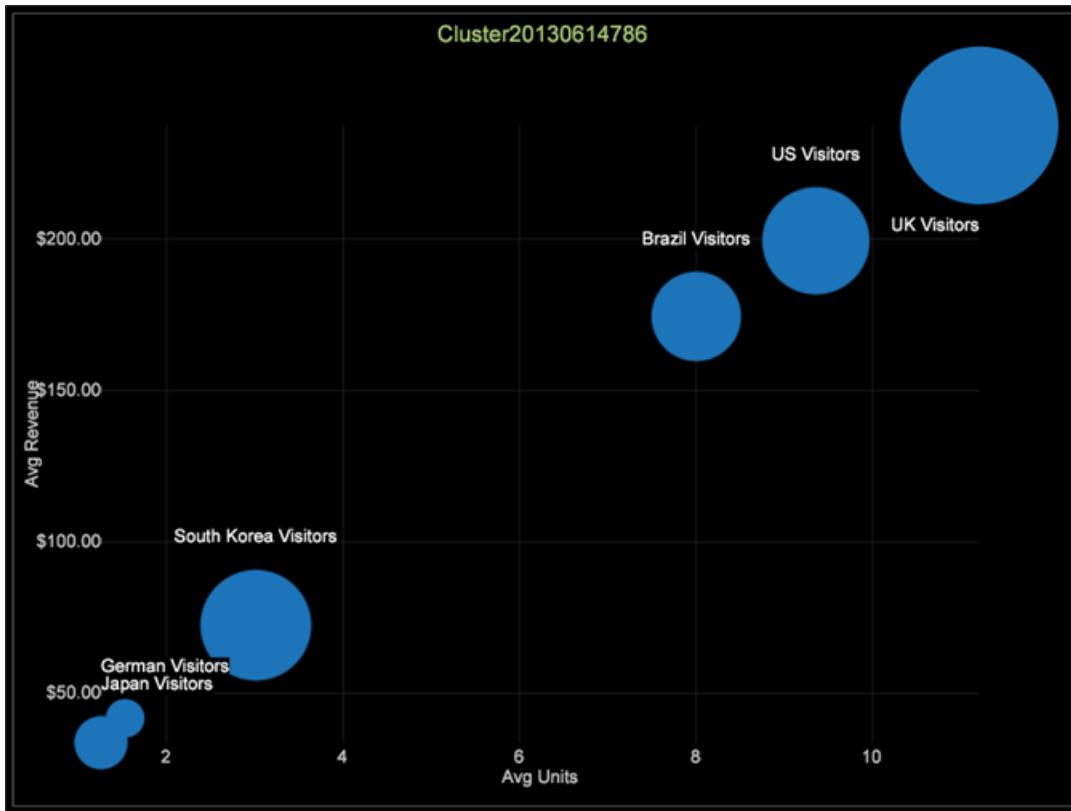
- Select **Visualization > Predictive Analytics > Clustering > Cluster Explorer > Clusters**. Choose the saved cluster dimension that you want to explore.



- Right-click the visualization title and click Toggle Labels to turn them on. You can use this menu to change the cluster dimension you are exploring, change the radius metric, add trend lines, and display correlation.

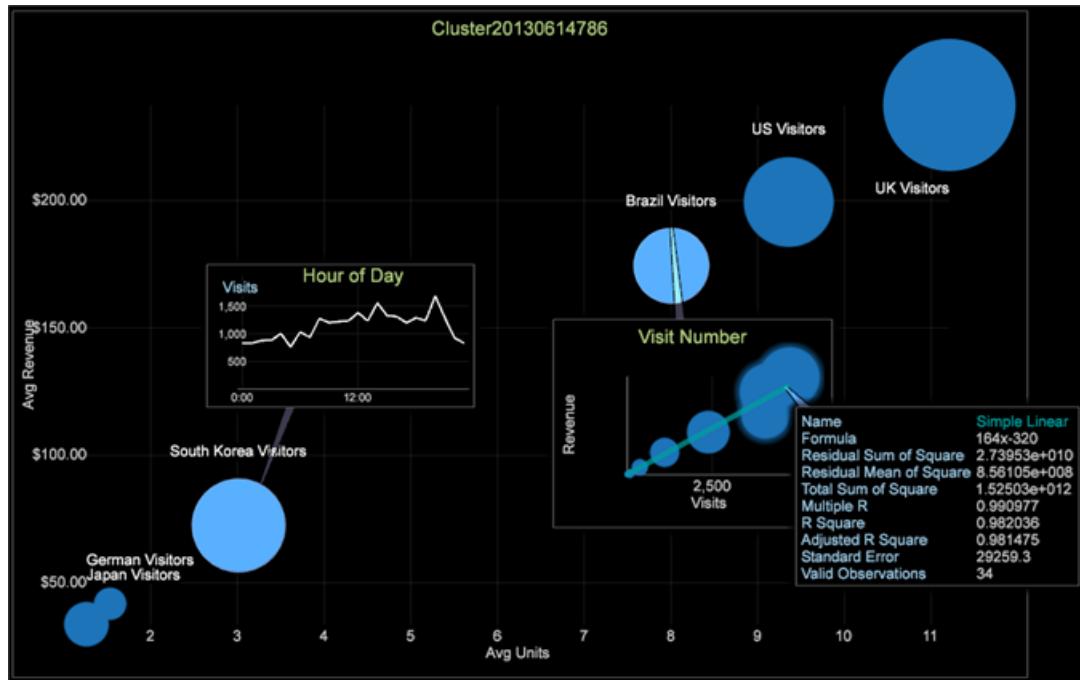


- To modify the Cluster Dimension you are exploring, change the radius metric, add trend lines, and display the correlation using this menu.



The Cluster Explorer displays the centers of the clusters along any two input axes at a time. This allows you to examine their separation in multidimensional space.

4. Right click the cluster to display the menu and choose from callout types: Image or Text Annotation, Metric Legend, Table, Line Graph or Scatter Plot.



Chord Visualization

The Chord visualization allows you to show both the proportion and correlation between metrics, displaying larger chords as an indication of a stronger correlation.

The Chord visualization lets you see identify correlations between metrics, allowing you to add and easily evaluate possible correlations. It also provides another view into any previously built [Correlation Matrix](#). Using the Chord visualization, you cannot identify a positive or negative correlation between the metrics—only that a correlation exists. In certain cases, determining a direct or inverse relationship can be identified by applying counter metrics.

1. **Open the Chord visualization.**

In the workspace, right-click **Visualization > Predictive Analytics > Chord**.

2. **Select a Dimension from the menu.**

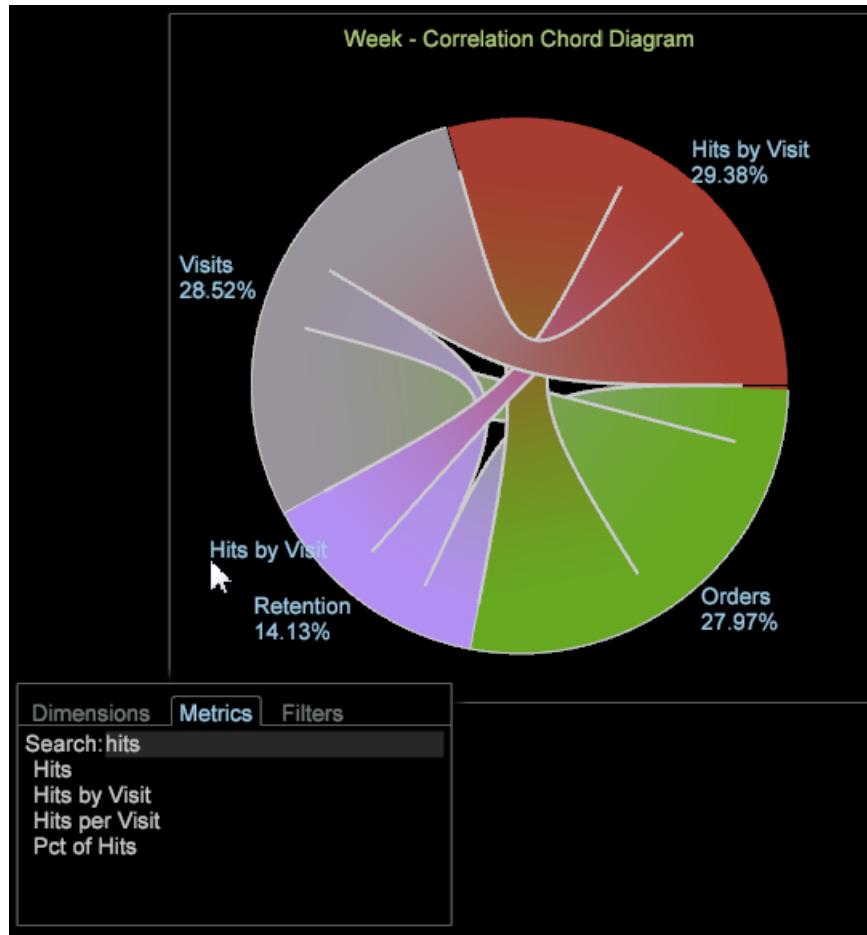
A blank visualization will open allowing you to select a dimension. The dimension name will appear at the top of the blank chord visualization.



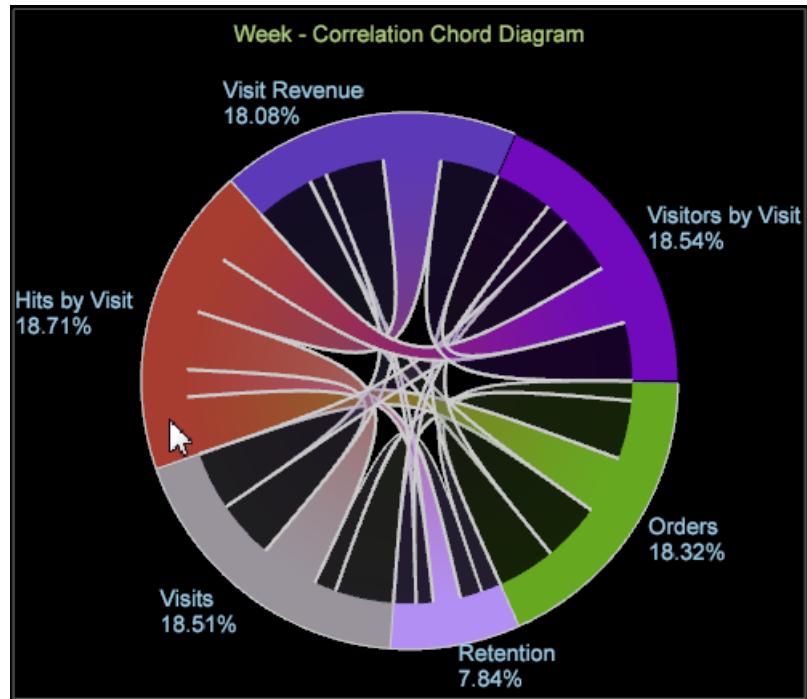
Note: If you already have a Correlation Matrix open in the workspace, you can also render it as a Chord visualization.

3. **Choose metrics to correlate.**

Drag metrics from the **Finder** by clicking **Ctrl-Alt** to drag metrics from the table to the chart. After two or more metrics are selected, the chart will automatically refresh and begin displaying correlation data. Continue adding metrics as needed to correlate data points.

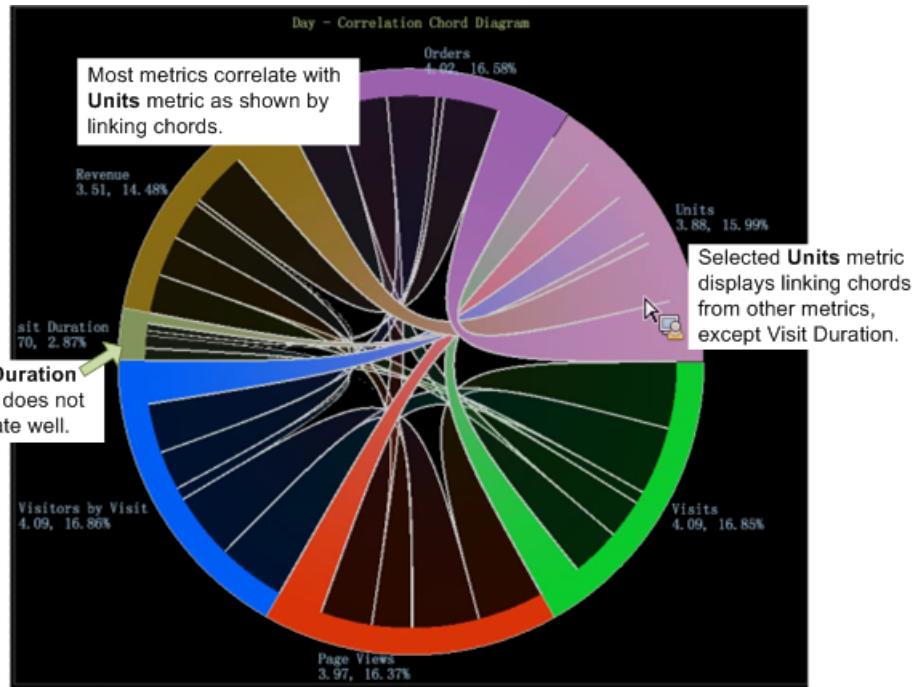


The Chord visualization displays the proportion of the whole represented by the area of each segment. Continue to add metrics as need to identify and investigate significant relationships.

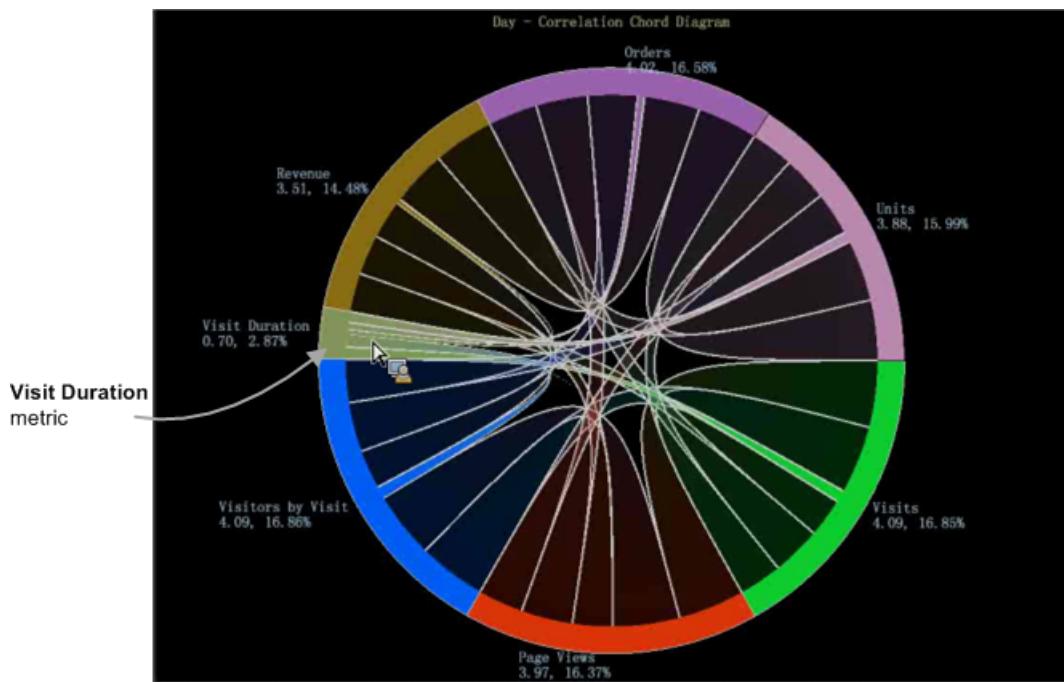


4. View the Chord visualization.

Hover over each metric in the visualization to see relationships. In the example, you can see a correlation between Units and most other metrics (except for the **Visit Duration** metric).

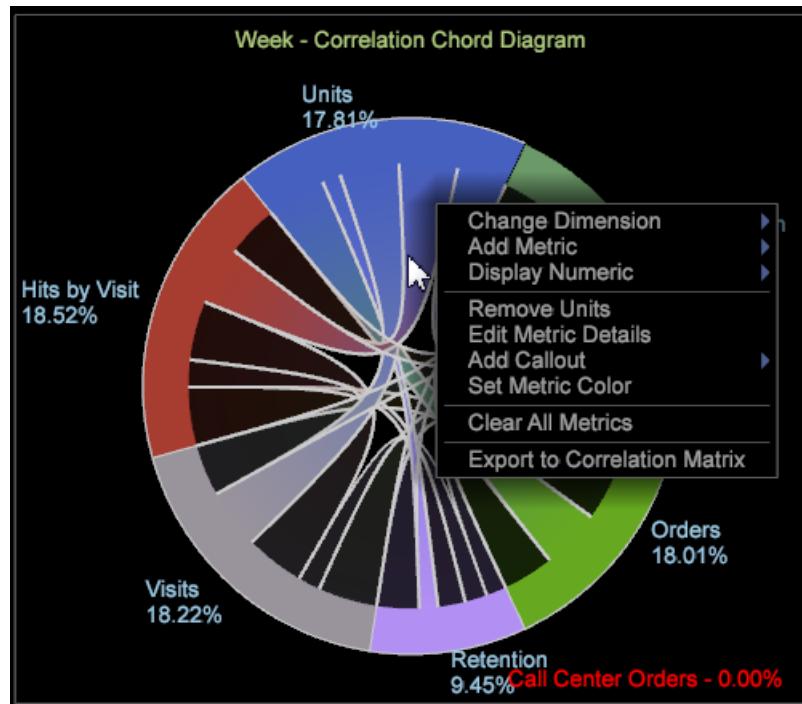


When you hover over the **Visit Duration** metric on the Chord visualization, you can see there is very little or at most weak correlation between all other metrics.



5. Change Settings.

Right-click the Chord visualization to open a menu to change the dimension, display the dimensions as absolute numbers or as percentages, remove the selected metric or all metrics, edit colors and details, and export values to a Correlation Matrix.



Correlation Matrix

Statistical correlations measure meaningful relationships to identify opportunities through advanced data mining.

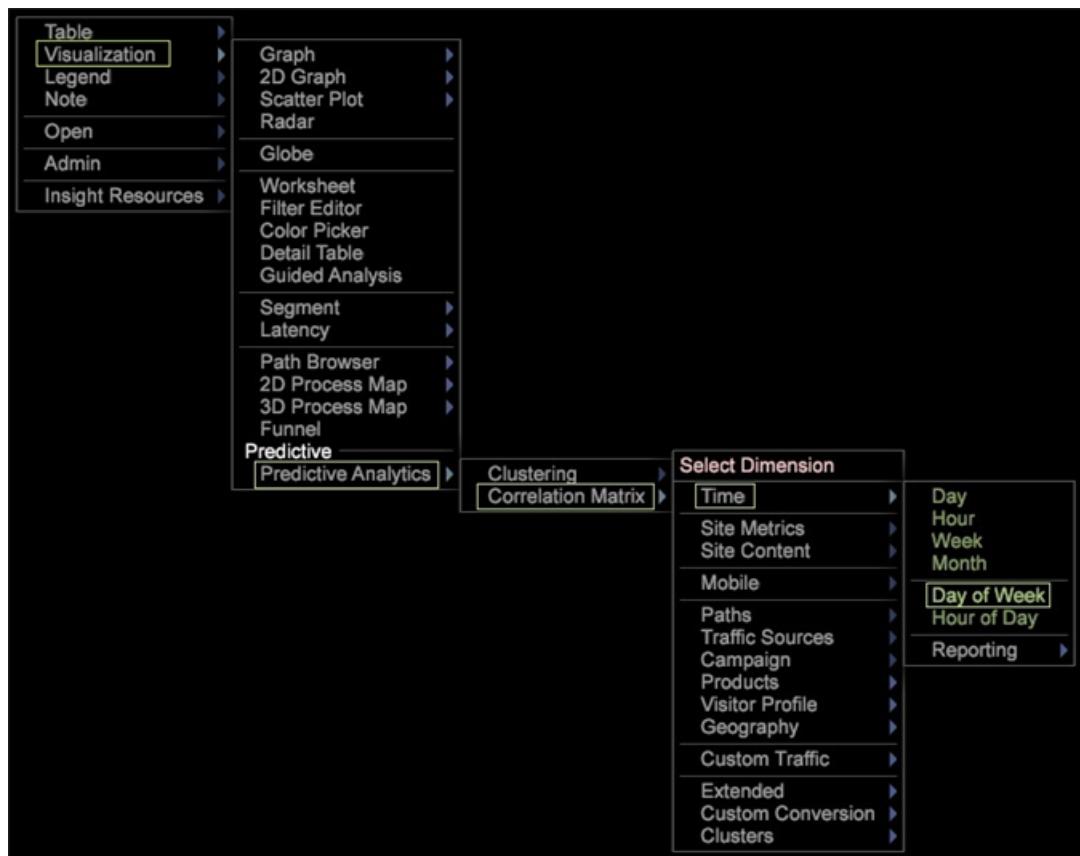
Employing the [About the Pearson Correlation Coefficient](#) on page 168, the Correlation Matrix furnishes you with relevant information to better identify the next steps in a marketing campaign, to improve site design, or to continue in-depth customer analysis for additional correlation dependencies.

Building a Correlation Matrix

The Correlation Matrix compares metrics over a countable or non-countable dimension. The matrix can then be modified to highlight correlations within the visualization through color picking or to render it as a text map, heat map, or both.

1. Open a Correlation Matrix.

Right-click **Visualization > Predictive Analytics > Correlation Matrix**. The dimension table will open.



Select a dimension, such as Time > Day of the Week from this menu. The correlation table will open with the dimension identified in the corner of the matrix and its associated metric placed in the row and column. For the Day of the Week dimension, **Visits** is the associated metric.

Day of Week	Visits
Mon	29,273
Tue	22,477
Wed	23,200
Thu	30,602
Fri	29,066
Sat	24,400
Sun	24,837

The correlation is 1.000 because you are comparing a metric against itself (which reflects a perfect, but unusable, correlation.)

2. Change one of the metrics.

Right-click and select **Change Metric** to change a metric in either the row or column. This sets up a correlation between two metrics of value.

For this example, change the **Visits** metric in the column to **Internal Searches**. Right-click and select **Metric** > **Custom Events** > **Custom Event 1-10** > **Internal Searches**.

Correlation Matrix	
Day of Week	Internal Searches
Visits	0.828

3. Add more metrics to the Correlation Matrix.

Right-click in a metric column or row. For example, from the Metric menu, add

Metric > **Custom Events** > **Custom Event 1-10** > **Sign in Error**.

Correlation Matrix	
Day of Week	Visits
Visits	1.000
 <ul style="list-style-type: none"> Visits Metric Editor Metrics Alt Metric Change Metric Remove Metric Mirror to Rows Recalculate Edit Colors 	

The new metric will appear in a column with a correlation number. You can add other metrics, such as **Email Signups**, to build out the table.

Correlation Matrix		
Day of Week	Visits	Signin Errors
Visits	1.000	0.681

Or add metrics to rows to compare against metrics in columns.

Correlation Matrix			
Day of Week	Internal Searches	Signin Errors	Email Signups
Visits	0.795	0.815	0.769
Visitors by Visit	0.865	0.889	0.869

Visitors by Visit

- Metric Editor
- Metric Legend
- Edit Metric Details

Metrics

- Add Metric
- Change Metric
- Label
- Remove Metric
- Mirror to Columns
- Recalculate
- Edit Colors

Visits

- Page Views
- Visitors by Visit

Traffic

- Custom Events
- Duration

Commerce

- Products

Traffic Source Clickthrough

Scoring

- Confidence

4. (optional) Constrain a metric by adding a dimension element.

Right-click in the workspace and select **Table**. From the open dimension table, press Ctrl + Alt and drag the element over a metric in a column or row. The element will appear next to the metric in brackets.

For example, for the **Visits** metric, you can constrain it by selecting the **Country** as **New Zealand**.

Correlation Matrix			
Day of Week	Internal Searches	Signin Errors	Email Signups
Visits[Country="new zealand"]	0.599	0.542	0.283
Visitors by Visit	0.606	0.587	0.334

Country

Country	#	Visits
guatemala		32.
new zealand		330.
ecuador		191.
croatia (local name - hrvats		81.
thailand		363.
finland		138.
p puerto rico		182.
i indonesia		124.
philippines		150.

Notice that when you select a dimension element, the correlation changes in all metrics based on the selected dimension element. Only the Visit metric will be constrained for "New Zealand" once the dimension window is closed.



Note: If changing a metric with a dimension constraint (by right-clicking and selecting **Change Metric**), the dimension element constraining the metric will be lost. You will need to add the dimension element again.

-
5. Create a *Binary Filter in the Correlation Matrix* on page 169 to further constrain the metric. Right-click the metric in the table and select Binary Filter from the menu.

Correlation Planning and Analysis Goals

The following are general goals for building a Correlation Matrix.

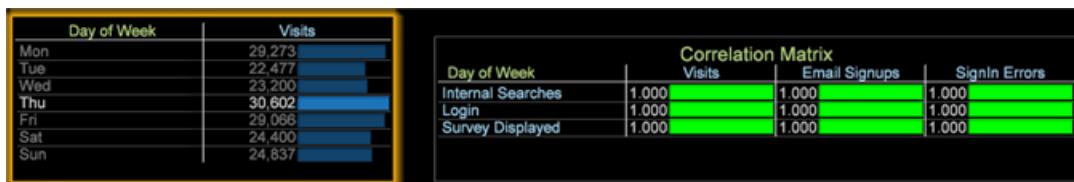
Identify the relationship between two metrics against a specified dimension. In the example, the matrix was built around the core dimension, Day of the Week, with the metrics Visit, Email Signups, and SignIn Errors compared against Internal Searches, Login, and Survey Displayed metric events.

Develop hypotheses to focus analysis. After running a correlation analysis, your next step is to look for dependencies and correlation of the metrics. For example, understanding that internal searches has an effect on email sign-ups provides a path to predict that relationship and to modify marketing campaigns or web site navigation design.

Identify metrics to include more advanced data mining algorithms. In most cases, the key metrics will be identified because they will be seen affecting multiple correlations. You can now take those key metrics and apply them to additional data mining analysis for deeper insight.

Correlation Matrix Feature Notes

Filtering and selecting on dimension elements within a table compares like values. For example, using Day of the Week dimension and then clicking into an element of that core dimension, such as clicking on a specific day within the Day of Week dimension table, renders a one to one match at 100% that provides no usable correlation. Because the root dimension was Day of the Week, any selection within the Day of the Week dimension table will alter the matrix to be a one-to-one correlation.



Day of Week	Visits
Mon	29,273
Tue	22,477
Wed	23,200
Thu	30,602
Fri	29,066
Sat	24,400
Sun	24,837

Day of Week	Correlation Matrix		
	Visits	Email Signups	SignIn Errors
Internal Searches	1.000	1.000	1.000
Login	1.000	1.000	1.000
Survey Displayed	1.000	1.000	1.000

However, the 1 to 1 correlation (when a single selection is made of all elements) is only on that specific day. If you make multiple selections then it does not necessarily remain a 1 to 1 correlation, and will not always yield a 100 percent match regardless of selecting 1 or 1+ days of the week.

Statistical correlations are not equal to the Correlated Data Model, the historical reference of Adobe Analytics products. The statistical correlation in data workbench is based on the *About the Pearson Correlation Coefficient* on page 168.

Display Correlation in a Scatter Plot. Right click the title on a Scatter Plot and choose **Display Correlation** from the **Visualization** menu. The Correlation value will display in the upper right section of the Scatter Plot.



Note: The Scatter Plot and Pearsons matrix will display "Calculation Error" if the application is unable to run the Pearson's correlation calculation. This is usually due to insufficient data, which can cause the equation to attempt to divide by 0.

About the Pearson Correlation Coefficient

The Pearson Correlation Coefficient is used in the Correlation Matrix as the algorithm to display the strength of the linear dependence between two variables.

This linear correlation is a statistical measure of the linear dependence, or correlation, between two variables to render a value between +1 and -1 inclusive, representing either a positive or negative dependence.

Here is the Pearson Correlation Coefficient

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

The Pearson's value is visualized in the Correlation Matrix, which depicts the correlation between two defined metrics. These metrics can be compared to one another over any countable or non-countable dimension in the dataset.

You can highlight these comparisons through contrasting colors using the color picker, or by comparing values in a text map and heat map, or both.

Correlation Analysis Use Cases

Defining dependent correlation points that make sense in your market is the essence of correlation analysis.

These use cases highlight the art of identifying relationships as correlation points applied to the science of the [About the Pearson Correlation Coefficient](#) on page 168.

Social Media and Website Correlation

Digital publishers want to maximize their understanding of the potential relationship between social media activity and visits to their website. For example, the digital publisher runs the correlation report between hourly Twitter mentions and visits for a two week period. The correlation is found to be $r = 0.28$, which indicates a medium, positive relationship between Twitter mentions and website visits.

Optimization for E-retailers

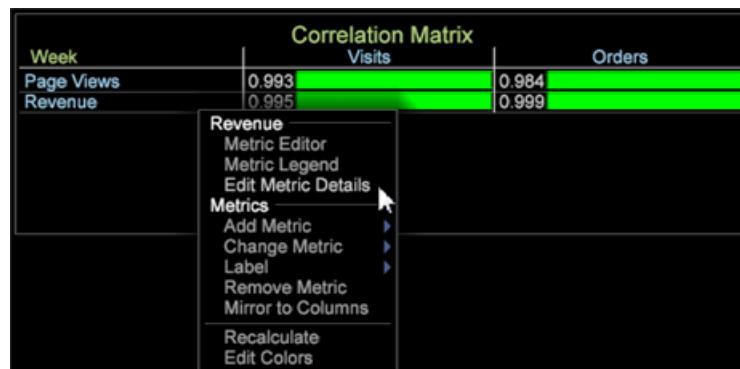
E-retailers are interested in driving increased revenue. For example, an e-retailer wants to compare a number of secondary success events (e.g., file downloads, product detail page views, internal search click-throughs, etc.) with weekly web revenue. They can quickly identify internal search click-throughs as having the highest correlation ($r = 0.46$), which may indicate an area for optimization.

Binary Filter in the Correlation Matrix

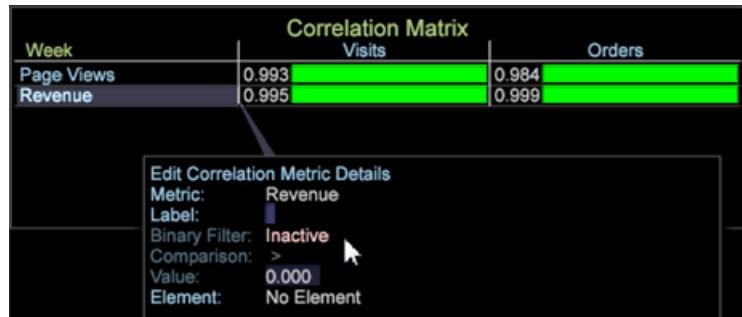
A **Binary Filter** in the Correlation Matrix lets you constrain values for one or both of the correlated metrics to better focus the comparison.

To set a binary filter on a Correlation Matrix:

1. From the Correlation Matrix, right-click a metric name.
2. Select **Edit Metric Details**.



The **Edit Correlation Metric Details** window will open.



A screenshot showing the Correlation Matrix and the **Edit Correlation Metric Details** window. The Correlation Matrix shows correlations between Week, Page Views, and Revenue. The matrix values are: Week vs Page Views (0.993), Week vs Revenue (0.995), Page Views vs Visits (0.984), and Page Views vs Orders (0.999). The **Edit Correlation Metric Details** window displays the following settings:

Metric:	Revenue
Label:	
Binary Filter:	Inactive
Comparison:	>
Value:	0.000
Element:	No Element

3. Set up a Binary Filter.

First, click the **Inactive** setting. It will toggle to set the filter as **Active** and display the **Comparison** and **Value** fields.

Then, select a **Comparison** operator and set its **Value** to set up a filter for the selected metric.



Important: The Binary Filter for Data Workbench 6.2 has been updated with new features, requiring you to rebuild any correlation matrix with a binary filter built in previous versions.

Adding Dimension Elements

You can also add a dimension element to constrain a metric. A metric can have only one element associated with it.



A screenshot showing the Correlation Matrix and the **Edit Correlation Metric Details** window. The Correlation Matrix now includes a dimension element for Revenue: Revenue[Week="01/21/13"]. The **Edit Correlation Metric Details** window shows the following settings:

Metric:	Revenue[Week="01/21/13"]
Label:	
Binary Filter:	Inactive
Comparison:	>
Value:	0.000
Element:	Week="01/21/13"

Below the matrix, a table shows the data for the dimension element:

Week	Visits
01/21/13	13,782
01/28/13	40,825
02/04/13	42,010
02/11/13	35,254

Right-click in the workspace and select **Table**. Open a dimension with its elements and drag to the **Element** setting in the **Edit Correlation Metric Details** window, or drop on a metric in the Correlation Matrix.

Decision Tree Builder

Decision trees are a predictive analytics visualization used to evaluate visitor characteristics and relationships. The Decision Tree Builder generates a decision tree visualization based on a specified positive case and a set of inputs.

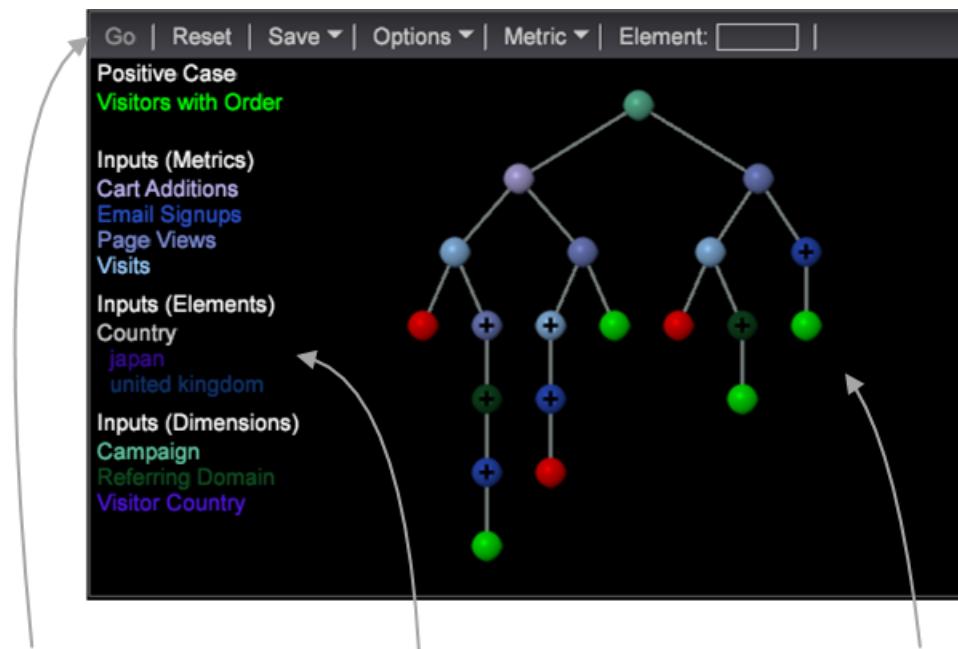
A Decision Tree is a binary classifier with a set of rules (or filters) identifying visitors who satisfy specific rules based on a positive case. A decision tree sets rules to classify visitors who satisfy (or do not satisfy) this positive case. These rules generate a tree map to provide a level of confidence to meet these positive case results.

A Decision Tree is built by examining inputs at each level and choosing the one that provides a maximum gain of information at a specified split point. Split points for each variable-level generates two sets:

- Values less than or equal to the split point, and
- Values greater than the split point.

Use decision trees to

- Perform meaningful analysis and interpretation in less time.
- Employ automated segment generation.
- Quickly make inferences from a model based on a large amount of data.



Toolbar and Menus	Input Listing	Tree Display
<p>The toolbar includes buttons and menu commands for the Decision Tree, including features to set the Positive Case and add Input Listings.</p> <p>Like other visualizations, the Element box lets you drag and drop Dimension and Elements, although you can also drag directly from the Finders pane.</p> <p>For additional information, see Decision Tree Options on page 90.</p>	<p>This area displays the inputs into the tree model. They are color coded to match nodes in the Tree Display area.</p> <p>Right-clicking on an input allows you to remove the input from the model and reset. If you hover over a tree node, it will display the split conditions along the branch to that node and the prediction at that node with its confidence value.</p>	<p>This area displays the tree model with leaf nodes color-coded based on its prediction: green for a True prediction of the Positive Case, and red for a False prediction.</p> <p>The split nodes are color coded to the inputs that match their selection condition. Hovering over a node displays information about the split and expands the inputs listing to display the split points along the branch and the distribution of the training set.</p> <p>Nodes below a threshold are not displayed by default. Click on an expandable node (indicated</p>

by a + symbol) to explore a branch. Click on the root node to return to the full tree display.

Building a Decision Tree

Set up a Decision Tree by identifying a positive case and adding metric and dimension inputs to evaluate the data and explore the decision tree.

Follow these steps to build a decision tree.

1. Open a new workspace.

After opening a new workspace, you might need to click **Add > Temporarily Unlock**.

2. To open the Decision Tree Builder, right-click **Visualization > Predictive Analytics > Classification > Decision Tree Builder**.
3. Set a **Positive Case**.

You can define a positive case for a decision tree by selecting dimensions in a Finder or dimension elements in a table, or by designing a filter in the Design Filter. In fact, the positive case can be a combination of multiple selections in the workspace including filters, dimensions, elements, and all types of Data Workbench visualization values.

- **Design and Apply a Filter** as a positive case. Right-click in the workspace and select **Tools > Filter Editor** to design and apply a filter.
- Add **Dimensions** as a positive case. In the workspace, right-click and select **Tools > Finders** (or select **Add > Finders** in the left pane). Type a dimension name in the **Search** field and then select a dimension.
- Add **Metrics** as a positive case. Right-click and select **Tools > Finders** or select **Add > Finders** in the left pane to open a Metrics table. Select a metric as your positive case.
- Add **Dimension Elements** as a positive case. Right-click in the workspace and select **Table** to open dimension elements, then select from the dimension elements to set your positive case.

4. Click **Options > Set Positive Case**.

This sets the positive case and lets you name it. The name will appear under the **Positive Case** heading in the workspace.



Note: When you set the positive case the Decision Tree uses the current workspace selection, which can be defined as the Visitors (or whatever top-level countable is defined, but in most cases Visitors) that match the current selection within the workspace. These combine as a single filter for a single positive case (not multiple positive cases).

Clicking **Set Positive Case** when there is no selection will clear the positive case.

5. (optional) Select **Set Population Filters** to define the visitor population to be classified.

If no population filter is applied, then the training set is drawn from all visitors (default is "Everyone").



Note: Click the **Show Complex Filter Description** to view the filtering scripts for the Positive Case and Population Filter.

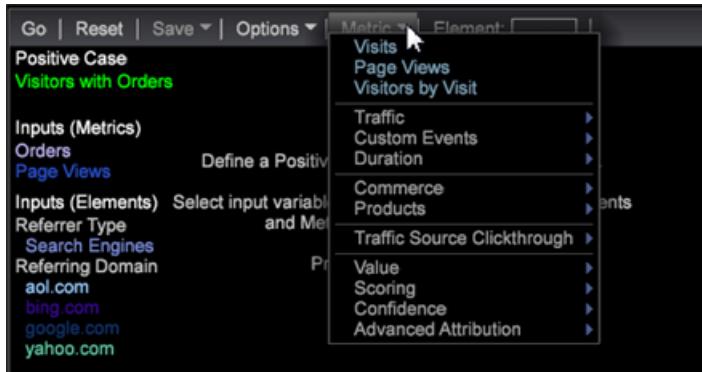
6. Add **Metrics, Dimensions, and Dimension Elements** as inputs.

You can select inputs by dragging and dropping from the Finder panels or from tables for individual dimension elements. You can also select from the **Metrics** menu in the toolbar.

- Add **Metrics** as inputs.

Select Metrics from the toolbar. Press **Ctrl + Alt** to drag one or more metrics to the Decision Tree Builder.

The metric will appear in the **Input (Metrics)** list as an input with unique color-coding.



- Add **Dimensions** as inputs.

In the workspace, right-click and select **Tools > Finder** and type the dimension name in the **Search** field. Press **Ctrl + Alt**, select a dimension, and drag the dimension to the Decision Tree Builder.

The dimension will appear in the **Input (Dimensions)** list with a unique color-coding.

- Add **Dimension Elements** as inputs.

In the workspace, right-click and select a Dimension table. Select Dimension Elements, press **Ctrl + Alt**, and drag the selected elements to the Decision Tree Builder.

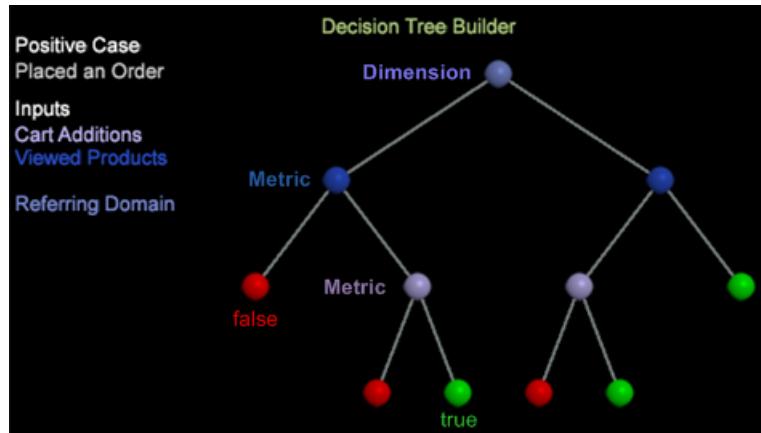
The dimension elements will appear in the **Input (Elements)** list with a unique color-coding.



Important: You can select up to a maximum of fourteen inputs to be evaluated. An error message will appear if too many inputs are added.

7. Select **Go** from the toolbar.

The decision tree will build based on the selected dimensions and metrics. Simple metrics such as Cart Additions will build quickly, while complex dimension such as Visit Duration with multiple data points will build more slowly with a percentage of the completion displayed as it converts. The tree map will then prune and open for user interaction. The dimension and metric inputs will be color-coded consistent with the node names.



The leaf node displays as green (true) or red (false) if the tree has been pruned and if there is a prediction of **True** or **False** following the pruned branches.

Attention: The training sample is pulled from the dataset for the tree builder to use. Data Workbench uses 80 percent of the sample to build the tree and the remaining 20 percent to assess the accuracy of the tree model.

8. Verify accuracy using the **Confusion Matrix**.

Click **Options > Confusion Matrix** to view the Accuracy, Recall, Precision and F-Score values. The closer to 100 percent, the better the score.

The Confusion Matrix gives four counts of accuracy of the model using a combination of values:

- Actual Positive (AP)
- Predicted Positive (PP)
- Actual Negative (AN)
- Predicted Negative (PN)



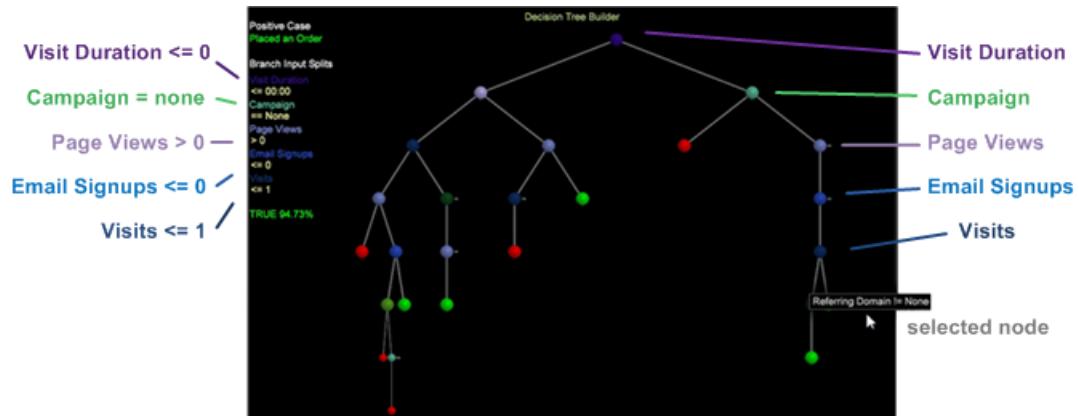
Tip: These numbers are obtained by applying the resulted scoring model of the 20 percent testing data withheld and already known as the true answer. If the score is greater than 50 percent, it is predicted as a positive case (that matches the defined filter). Then, Accuracy = $(TP + TN) / (TP + FP + TN + FN)$, Recall = $TP / (TP + FN)$, and Precision = $TP / (TP + FP)$.

9. Explore the decision tree.

After generating a decision tree, you can view the path of the prediction and identify all visitors who meet the defined criteria. The tree identifies the input split for each branch based on its position and color-coding. For example, if you select the Referring Domain node, the nodes leading to that split are listed by color-code to the left of the tree.

You can make selections of the leaf nodes to select branches (rule sets) of the decision tree.

For this example: If the visit duration is less than 1, no campaign exists, at least one page view exist, no email signups, and there was at least one visit. The projections on this meeting criteria and placing an order is **94.73** percent.



Decision Tree interaction: You can select multiple nodes on the tree using the standard **Ctrl-click** to add, or **Shift-click** to delete.

Color-coded nodes: The color of the nodes matches the color of the input dimensions and metrics as assigned by Data Workbench.

Bright green and red nodes at the leaf-level of a pruned branch predicts the node as True or False.

● Bright green	Identifies that the node equals true and that all conditions are met.
● Bright red	Identifies that the node equals false and not all conditions are met.

10. Save the Decision Tree.

You can save the Decision Tree in different formats:



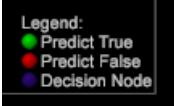
- Predictive Markup Language (**PMML**), an XML-based file format used by applications to describe and exchange decision tree models.
- **Text** displaying simple columns and rows of true or false, percentages, number of members, and input values.
- A **Dimension** with branches corresponding to predicted outcome elements.

Decision Tree Options

The Decision Tree menu includes features to set the positive use case, filters, leaf distribution options, confusion matrix, and other advanced options.

Toolbar buttons	Description
Go	Click to run the decision tree algorithm and display the visualization. This is grayed-out until there are inputs.
Reset	Clears inputs and decision tree model and resets the process.
Save	Save the Decision Tree. You can save the Decision Tree in different formats: <ul style="list-style-type: none"> • Predictive Markup Language (PMML), an XML-based file format used by applications to describe and exchange decision tree models.

Toolbar buttons	Description
	<ul style="list-style-type: none"> • Text displaying simple columns and rows of true or false, percentages, number of members, and input values. • A Dimension with branches corresponding to predicted outcome elements.
Options	See table below for Options menu.

Options menu	Description
Set Positive Case	Defines the current workspace selection as the model's Positive Case. Clears the case if no selection exists.
Set Population Filter	Defines the current workspace selection as the model's Population Filter and will be drawn from visitors who satisfy this condition. The default is "Everyone."
Show Complex Filter Description	Displays descriptions of the defined filters. Click to view the filtering scripts for the Positive Case and Population Filter.
Hide Nodes	Hides nodes with only a small percentage of the population. This menu command displays only when the decision tree is displayed.
Confusion Matrix	<p>Click Options > Confusion Matrix to view the Accuracy, Recall, Precision and F-Score values. The closer to 100 percent, the better the score.</p> <p>The Confusion Matrix gives four counts of accuracy of the model using a combination of values:</p> <ul style="list-style-type: none"> • Actual Positive (AP) • Predicted Positive (PP) • Actual Negative (AN) • Predicted Negative (PN) <p> Tip: These numbers are obtained by applying the resulted scoring model of the 20 percent testing data withheld and already known as the true answer. If the score is greater than 50 percent, it is predicted as a positive case (that matches the defined filter). Then, Accuracy = $(TP + TN) / (TP + FP + TN + FN)$, Recall = $TP / (TP + FN)$, and Precision = $TP / (TP + FP)$.</p>
Display Legend	<p>Allows you to toggle a legend key on and off in the Decision Tree.</p>  <p>This menu command displays only when the decision tree is displayed.</p>
Advanced	Click to open Advanced menu for in-depth use of Decision Tree. See table below for menu options.

Advanced menu	Description
Training Set Size	Controls the size of the training set used for the model building. Larger sets take longer to train, smaller sets take less time.
Input Normalization	Allows the user to specify whether to use the Min-Max or the Z Score technique to normalize inputs into the model.
SMOTE Over-Sampling Factor	When the Positive Case does not occur very often (less than 10 percent) in the training sample, SMOTE is used to provide additional samples. This option allows the user to indicate how many more samples to create using SMOTE.
Leaf Class Distribution Threshold	Allows you to set the threshold assumed for a leaf during the tree building process. By default, all members of a node must be identical for it to be a leaf (prior to pruning stage).

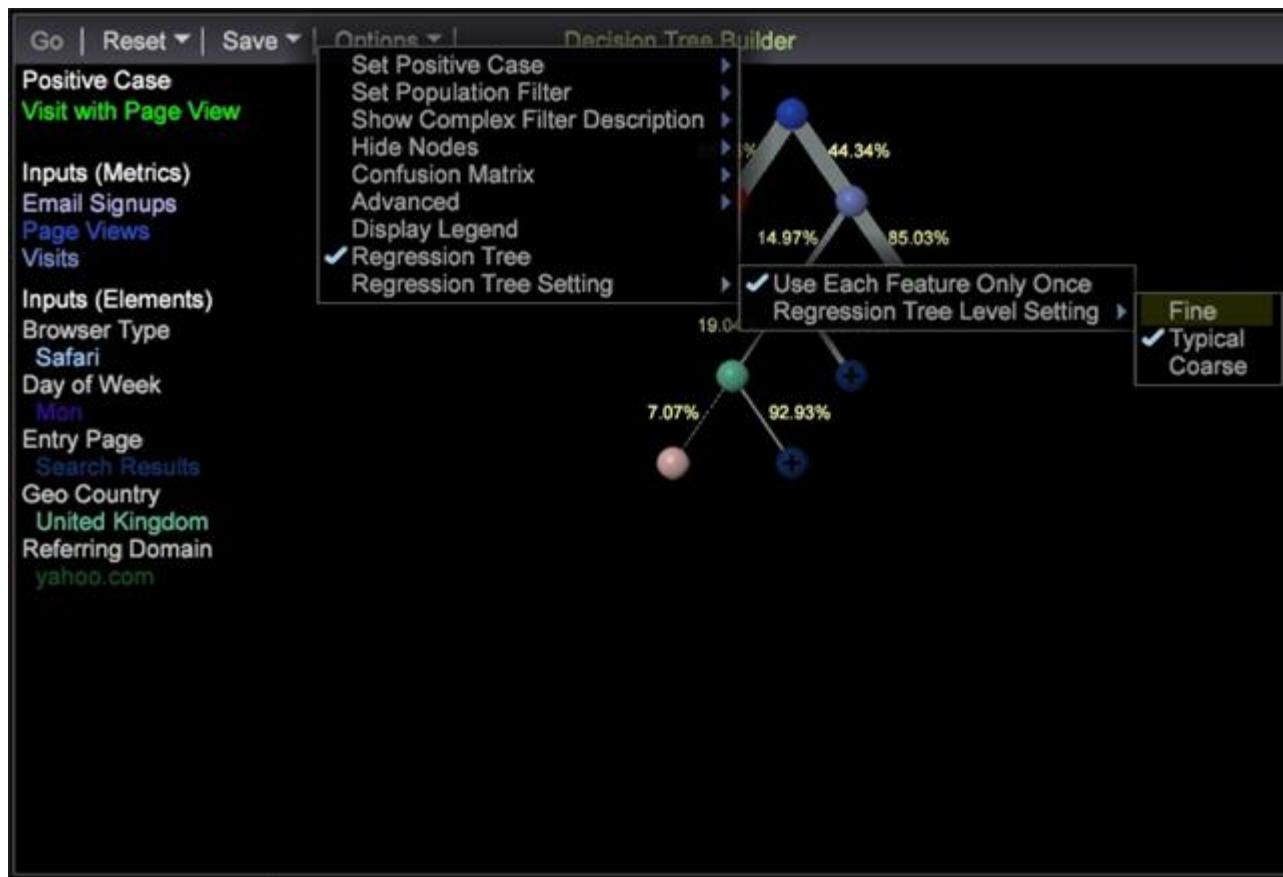
Regression tree option for decision tree

Evaluate a Decision Tree using the Regression Tree option with new sampling and visualization features.

Evaluate a Decision Tree using the Regression Tree option by right-clicking and selecting Options > **Regression Tree** within a Decision Tree visualization.

Updated Decision Tree builder: The new algorithm was introduced for building a *Decision Tree*. It handles more general data and provides a more informative visualization to improve the precision of the prediction.

Improved data sampling module: An updated adaptive sampling scheme helps Decision Tree and Propensity Score achieve higher precision results.



Green and red indicate true or false. The saturation of color—such as deep red versus light red—is used to indicate probability. For example, a node with deep red has a very high probability to be false, while a node with light red has a lower probability to be false. A node with deep green has a very high probability to be true.

All Decision Trees have varying branch widths to indicate the level of traffic for that branch of the tree.

In a Decision Tree visualization, right-click and select Options > **Regression Tree**. When selected, additional settings are provided:

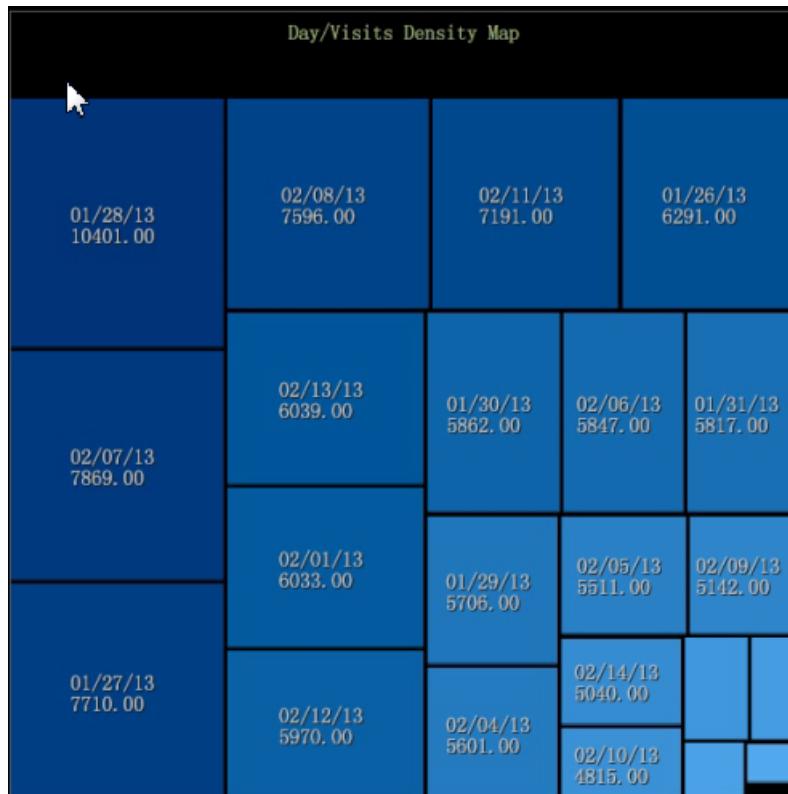
Regression setting	Description
Use Each Feature Only Once	Selecting this option will not use a feature more than once (like the original decision tree)—so if you have five inputs, the tree will be no more than five levels and the tree structure will look similar to a Decision Tree (but a bit more complicated). This option will make the tree building

Regression setting	Description
	fast by using each feature only once (like an original Decision Tree). Using this feature is a default setting.
Regression Tree Level Setting	<p>This option controls the complexity of the Regression Tree. Depending on your data, you might need to build a <i>Fine</i> tree (with a complicated structure with more nodes) to get a more meaningful tree classification. If you have much data, then a relatively <i>Coarse</i> tree (less complicated with fewer tree nodes) could work well.</p> <p> Note: <i>Typical</i> is the default setting. There are some extreme cases where the <i>Typical</i> setting doesn't work as well and the <i>Coarse</i> or <i>Fine</i> setting can provide a better view of the data.</p>
	<i>Fine</i> : The most complex tree with the most granular levels of reporting and most branches.
	<i>Typical</i> : The average level of granularity and branches.
	<i>Coarse</i> : The least complex tree with the fewest defined categories and fewest branches.

Density Map

The density map visualization displays elements as shaded rectangles within a square map.

The sizes of the rectangles are dependent on element values, where larger values are represented by rectangles of larger area. Similar to a pie chart, this visualization allows you to quickly see which elements constitute the greatest percentage of the selected dimension.



To create a density map:

1. Open a new workspace.

After opening up a new workspace, you may need to click **Add > Temporarily Unlock**.

2. Click **Visualization > Density Map**.
3. Select a **Dimension** from the menu.

For example, select **Time > Days**.

In contrast, selecting **Time > Hours** would give you more elements with smaller values displaying as smaller rectangles.



Note: You will want to pick a dimension with multiple elements per your needs. The current limit is 200 of the largest elements for each dimension.

4. You can change dimension views by opening **Visualization > Table** and selecting across elements from the table to display in the map.



The map will respond to selections from the table.

5. Hovering over small elements will display their name and value in text that appears near the mouse cursor.
6. Mask elements by right-clicking and selecting **Mask**, then choose an option.

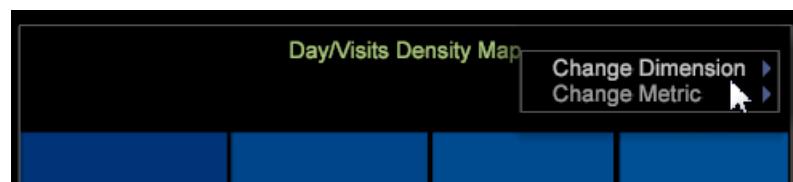


To display all masked nodes, select **Unhide All**.

7. Spotlight elements by right-clicking and selecting **Spotlight**, then choose an option. Spotlighting lets you highlight and dim elements in a range.
8. Add a color legend to the workspace. You can identify values in the map using the color legend.

You can add a color legend to the workspace and the nodes will change color based on the additional dimension of data.

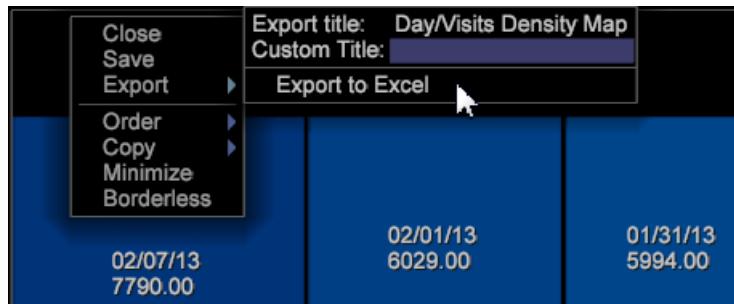
9. Change the dimension or metric by right-clicking the map title and selecting from the menu.



10. Add callouts by right-clicking a cell and selecting **Add Callout**. You can select from different types or visualizations from the menu.



11. As in all visualizations, you can right-click above the title bar for basic commands to Close, Save, Export to Microsoft Excel, Order, Copy, Minimize, and Borderless to display a visualization without a border.



12. The Density Map lets you select and deselect multiple elements similar to other visualizations:

- Left-click to select an element.
- Ctrl + click to select multiple elements.
- Shift + click to deselect an element.
- Right-click within selected elements to open a menu. Then choose **Deselect** or **Deselect All** to clear selected elements.

Additional Options

Right-click the Density Map to open a menu with these options:

Option	Description
Add Callout	<p>Add a text or graphic as a callout in the visualization to further identify or describe an element.</p> <p>You can also select a blank Metric Legend, Table, Line Graph, or Scatter Plot based on the selected element in the Density Map. You can then add metrics and dimensions to these blank visualizations as needed.</p>
Mask	<p>Masking options let you hide selected elements. Right-click to display Mask options.</p> <p>Hide This Element—Choose this option to mask a single element that you have selected.</p> <p>Hide Selected—Choose this option to mask multiple elements that you have selected.</p> <p>Show Top—Choose this option to display only the top 100, 50, 25, or 10 top elements based on the values in the Density Map.</p> <p>Show Bottom—Choose this option to display only the bottom 100, 50, 25, or 10 top elements based on the values in the Density Map.</p>
Spotlight	<p>Spotlighting lets you highlight and dim elements in a range. Right-click to open a menu of options.</p> <p>Show Top—Choose this option to highlight only the top 100, 50, 25, or 10 top elements based on values in the Density Map.</p> <p>Show Bottom—Choose this option to highlight only the bottom 100, 50, 25, or 10 top elements based on values in the Density Map.</p>

Option	Description
Deselect	Select these commands to deselect the current element, if selected, or deselect all elements that are selected.
Deselect All	

Detail table

Navigation title:

Detail tables enable you to view additional information about a subset of data, which is defined by the selections that you make in other visualizations.

The additional information that you see is a sampling of all of the available data.

Domain	6/444 Visitors City	Page Views	18/1K Page Views Page
cox.net	hampton, va	7	/hme/online bank /hme/s1/red/Account login /hme/rts/Our Rates /hme/s1/red/Account login /hme/online bank /pro/evr/t1/free checking account online /pro/evr/t8/Depositing Is Easy
snet.net	meriden, ct	0	
comcast.net	decatur, ga	7	/hme/online bank /hme/online bank /pro/wc/world currency investing /pro/wc/cd/t1/world currency certificates /pro/wc/cd/rts/pop/world currency cd rat /pro/wc/cd/t2/ a Currency CD? /hme/online bank
rr.com	houston, tx	2	/hme/online bank /hme/s1/red/Account login
swbell.net	houston, tx	2	/metals/index.asp /favicon.ico
comcast.net	arlington, tx	0	

The following table describes the elements of a detail table.

Detail Table Elements

Element	Color	Description
Level	Pink	Any countable dimension for which you want to view detailed attribute and metric information. The level is preceded by the number of elements displayed out of the number of elements available, for example 6/444 indicates that 6 elements are being displayed out of a possible 444. In the example above, the level Visitors indicates that all of the detail provided is based on visitor. The level Page Views indicates that all of the detail provided is based on page view. Viewing multiple levels at the same time is useful when you want to analyze data that has different countable parents.
Attribute	Green	Any dimension that is one-to-many or one-to-one with the level, such as City to Visitors. Each row displays the element related to each element of the level you selected. In the example above, the Domain and City attributes list the domain and city for each of the sample visitors.

Element	Color	Description
Metric	Blue	Metric detail about the level you have selected. In the example above, with the level set to Visitors, the metric Page Views displays the number of page views for an individual visitor, while the Page Views level provides the detail about each of those page views.

Let's say you are working with website data and want to find out what pages visitors in particular cities and from particular domains visited during a particular time frame.

First you need to create a visualization that displays the time frame in which you are interested, then you need to select that time frame. Now you can add a detail table to view the desired details for a sample number of visitors in the dataset.

To view the details described above, you must complete the following steps:

1. Right-click within the details table and click **Add Level > Visitor**.
2. Right-click within the details table and click **Add Level > Page View**.
3. Right-click the **Visitors** level heading and click **Add Attribute > Geography > Domain**.
4. Right-click within the Visitors level heading and click **Add Attribute > Geography > City**.
5. Right-click within the Visitors level heading and click **Add Metric > Page Views**.
6. Right-click within the Page Views level heading and click **Add Attribute > Page > Page**.

The following sample workspace shows you the related details for a random sampling of six visitors to the site during the time frame you specified.

Domain	6/444 Visitors City	Page Views	18/1K Page Views Page
cox.net	hampton, va	7	/hme/online bank /hme/s1/red/Account login /hme/rts/Our Rates /hme/s1/red/Account login /hme/online bank /pro/evr/t1/free checking account online /pro/evr/t8/Depositing Is Easy
snet.net	meriden, ct	0	
comcast.net	decatur, ga	7	/hme/online bank /hme/online bank /pro/wc/world currency investing /pro/wc/cd/t1/world currency certificates /pro/wc/cd/rts/pop/world currency cd rat /pro/wc/cd/t2/ a Currency CD? /hme/online bank
rr.com	houston, tx	2	/hme/online bank /hme/s1/red/Account login
swbell.net	houston, tx	2	/metals/index.asp /favicon.ico
comcast.net	arlington, tx	0	

Add a level

- Right-click within the detail table and click **Add Level > <dimension name>**.



Remove a level

- Right-click the existing level heading and click **Remove Level** > *<dimension name>*.



Add attributes and metrics

- Right-click an attribute or metric heading and click **Add Attribute** > *<attribute name>* or **Add Metric** > *<metric name>*.



Remove attributes and metrics

- Right-click the column that you want to remove and click **Remove Attribute** > *<attribute name>* or **Remove Metric** > *<metric name>*.



Export to Microsoft Excel

For information about exporting windows, see [Export window data](#) on page 228.

Filter editors

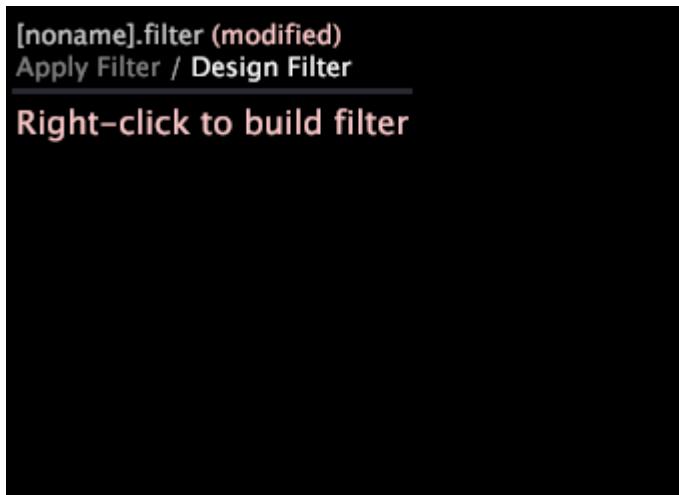
A filter editor is a special visualization for building and editing filters.

While other visualizations in Data Workbench display data from your dataset, a filter editor enables you to specify which portion or subset of the data displayed in the visualizations.



Save Filter button. When you select an item within a workspace—a detail table, visualization, filter, or other item—a **Save Filter** button opens on the toolbar allowing you to create a filter for the item selected. The option for entering a name is provided when the Save Filter icon is clicked. The filter can then be saved to the profile.

You create a filter by opening a filter editor window in your workspace and adding one or more conditions. To assist you in defining the conditions of your filter, editable fields in the filter editor window appear as text boxes or are highlighted when you roll your mouse over them. You simply click within the text box to add text or right-click the highlighted field to see what options are available. The filter that you create reads like a sentence, making it easy to understand what subset of the data is included or excluded. If you would like, you can create a description for each filter condition so that its label in the filter editor is more meaningful to you or your team.



In addition, you can use the filter editor to open and edit any `.filter` file.



Note: When you open the filter for the first time, the filter editor updates the filter's definition to match the filter editor format.

Working with filter conditions

Navigation title:

Information about working with filter conditions including creating a new filter and adding a condition to a new filter.

Create a filter

- Open a filter editor in your workspace by right-clicking **Add Visualization > Filter Editor**.
- or-
- If you already have a filter editor open and a filter loaded, right-click the current filter's name and click **New Blank Filter**.

Add a condition to a new filter

1. Create a new filter. Make sure that Design Filter is highlighted (as opposed to Apply Filter), indicating that you are working in Design Filter mode.
2. Right-click within the area marked **Right-click to build filter** and select one of the following options:
 - To create an inclusion filter, click **Include group with**.
 - To create an exclusion filter, click **Exclude group with**.
3. Select the type of condition to add to the filter.

The following table provides descriptions of the available filter condition types:

Condition Type	Description
workspace selection	<p>Defines a filter condition based on the selections in the workspace. This option is available only if one or more selections exist within the workspace.</p> <p>To view more information about the selection, right-click the condition and click View Details. A callout appears for the condition.</p> <p>If you make another selection in the workspace, you can add the selection as a subcondition of the first selection. The selections are grouped together as logical ANDs. Therefore, the data that is included or excluded by the condition must satisfy all of the workspace selections.</p>
at least one	Defines a filter condition based on the existence of at least one (any) element of a dimension that you choose. To edit the condition, right-click the condition and click Change condition to. Click one of the available dimensions.
formula	<p>Defines a filter condition based on the formula that you enter. You must use the appropriate syntax for the filter to work.</p> <p> Note: For information about the syntax for defining filters, see Syntax for filter expressions on page 521.</p>
metric value	<p>Defines a filter condition based on a metric value that you specify.</p> <p>To define the condition, follow these steps:</p> <ul style="list-style-type: none">• Right-click [choose level] > Change level to select the level and metric from a list of dimensions in your dataset.• Right-click [choose metric] > Change metric to select the metric from a list of metrics in your dataset.• Right-click less than and click Change comparison to select one of the available comparison conditions (less than, more than, exactly, at least, or at most).• Type the desired value for the metric.
first/last	<p>Defines a filter that lets you include or exclude a level with a specified dimension. For example, you might specify a first/last filter to include (or exclude):</p> <p>Sessions whose last Page View has a Page of /home/rts/Our Rates.</p> <p>To define a First/Last condition:</p> <ul style="list-style-type: none">• Choose Include group with or Exclude group with > first/last as a new condition in the Filter Editor.• Right-click [choose container] > Change container to select the container.• Right click first or last to specify the level.• Right click to specify a dimension, then type a value into the available field.• Click Apply.

The filter in this example defines a first/last filter for users whose last page view was /hme/rts/Our Rates:



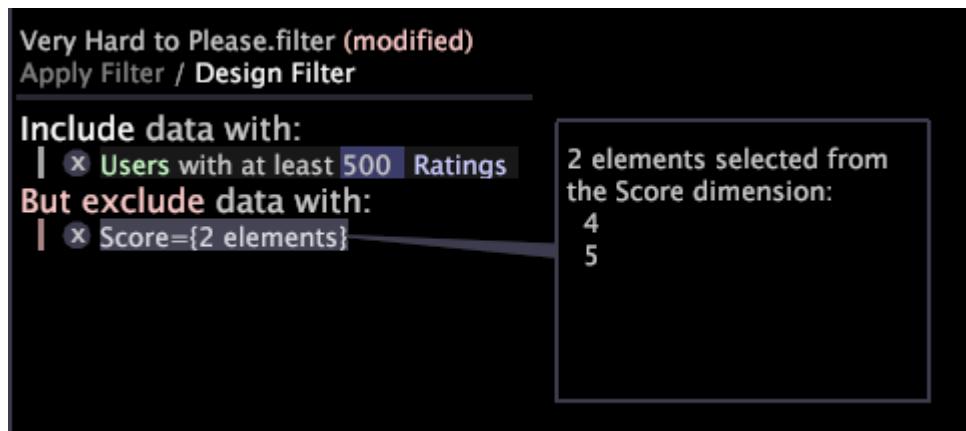
4. (Optional) To add more conditions to your filter, right-click within the area of the window where you are building your filter and select the type of filter (see Step 2) and the condition rule (see Step 3).



Note: Multiple inclusion conditions are grouped together as logical ORs. Therefore, the data included by the filter must satisfy at least one of the defined inclusion conditions. Multiple exclusion conditions also are grouped as logical ORs. To be excluded, data must satisfy at least one of the exclusion conditions.

The filter in this example defines a subset of data consisting of movie viewers (users) who rated lots of movies but did not give any one movie a high score (4 or 5). This filter (appropriately named Very Hard to Please) consists of two conditions:

- **A metric value condition:** The condition includes users who have rated at least 500 movies.
 - **A workspace selection condition:** The condition excludes users who gave any one movie a score of 4 or 5.
- The callout tells you that 4 and 5 were the elements selected from the Score dimension.



Delete a filter condition



Note: You can delete conditions only when you are working in Design Filter mode. If you have applied a filter to your workspace, you must click Design Filter to return to Design Filter mode before you can delete one or more of the filter's conditions.

- Click the **x** to the left of the condition to delete it.

Edit a condition description

You can add descriptions to each of the conditions that you add to a filter. You can edit or remove the descriptions as desired.



Note: Descriptions of conditions appear only when you are working in Design Filter mode.

- Right-click the condition and click **Edit description**.
 - To add or edit a description, type the description in the **Edit condition description** field. The description appears in quotation marks above the condition in the filter editor window.



- To remove a description, click **Remove description**. The condition remains in the filter editor window.

Save and load filters

Navigation title:

Steps to saving and loading filters.

1. Right-click the filter's name (or **[noname]** if the filter is new) and click **Save Filter**.
2. Enter the desired name for the filter in the **Save As** label.



3. Click **Save Filter**.

After you save the filter, you can load it to use in other workspaces. See below how to load an existing filter or reference it by name in a **Metric Editor**. See [Work with derived metrics](#) on page 461.



Note: If you would like all users of a profile to use the filters that you save, you must upload them to the appropriate profile using the **Profile Manager**. See [Publish files to your working profile](#) on page 459.

To load an existing filter

- Right-click the filter's name and click **Load Filter > <filter name>**.

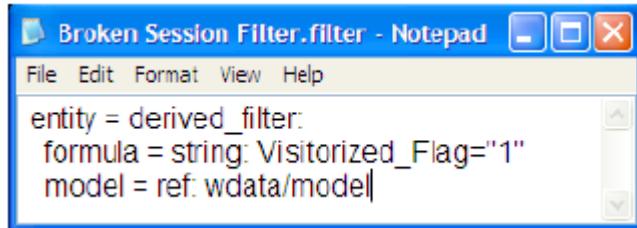
Edit filter files manually

Navigation title:

Steps to editing a .filter file manually.

1. In the **Profile Manager**, click **Filters**, then locate the file that you want to edit.

-
2. If the file is located on the server, right-click the check mark in the appropriate column next to the name of the filter and click **Make Local**.
 3. Right-click the check mark for the filter in the **User** column and click **Open > in Notepad**. The **.filter** file opens.



```
entity = derived_filter;
formula = string: Visitorized_Flag="1"
model = ref: wdata/model|
```

4. Edit the file as necessary. For the syntax to use when defining filters, see [Syntax for filter expressions](#) on page 521.

Apply a filter to a workspace

Navigation title:

Steps to apply a filter to a workspace.

- In the filter editor window, click **Apply Filter**.

The border of the filter editor window turns yellow to indicate that the filter has applied a selection to the workspace.

Funnel Visualization

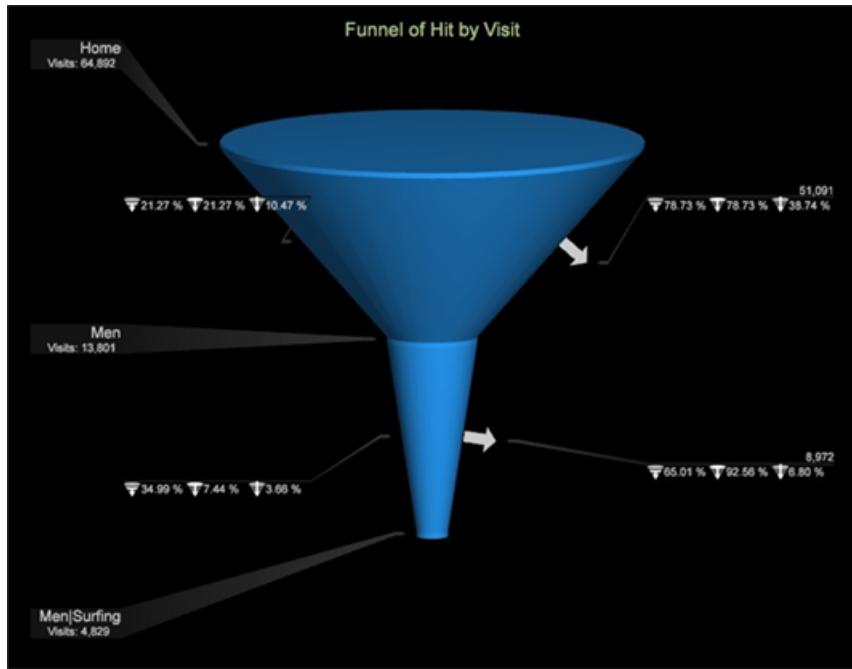
The Funnel visualization lets you identify where customers abandon a marketing campaign or divert from a defined conversion path while interacting with your website or cross-channel campaign.

The funnel visualization identifies converting page views to purchases, and lets you see where customers fall out of the process. Gaining visibility into customer decisions at each step lets you understand where they are being deterred, what path they tend to follow, and when customers leave your site and where they go.

Proper understanding of customer navigation allows architects to design and target web pages based on trending interest, and lets marketers better interpret customer actions and interactions against specific campaigns.

About the Funnel

The Funnel visualization is much like the Path Browser in building a path to analyze your visitor's sequence of hits, to identify who is falling out (who left the path) or falling through (who followed the path). It also identifies where visitors went after each step along the campaign path, and where they navigated to after falling out or falling through the defined steps.



In addition to web data, you can perform funnel analysis across all types of cross-channel data supported by the platform. Any data element from any source can be represented in the Funnel visualization.

The Funnel visualization provides various levels of data:

- **Funnel by Visit.** This funnel is based on a "per visit" interaction.
- **Funnel by Visitor.** This funnel is based on a "per visitor" interaction. This view shows the data based on the visit organized by visitor.
- The **Clip dimension and Level dimension.** Funnel dimensions can be modified by choosing Clickthrough, Hit, Product, Visit or Visitor Level dimensions.
- **Metrics.** You can change the underlying Funnel metrics from Visits used to build the funnel to any metric in your profile. The Funnel visualization allows you to drill into your data and analyze the patterns of visitors with multiple baseline metrics.

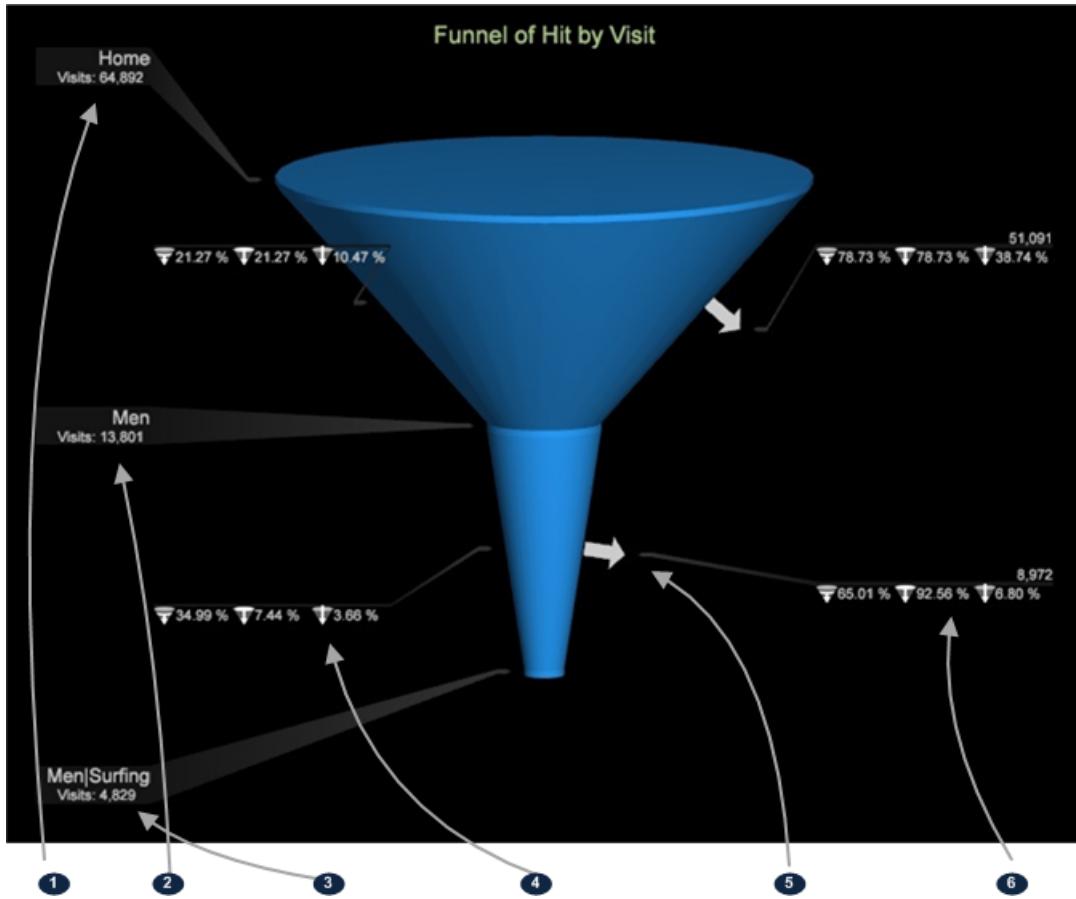


Note: In the Funnel visualization, you can apply any dimension at the countable level. This is in contrast to the Path Browser and Process Maps that restrict your choice of metrics. Analysts have multiple choices when applying a metric in the Funnel visualization.

Funnel Features

The Funnel visualization includes features to build a funnel with multiple dimensions, raw visitor numbers, percentage of visitors at each step, and separate scopes.

Here are the basic features of the funnel visualization.



1	First Element	First funnel step in the process.
2	Third Element	Third funnel step in the process. ⚠ Note: The selected elements do not have to be from the same dimension.
3	Fall-through Percentage	Percentage who completed the defined path displayed in three scopes.
4	Fallout Browser	Fallout arrow. Right-click and select Add Path Browser to see what other path visitors took.
5	Percent Fallout	Percentages that describe three scopes of fallout for users that did not complete the path. Percentages are presented in three scopes: 
		⚠ The percentage of fallout from the step previous to this step.
		⚠ The percentage of fallout from the first step in the funnel.



The percentage of fallout based on the total number of visitors.

Funnel Steps

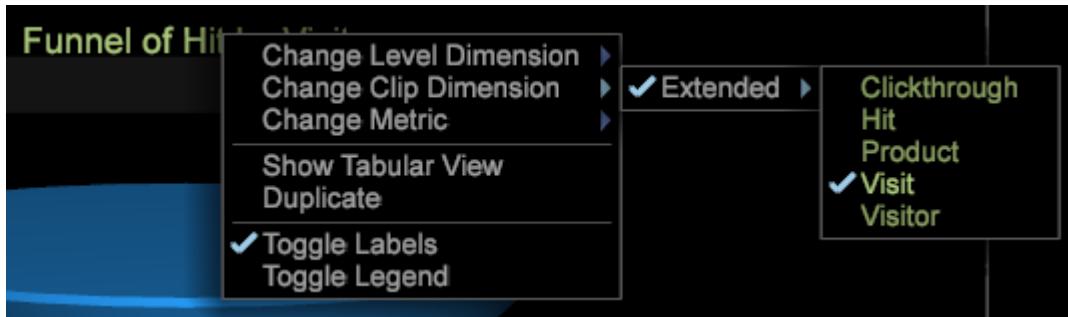
The disks in a funnel represent the steps in the navigation, the cones represent the fall-through from one step to the next, and the arrows represent the fallout. Clicking a cone will select the users who fell through at that point and include them in the current workspace filter. Clicking an arrow will select the visitors who fell out.



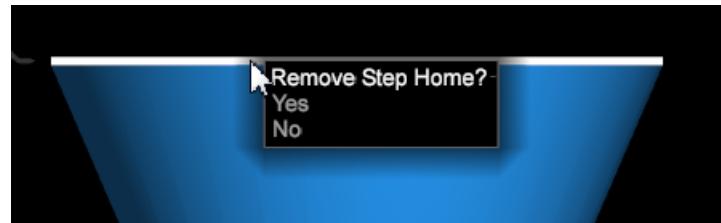
Note: The Funnel visualization has a limit of eight steps that can be applied.

Additional Funnel Features and Functionality

- **Adjust the clip and level of the funnel.** Select the Funnel option from the Visualization menu. After the funnel is created, you can right-click on the title to adjust the clip and level to any countable metric in your system.



- **Drag more elements.** Add more elements to your funnel by dragging and dropping them from the Dimension table to the funnel using the <Ctrl> + <Alt> keys. You can drag multiple steps at the same time from the Dimension table by selecting multiple items (using <Ctrl> + click) and then dragging them to the Funnel visualization using the <Ctrl> + <Alt> keys..
- **Delete a step:** Delete elements by right-clicking on the step in the visualization and clicking **Yes**.



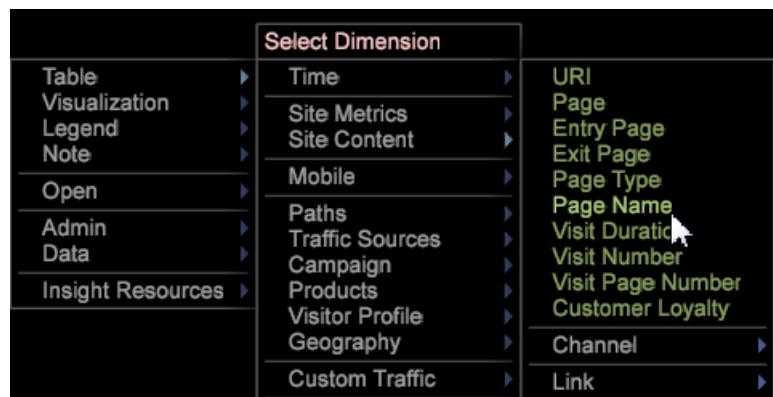
- **Rearrange the steps you have dragged to the funnel.** Simply click the step to select it and drag it to another position to rearrange the steps.
- **Open a Path Browser.** You can see more detail about where customers fall through or fall out of the process through the [Adding a Path Browser](#) on page 154 feature.
- **Add more steps.** You can add a maximum of eight steps to each funnel visualization.
- **Change the metric.** The metric can be changed so the steps are counting visits or some other metric at each step. Available options vary by dataset.

-
- **Display in a tabular view.** Right-click the title to display the Funnel Visualization menu and click **Show Tabular View**. Once in tabular view, you can select **Show Graph View** to return to graphic representation of the funnel. To open the Tabular View, right-click on the title and select Show Tabular View from the menu.
 - **Compare sequences.** An efficient way to compare two similar sequences is to display their two visualizations side-by-side. You can also display both the tabular view and the graph view side-by-side using the Duplicate feature. To open, right-click on the title and select Duplicate from the menu.

Building a Funnel

Follow these steps to build a new funnel visualization

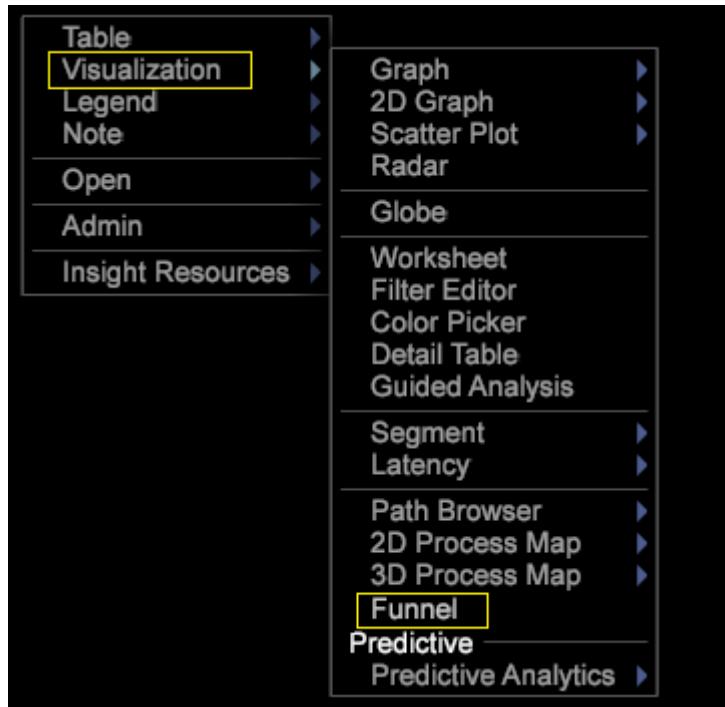
1. From a new workspace, open a dimension table.



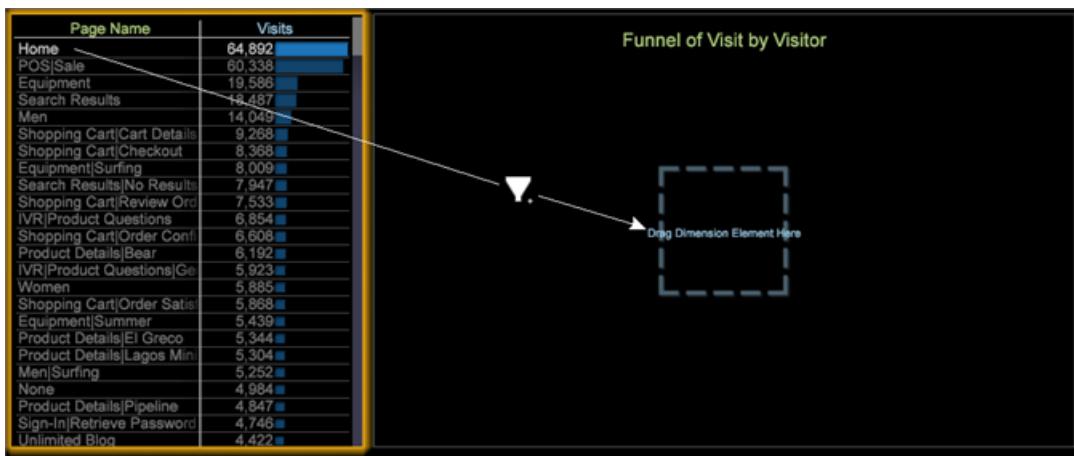
For example, right-click on the workspace and select **Table > Site Content > Page Name**

2. Open a funnel visualization.

For example, right-click on the workspace and select **Visualization > Funnel**.



3. Add a dimension by selecting the element, then drag while then holding <Ctrl> + <Alt>.
4. Drag the element into position over the target zone, **Drag Dimension Element Here**.



5. Add additional dimensions.

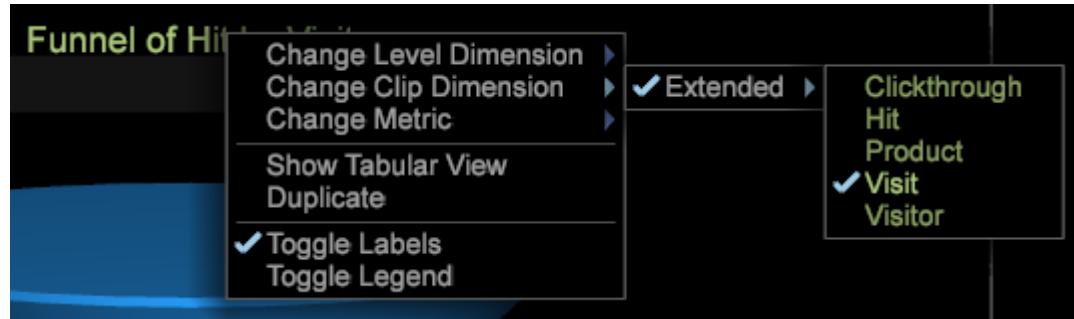
You can add up to a total of eight steps to your funnel.

6. Release the table selection you made to view your funnel based on a total audience.

Funnel visualizations follow the pattern of all other visualizations, allowing the selected element to control the segment of data that you are visualizing.

7. Modify the Funnel clip dimension, level dimension, and metric.

Right-click on the title to adjust these options based on the metrics and dimensions defined in your custom profile.



From the Funnel, you can identify your fall out and fall through rates and the percentage of customers converting sales and following campaigns or content paths. For additional information, you can open a [Adding a Path Browser](#) on page 154 to see a detailed list of the pages taken before arriving at the selected page, and the paths taken after they left.

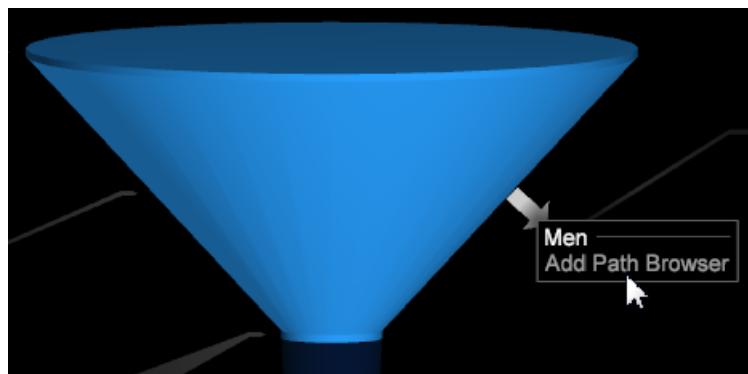
See additional [Funnel Features](#) on page 149.

Adding a Path Browser

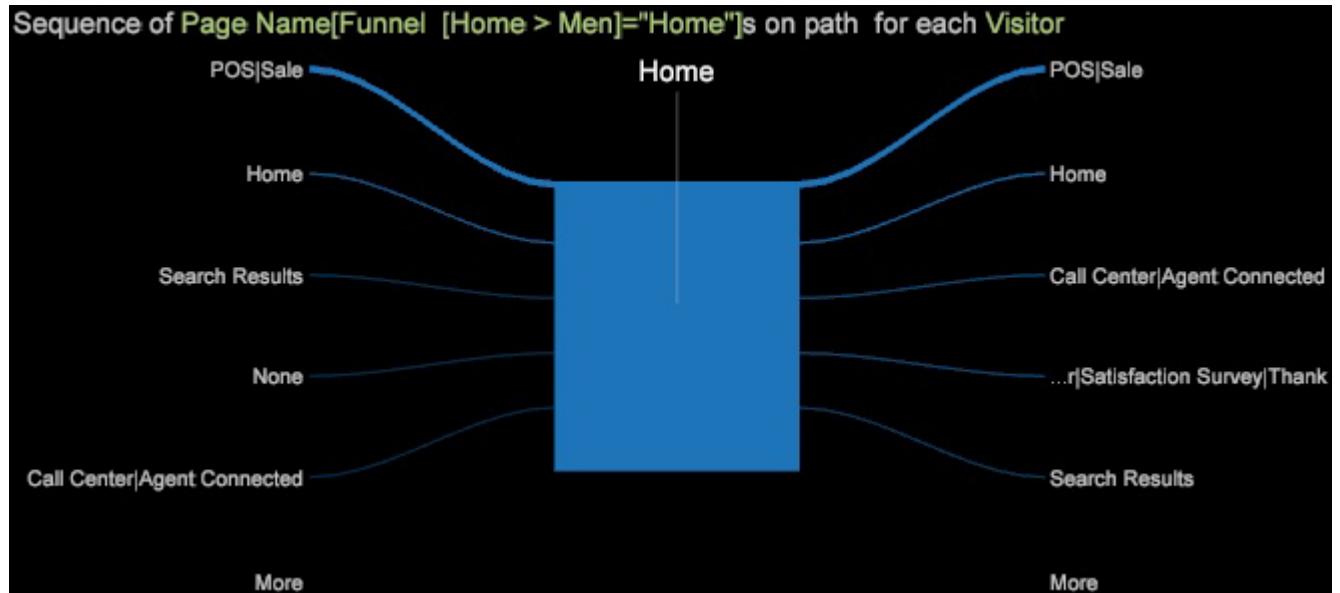
From the Funnel visualization, you can open a Path Browser to identify additional detailed visitor fallout and fall through paths.

You can identify where visitors "fallout" (who left the path) or where visitors are "falling-through" (who followed the path). You can open multiple Path Browsers from the same Funnel visualization by selecting fallout arrows or fall through cones.

1. Open the Path Browser for the fall out visitor, the visitors who left the site during a specific step. Right-click the fallout arrow and select the Add Path Browser option.



From the Path Browser, you can see where visitors went before hitting a selected element in the Funnel (on the left), and where they navigated to after leaving the page (on the right).



In the Path Browser visualization, the width of the lines going in and out of the selected dimension identifies the volume of traffic. For example, before hitting the Men's page, a few visitors came in through other routes, but most came in through the Home page. When leaving, most went to the Men Surfing page.

2. Open the Path Browser for the "fall-through" visitors. Right-click the cone in the funnel to identify the path for visitors who fell through, or moved to the next step, of the funnel.
3. Click the **More** button at the bottom of the Path Browser visualization to open a table with all visits listed in a tabular view for the previous and current visits.

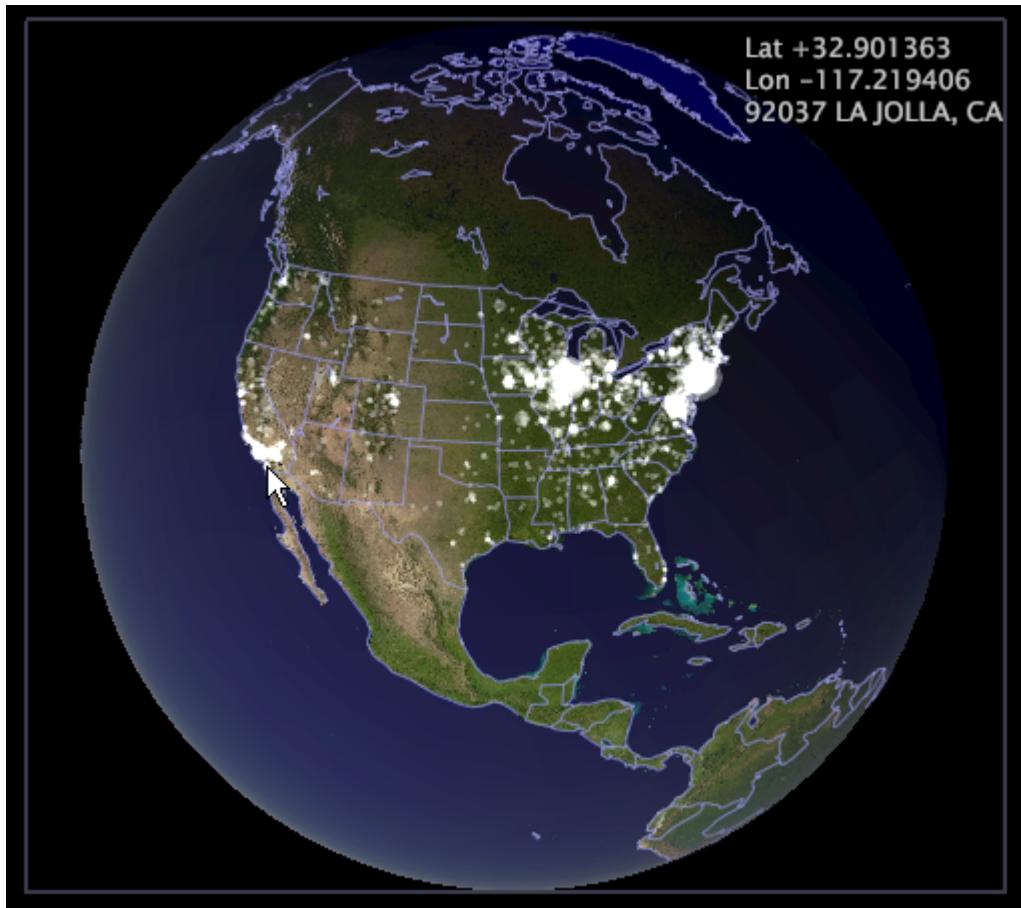
More Page Name[Funnel [Home > Men] = "Home"]s on path	Visits
POS Sale	1,132
Home	253
IVR Product Questions	30
IVR Product Questions General Questions	41
None	73
IVR Product Questions Specific Product Error: Product Not Fo	4
Search Results	74
Product Details Bear	34
Equipment Surfing	30
Product Details Lagos Mini Longboard	20
IVR Welcome	28
Call Center Agent Connected	107
Call Center Satisfaction Survey	31
Men	0
Men Winter	0
Men Winter Skiing	0
Product Details Midweight	0
Men Surfina	5

Globes

Navigation title:Globes

Data Workbench can display topographical imagery on a detailed three-dimensional (3-D) map of the world.

The following globe displays the geographic distribution of a company's customers.



Working with globes

Navigation title:

When you move your mouse over the map, any available data corresponding to the geographic position of your pointer location is displayed in the upper-right corner of the visualization.

You also can zoom and rotate the globe to focus on a desired geographical region.

To move a globe within a window

- Click the top border of the visualization and drag the map to the desired location.

To zoom in on a globe

1. Place your mouse pointer in the globe window.
2. Press both mouse buttons and move the mouse forward (away from you).

To zoom out on a globe

1. Place your mouse pointer in the globe window.
2. Press both mouse buttons and move the mouse backward (toward you).

To rotate a globe

- Right-click within the visualization and drag the mouse in the desired rotational direction.

To toggle lighting in a globe

You can turn on or turn off lighting in a globe based on the position of the sun given the time of day. The time is set by the Data Workbench server.

-
- Right-click within the visualization and click **Toggle Lighting**. An X to the left of the menu item indicates that lighting is enabled.

Display layers

Navigation title:

The Geography profile stores a collection of imagery layers and files.

When you display a globe, you can select which of the available layers to display for a particular analysis task:

- **Blue Marble 2km:** This layer displays the globe. When this layer is not selected, the globe is not visible.
- **IP Coordinates:** This element point layer displays the latitude and longitude of locations in your dataset based on visitor IP addresses.
- **Zip Points:** This layer displays the latitude and longitude of locations in your dataset based on a list of United States ZIP Codes provided by Adobe. The `Zip_Points.txt` lookup file contains the ZIP Code, latitude, and longitude data to be displayed, while the `Zip_Points.layer` file contains the configuration parameters needed to display this data on the globe. Both of these files are stored in the `Profiles\Geography\Maps` folder within the Data Workbench server installation directory.
- **Other available layer names:** Each layer name represents a `.layer` file stored in the `Profiles\Geography\Maps` folder, `Profiles\IP Geo-location\Maps` folder, or `Profiles\IP Geo-intelligence\Maps` folder within the Insight Server installation directory.



Note: To have the IP Geo-location or IP Geo-intelligence profile, you must subscribe to either the IP Geo-location or IP Geo-intelligence data service, respectively.

To control the order in which your layers display, you can use an `order.txt` file. See [Customize a menu using order.txt files](#) on page 486.

To toggle layers in a globe

- Right-click within the visualization and click the desired layer name. An X to the left of the layer indicates that the layer is visible.

Make selections on a globe

Navigation title:

When you make a selection on a globe from a particular layer, all of the other layers show data related to the selected data only.

You can use multiple globes together to display different subsets of layers, side by side. See [Make selections in visualizations](#) on page 281.

Graphs

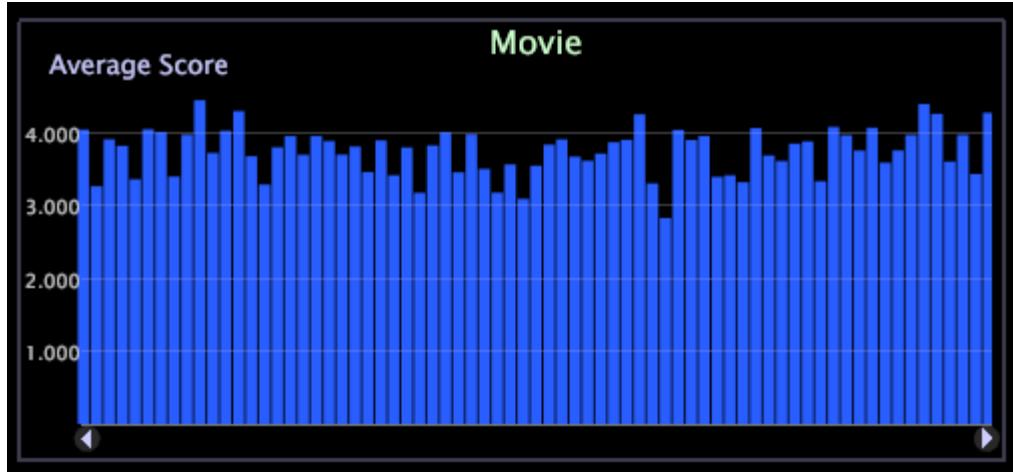
Navigation title:Graphs

A graph can display one or more metrics across one or two dimensions of data.

Graphs are very powerful visualizations, and learning to use them enables you to answer a large variety of questions.

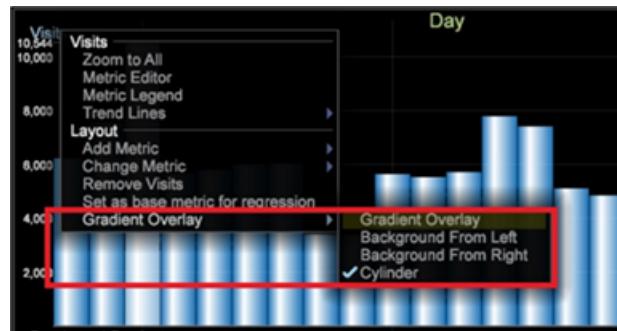
Graphs lets you easily see trends in business metrics over continuous data because you can select a subset of a dimension over which the metrics are graphed (such as a specific time frame) and zoom in to look at the data for that period in greater detail. For instructions to zoom in a graph or table, see [Zoom in or out in a visualization](#) on page 292.

In the following example, you see a bar graph in which the number of sessions associated with each day is depicted by the length of the bar. When you move the mouse pointer over the bar for a particular day, the numerical value represented by the bar displays.



Gradient Overlay on a Graph

You can right-click and select a **Gradient Overlay** from the menu to better view graph elements.

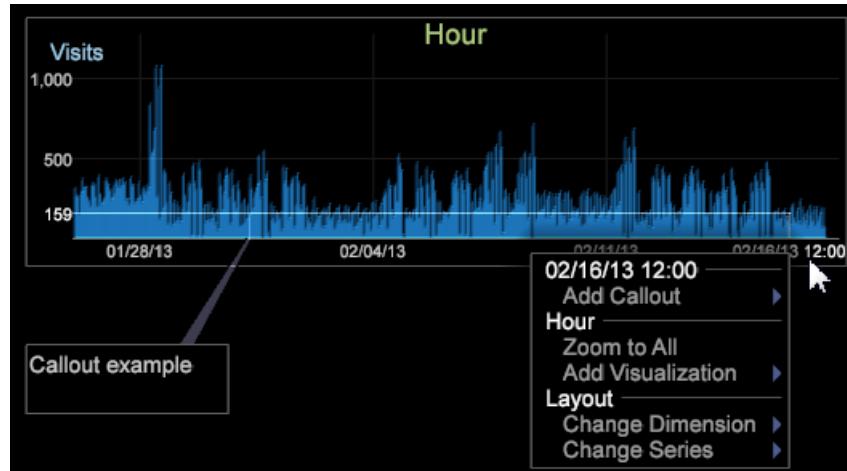


- **No Gradient Overlay** (default). Select to show bars without gradient overlay applied.
- **Background from Left**. Select to show gradation of colors across all bars from left to right.
- **Background from Right**. Select to show gradation of colors across all bars from right to left.
- **Cylinder**. Select to show gradation of colors from the center of each bar to the edge of each bar.

Adding Callouts to a Graph

You can also add callouts to linear graphs. See [Adding callouts to a workspace](#) on page 293.

To add a callout to Graph visualizations, you need to right-click at the bottom of the visualization (the base axis) to open a menu and select **Add Callout**.

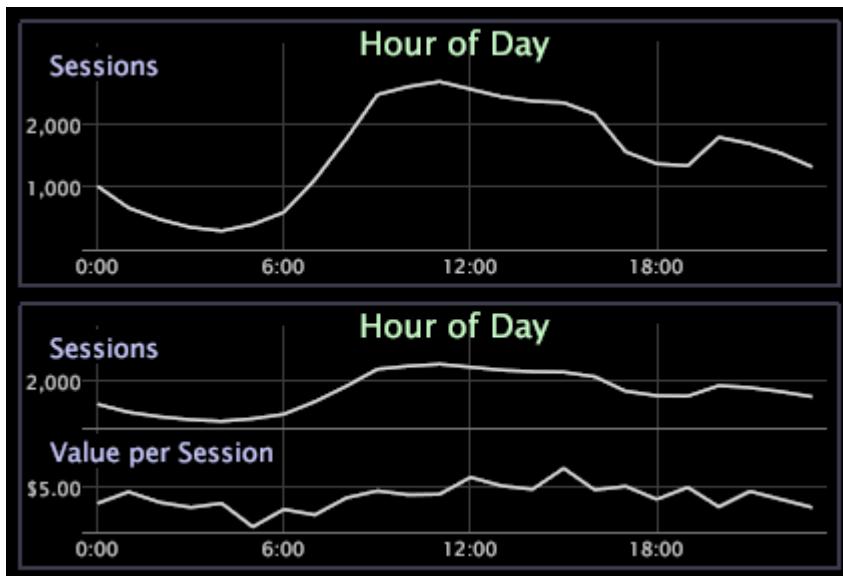


Add and change a metrics

Navigation title:

Steps to add a new metric or to change a metric.

The following examples show the Sessions metric graphed over the Hour of Day dimension (in the top line graph) and the same line graph with a second metric, Value per Session, added to the left axis (in the bottom line graph).



To add a new metric

- Right-click an element or the label of a metric and click **Add Metric** > *<metric name>*. The metric is added below the metric that you selected originally.

To change a metric

- Right-click an element or the label of the metric that you want to change and click **Change Metric** > *<metric name>*. The new metric replaces the original metric.

Change a dimension

Navigation title:

Steps to change a dimension.

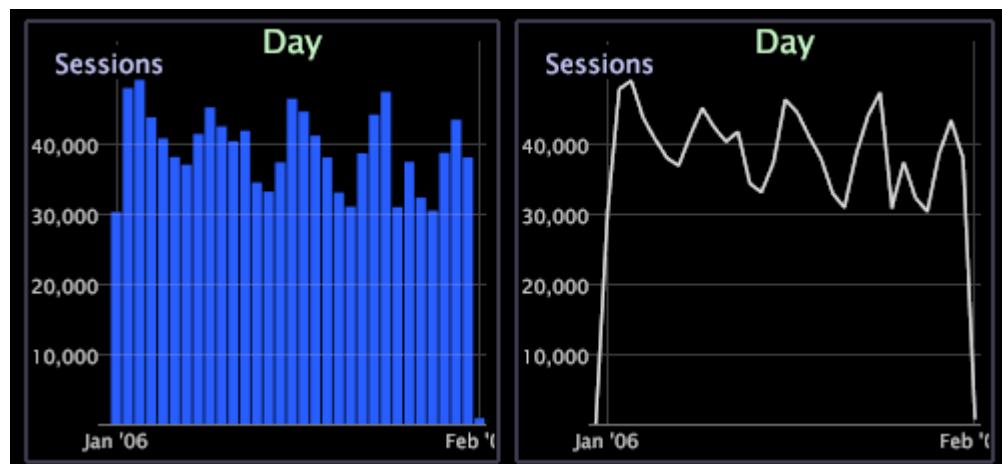
-
- Right-click an element or the label of the dimension that you want to change and click **Change Dimension > <dimension name>**.

Change a graph display

Navigation title:

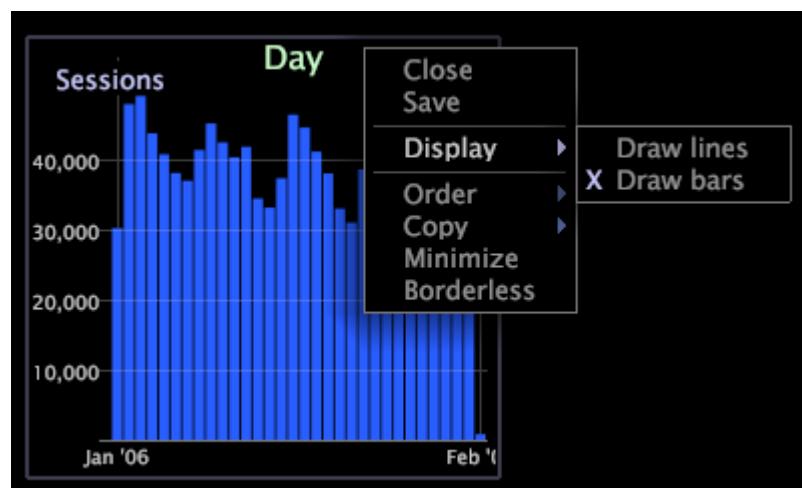
By default, graph visualizations display bars, but you can easily display the bars as lines and vice versa.

Both display types are shown in the following example.



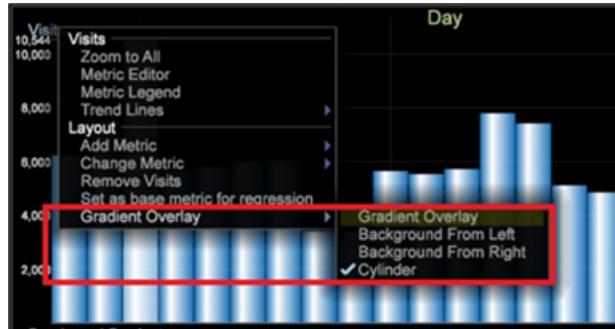
To select the display type

Right-click the dimension label at the bottom of the graph and click **Display > Draw lines** or **Display > Draw bars**. An X appears to the left of the active display mode.



Gradient Overlay on a Graph

You can right-click and select a **Gradient Overlay** from the menu to better view graph elements.



- **No Gradient Overlay.** Select to show bars without gradient overlay applied.
- **Background from Left.** Select to show gradation of colors across all bars from left to right.
- **Background from Right.** Select to show gradation of colors across all bars from right to left.
- **Cylinder.** Select to show gradation of colors from the center of each bar to the edge of each bar.

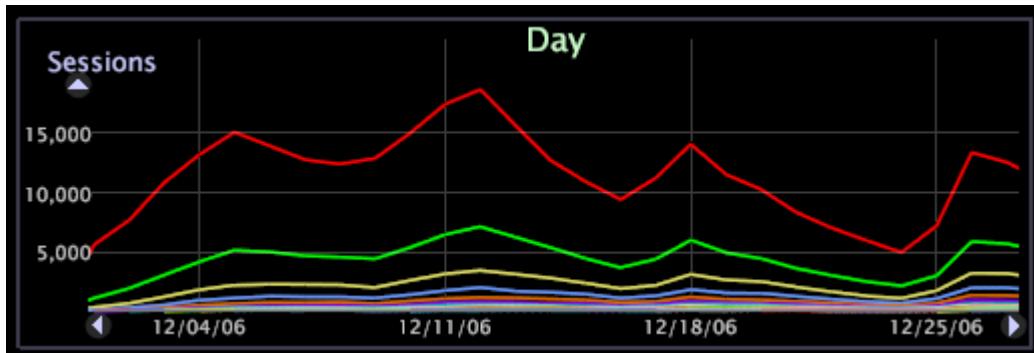
Change a series

Navigation title:

Elements within a dimension can be color-coded as a series within a line graph so that each element can be more easily tracked across other visualizations in the workspace.

The series feature applies to graphs only when the line graph is displayed as lines, not bars.

When a series is active, data for the different elements of the dimension is shown in alternating colors. It is useful to add a table to the workspace showing the same dimension to provide a series legend for the selected dimension. See [Enable a series legend](#) on page 421.



To enable a series

- Right-click an element or the label of the dimension and click **Change Series > <dimension name>**.

To disable a series

- Right-click an element or the label of the dimension and click **Change Series > None**.

Regression Analysis Graph

The Bar Graph in Data Workbench now includes a regression comparison for multiple metrics across multiple graphs.

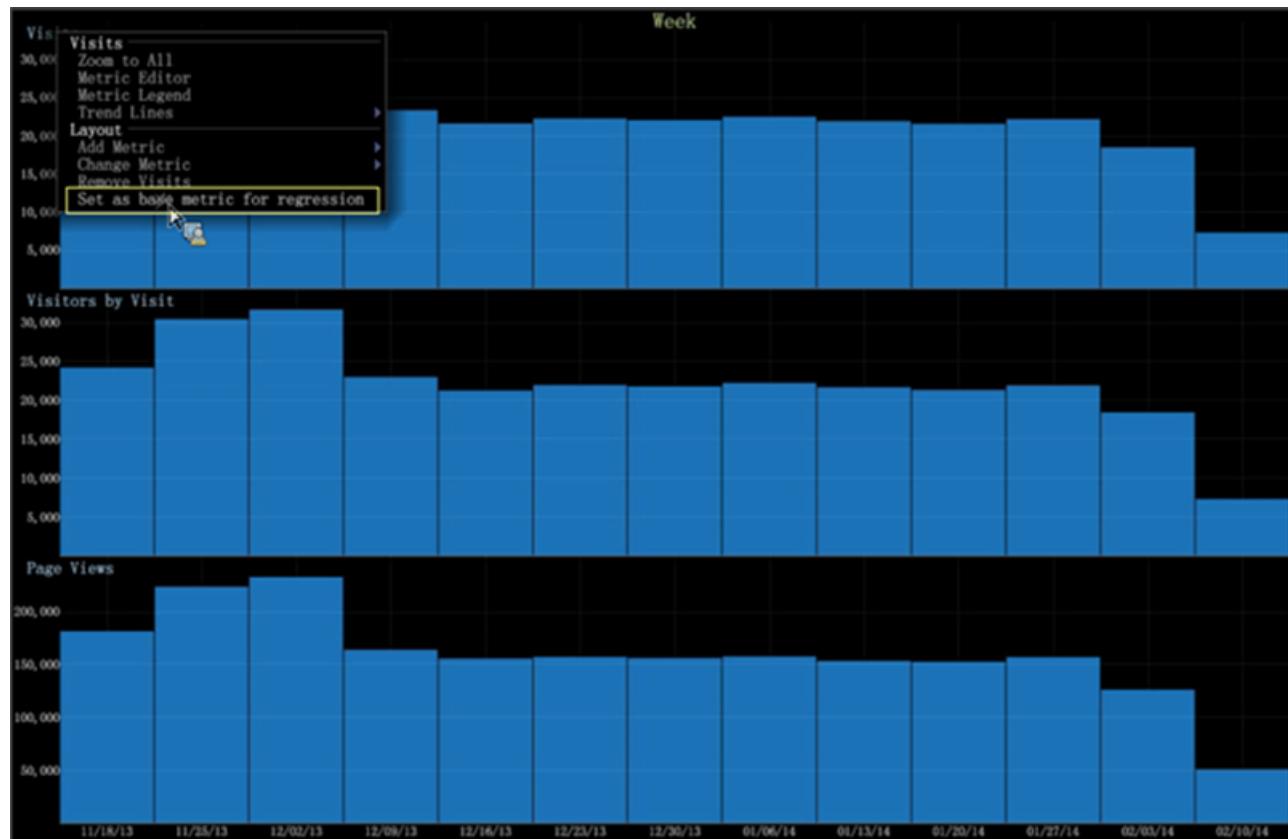
[Bar graphs](#) in Data Workbench let you regress metrics in one graph to metrics in another graph. If you have multiple graphs, you can compare a metric (as the independent variable) to a graph evaluating other metrics (as dependent variables). This lets you determine the strength of the relationship between one dependent variable (the metric established first) and a series of other changing metrics (regressions with the established dependent metric).

The regression analysis on a graph visualization allows analysts to perform "what-if" scenarios. For example, if visits increase to this level, what impact will this increase have on revenue?

Setting up Regression Analysis

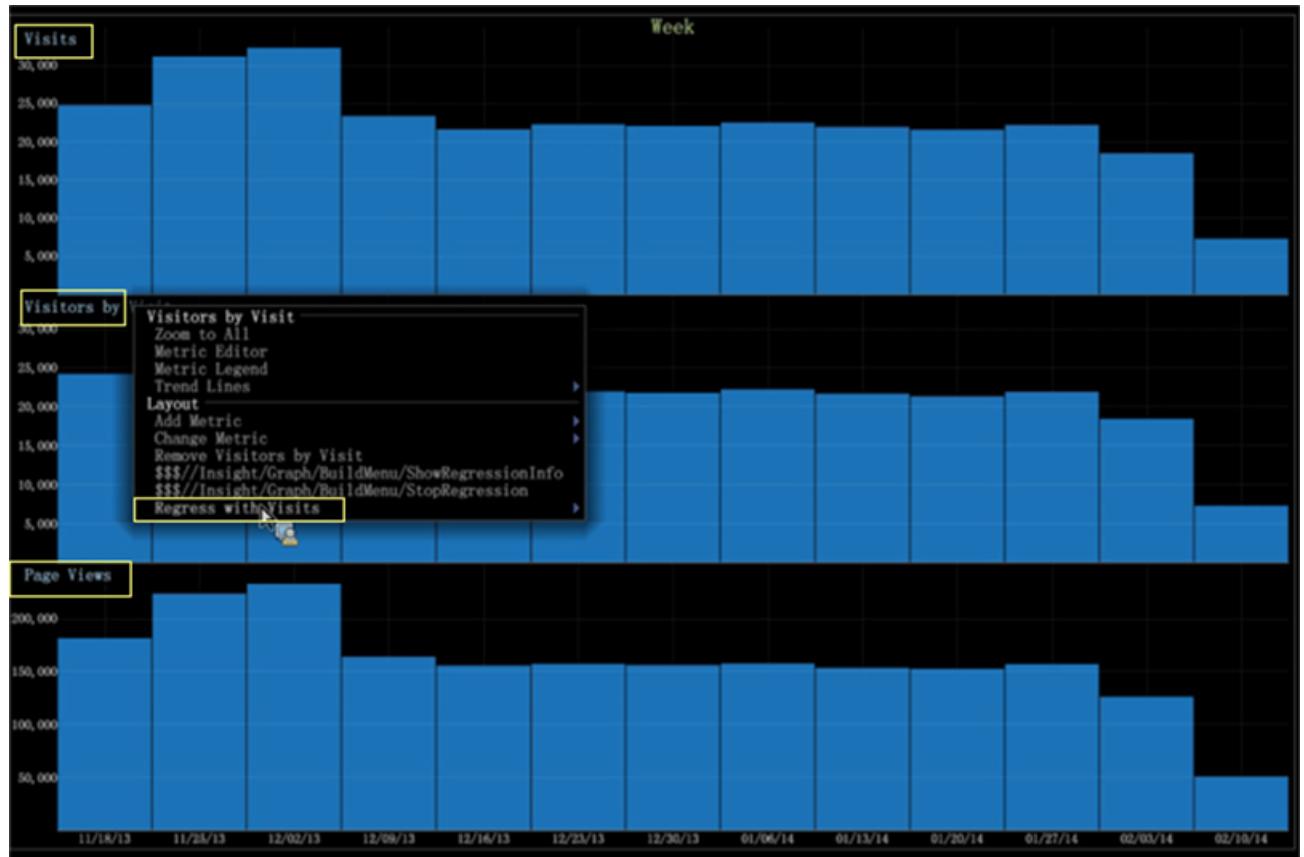
1. Select graph as a dependent metric for a regression comparison.

Right-click on the graph and select **Set as base metric for regression**.



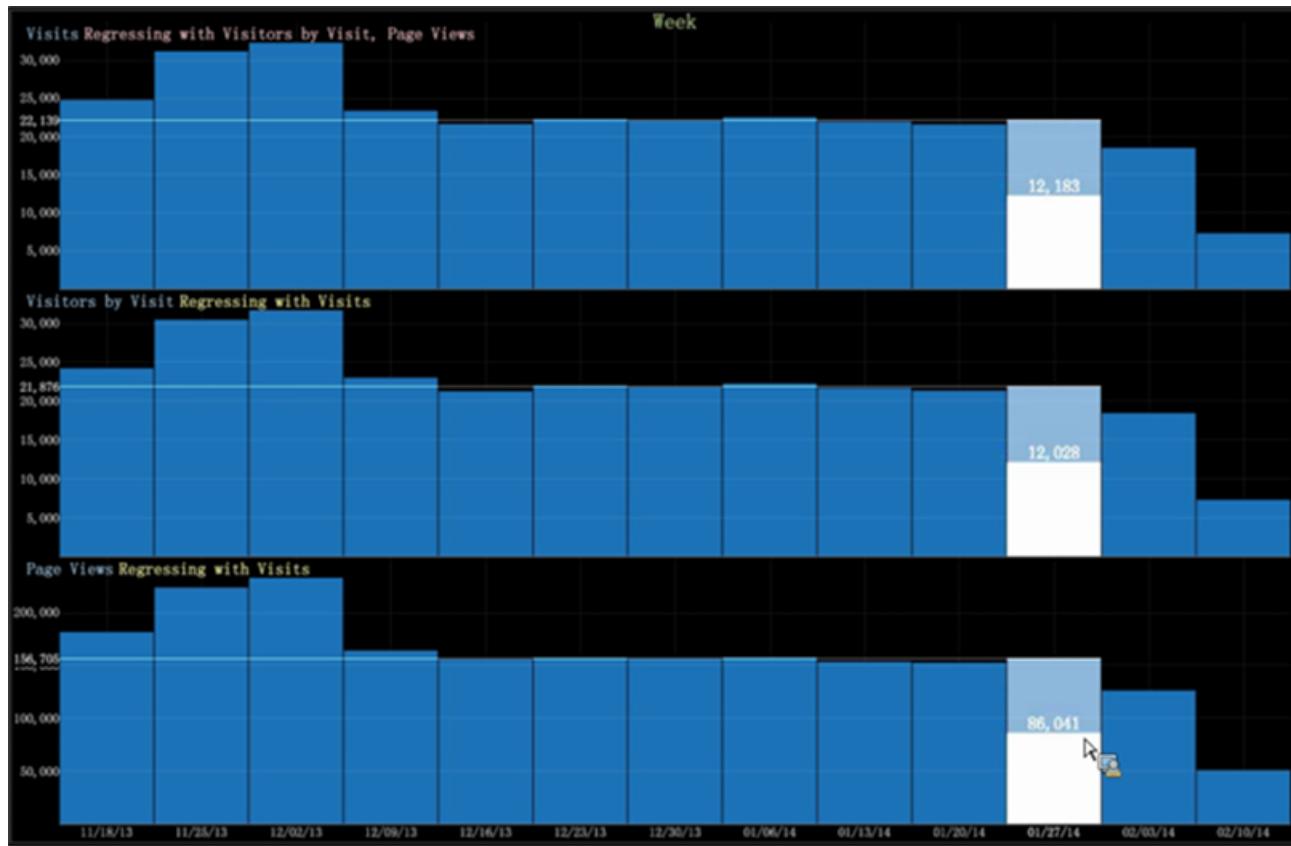
2. Set other metric graphs as independent variables.

Right-click metric and select **Regress with <base metric name>** for other metrics.

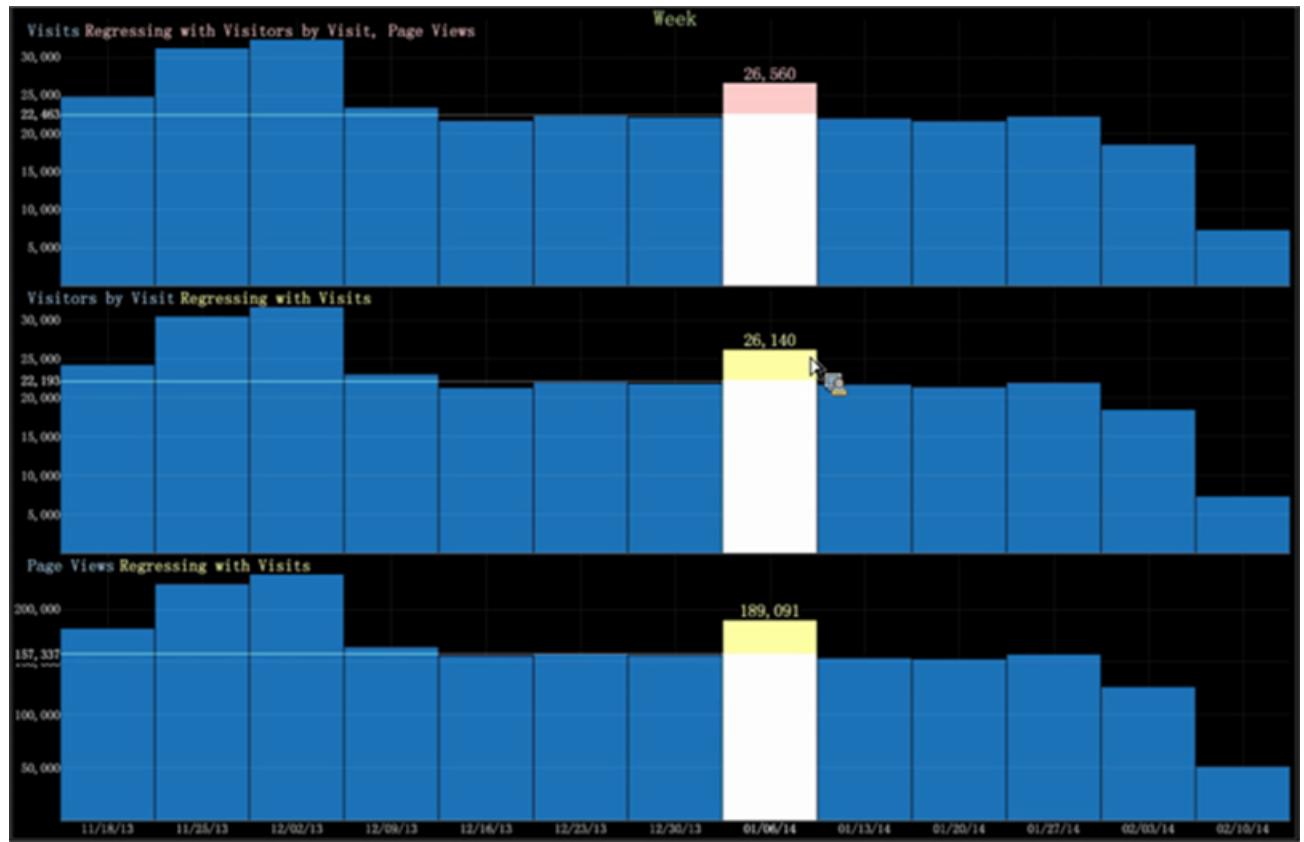


3. View regression by right-clicking on the graph to move the bar up and down.

If you right-click on the graph for a specific value, you can then see the regression ratios for each metric based for upward or downward values.



For example, if my Page Views decrease to 86,041, then the other metrics will have these values: Visits at 12,183 and Visitors by Visit at 12,028.



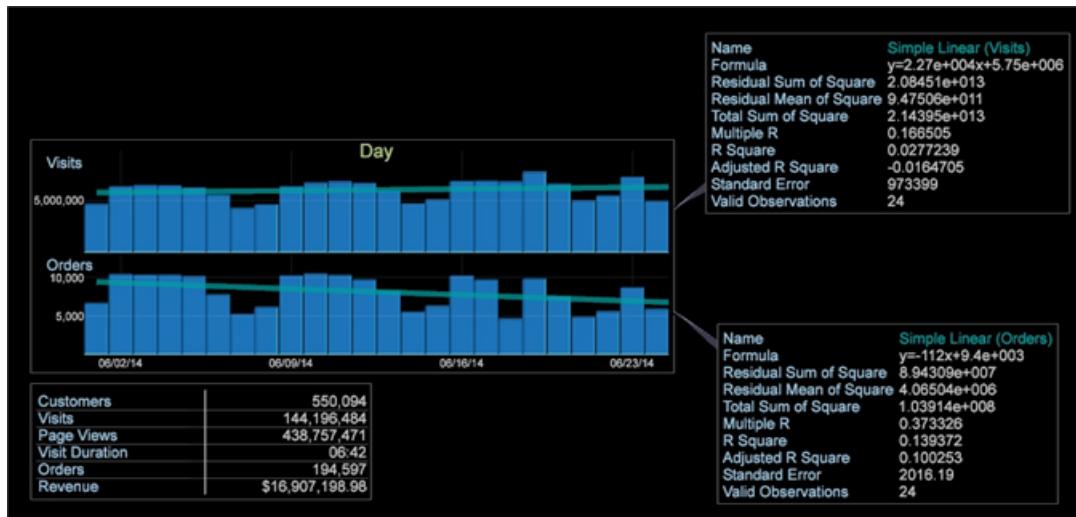
If Visitors by Visits values increase to 26,141, then the other metrics will be Visits at 26,560 and Page Views will be at 189,091.

Trend Lines

Trend lines lets you overlay graphs to compare and interpret data.

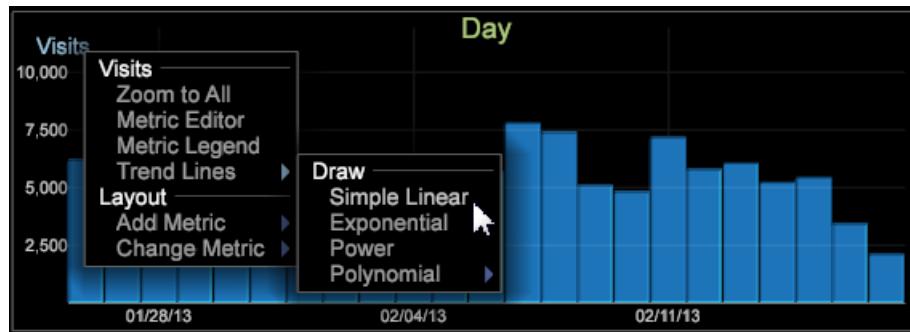
Like the [Scatter Plot](#) visualization, you can now set trend lines on a graph visualization to display the rate of change based on linear, exponential, power, or polynomial lines. The Trend Line feature allows you to overlay trend lines on a graph, most commonly over a Time dimension.

For example, in this graph comparison, we can see that Visits are trending up, but Orders are trending down.



To add a Trend Line

1. Open a graph and right-click the metric name in the upper left corner.
2. Click **Trend Lines** and select from the options.



You can select the trend line to appear over the graph as **Simple Linear**, **Exponential**, **Power**, or **Polynomial**. Polynomial will create a polynomial regression trend line. Simple Linear will create a trend line as the rate of change along the regression line. Exponential calculates a trend line as $y = b * \exp(a * x)$ and Power as $y = b * x^a$.

The trend will be calculated and rendered on the graph, and a callout will open displaying detailed information of the trend equation.

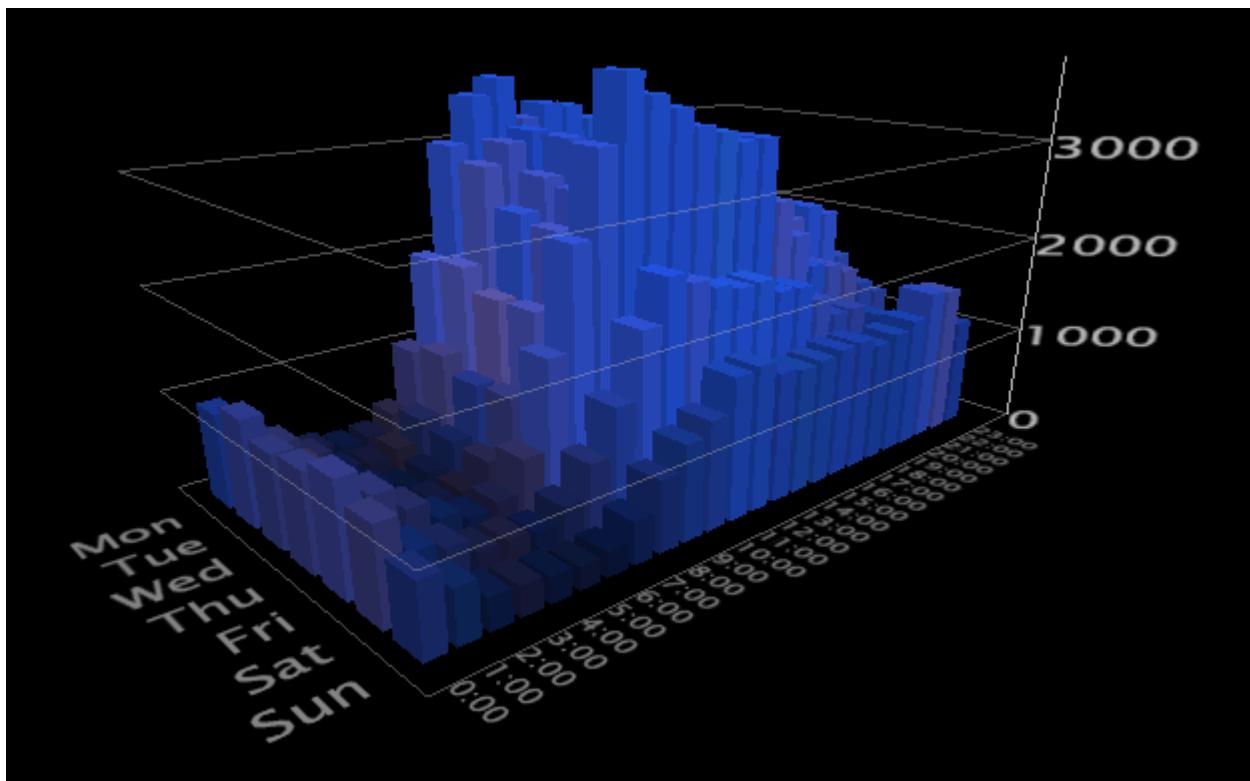


2D graphs

Navigation title:

Two-dimensional (2D) graphs display a metric across two dimensions.

See the following example:



To rotate a graph

- Right-click the graph and drag the mouse in the desired rotational direction.

To zoom in or out on the graph

-
- With the mouse pointer anywhere in the graph window, press both mouse buttons at the same time and push the mouse away from your to zoom in and toward you to zoom out. You also can use your mouse wheel if one is available.

To scale the vertical axis

- Move the mouse over the vertical axis to highlight a plane of the graph.
- While the plane is highlighted, click+drag the mouse up or down to scale the vertical axis of the graph.

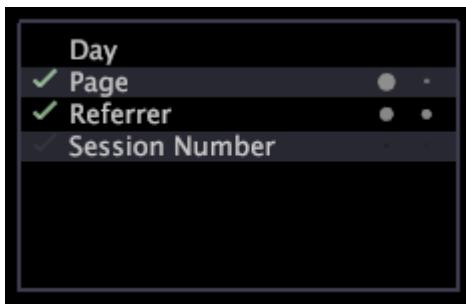
See [Zoom in or out in a visualization](#) on page 292.

Guided analysis

Navigation title:

A guided analysis visualization provides cues for further analysis based on the selections that you make in a workspace.

This visualization helps you identify which dimensions can provide you with more information, that is, those that are most strongly correlated with your selections. The guided analysis visualization in your Adobe application displays the dimensions relevant to your dataset, as in the following **Site** guided analysis visualization.



Note: If a dimension name displays in red, it is not defined in your dataset.

When you make a selection within a workspace, the guided analysis visualization displays check marks to the left and dots to the right of the dimensions to show which ones provides the most relevant information:

- Check marks** identify the dimensions whose values differ from the benchmark in a statistically significant way (that is, the difference between the selection and the benchmark is not due to random chance).
- Dots** indicate the degree by which the selection differs from the benchmark. The first dot represents the U statistic, and the second dot represents the V statistic. See [Statistical measures](#) on page 372. The brighter and larger the dots, the greater the difference, and the more relevant the information for the dimension based on your selection (that is, brighter, larger dots represent more useful information).

Statistical measures

Navigation title:

To help with the statistics, Data Workbench provides three statistical measures in the guided analysis visualization.



Note: Although mathematics can help you make judgments about the correlations in your data, the context surrounding the data also must be taken into account.

- Chi Sq p** is a test of statistical significance that controls the appearance of the check mark in the visualization. Mathematically, it is a probability that we can reject the null hypothesis, which states that the differences observed between the two groups can be explained by random variation. Practically, if the Chi Sq p value is less than almost 100% we can ignore the correlation regardless of its measured strength (as described in the following U statistic and V statistic sections).

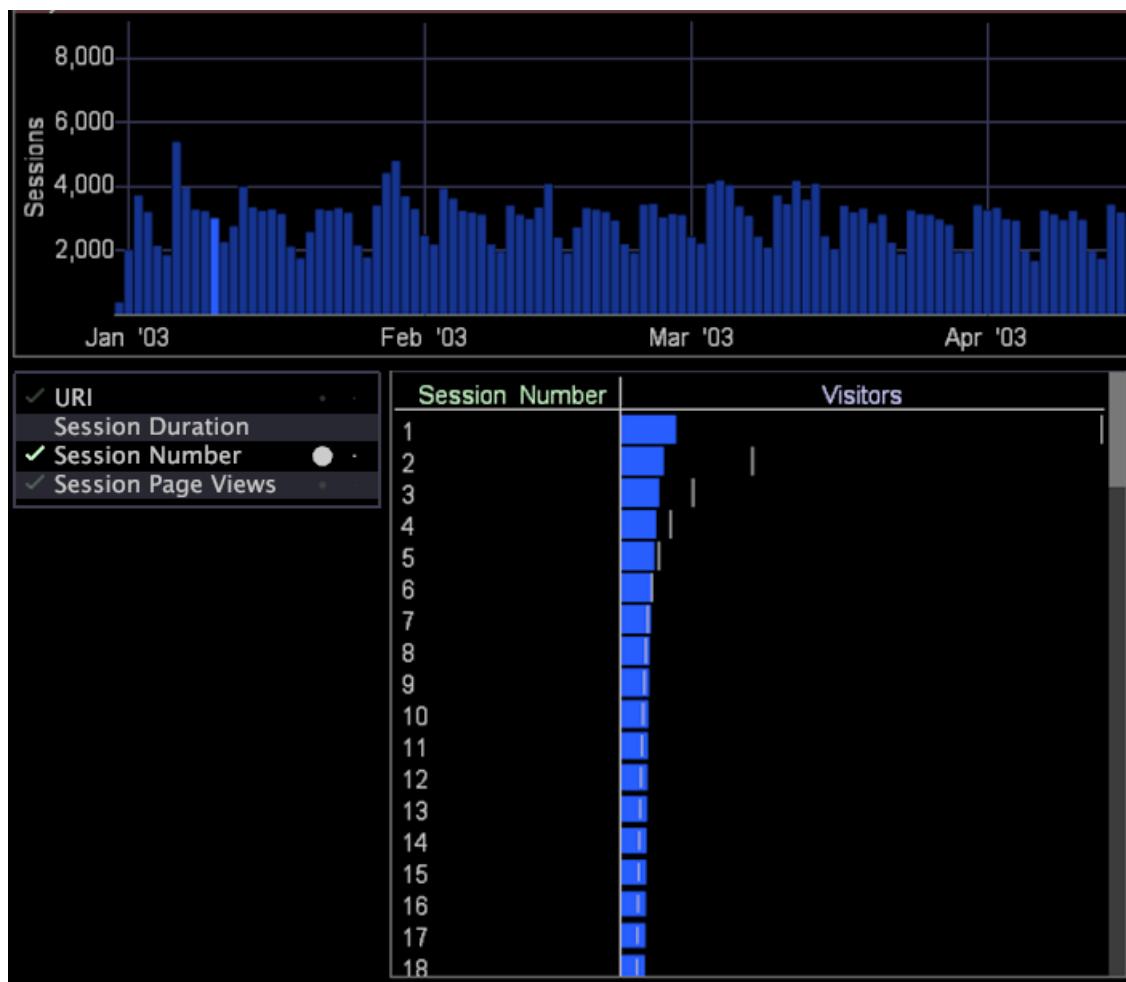
- **U statistic** is a measure of strength of statistical correlation. Mathematically, it comes from a branch of mathematics called information theory and is closely related to the concept of mutual information between the distributions of the two groups. Alternatively, it can be thought of as the compressibility of one group given an optimal coding scheme for the other group. Practically, this measure performs very well in the common case of a dimension with many elements containing few visitors. The measure varies from 0 (weak) to 1 (strong).
- **V statistic** is also a measure of strength of statistical correlation. Mathematically, it is related to the familiar Cramer's V statistic, differing only by a normalization step intended to improve symmetry of the measure with respect to selection inversion. Practically, this measure works reasonably well with many types of dimensions and is related to a commonly-used statistical measure. The measure varies from 0 (weak) to 1 (strong).



Note: The U and V statistics were selected to complement one another—each tuned to detect types of correlations to which the other one might not respond as strongly.

Using this visualization as your guide, you can add other visualizations to your workspace to provide more insight into your data based on the selection.

The following **Site** example contains a bar graph that shows sessions for days in January, February, March, and April. Note that one day in January is selected.



The guided analysis visualization in the lower-left corner of the workspace indicates that the Session Number dimension provides useful information about the selected day.

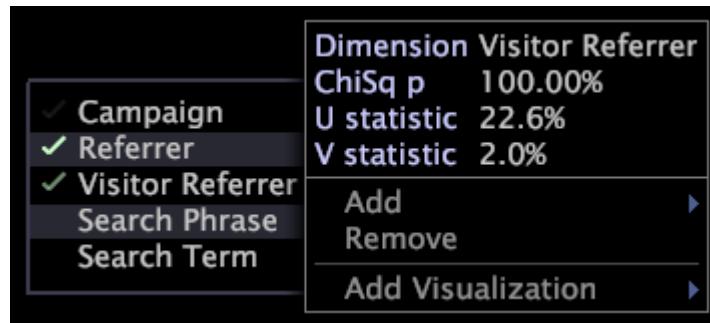
By examining the Session Number bar graph in the lower-right corner of the workspace, you can see that the data for Session Number 2 is much lower than the benchmark. Thus, we can conclude that, as a percentage, fewer second sessions occurred on the selected day than is usual. To view a bar graph for any of the dimensions listed in the guided analysis visualization, simply select the dimension by clicking the dimension with your mouse.

Display the Dimension menu

Navigation title:

Steps to display information or options for a dimension.

1. In the guided analysis visualization, right-click a dimension name.



Note: If a dimension name displays in red, it is not defined in your dataset.

2. View the statistical information about the dimension, which is described earlier in this section.
3. (Optional) Click one of the following options:
 - **Add** to add another dimension.
 - **Remove** to remove the selected dimension.
 - **Add Visualization** to add a visualization for the selected dimension.

Latency Analysis

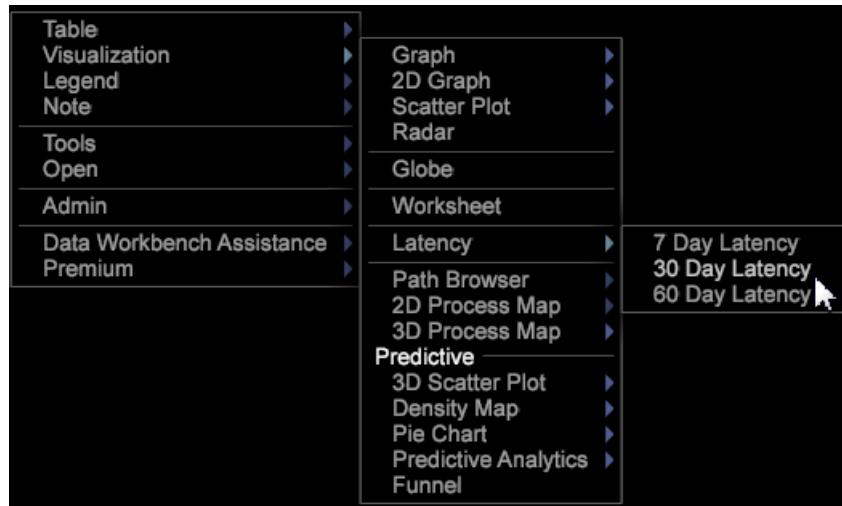
The Latency visualization lets you analyze latent customer behavior within a set number of days before or after an event occurred after a campaign or other type of event or period of time.

The **Latency** visualization allows you to set a metric to identify behavior before, during, or after an event occurred to determine its effect on customer behavior. For example, you can identify the effectiveness of a marketing campaign by viewing the revenue a week after an event occurred. Or you can show customer behavior a week previous to the event as a baseline to see an event's effect on behavior.

Using the Latency visualization, you can change the Time dimension from day, hour, week, or other time period, to visit, click through, hit, or other countable dimension.

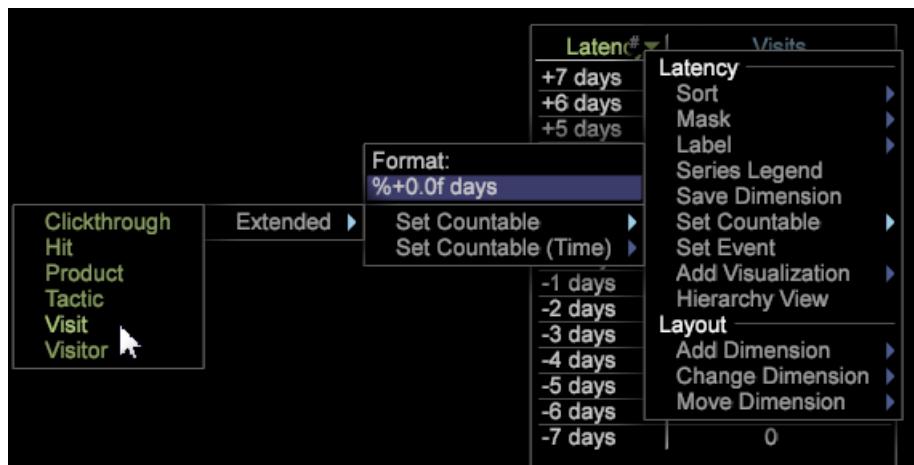
Setting up Latency Analysis

1. Open a workspace and right-click **Visualization > Latency**.
2. From the menu, select the number of days to analyze for latency events.



3. Right-click on the **Latency** label to open menu.

- Using the **Set Countable** option, set a countable dimension such as visit, clickthrough, hit, and others.
- Using the **Set Countable (Time)** option, set a time dimension such as day, hour, week, day of week, and hour of day.



4. Modify the Latency visualization.

Right-click on the **Latency** header to select menu options to change countables, dimensions, to re-label, or to select or change other settings.

Latency tables

Navigation title:

Latency table visualizations are tables that include a latency dimension, which is a type of derived dimension that measures the time that has elapsed since a particular event occurred.

You define the event by making selections within one or more visualizations and setting those selections as the event using the Set Event context menu option. Latency tables are especially useful for tracking activity related to a campaign or to a particular customer order in which you are looking for a time correlation.

In **Site**, latency tables provide information about the visitor sessions that occurred as many as seven days before or after the event, but you can configure latency tables to provide information about different countable and time dimensions. See [Configure a latency table](#) on page 501.

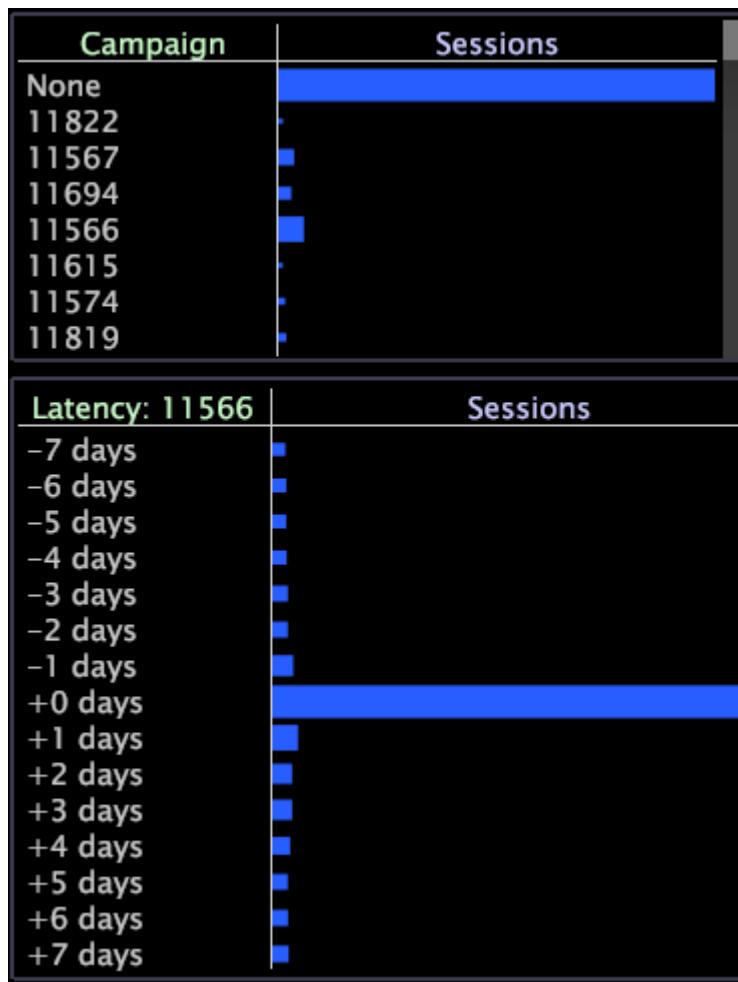
Elements of the parent dimension, such as a session, which are part of the specific event that you selected, have a latency of zero. All other elements are assigned latencies that reflect the distance (in the appropriate time dimension) from the event.

The following example illustrates how you might use the latency table.

Identify value events in relation to a campaign

Let's say you want to track the activity of customers during the seven days before and after they responded to a particular advertising campaign. To view the latency for a particular advertising campaign, you would set the campaign of interest as the event for the latency table.

Latency in the workspace below is based on selection of Campaign 11566 (the sessions in response to this campaign).



A latency of “+0 days” identifies the sessions in response to Campaign 11566, as well as all other sessions for those same customers that occurred on the same day.

A latency of “-2 days” identifies the sessions for those same customers that occurred two days before the customers responded to the campaign.

A latency of “+7 days” identifies the sessions for those same customers that occurred seven days after they responded to the campaign.

In addition to the procedures listed in the following sections, you can perform all of the same tasks that you can perform in a table, such as sort elements, mask elements, add a series legend, export data, and so on. For more information, see [Tables](#) on page 414.

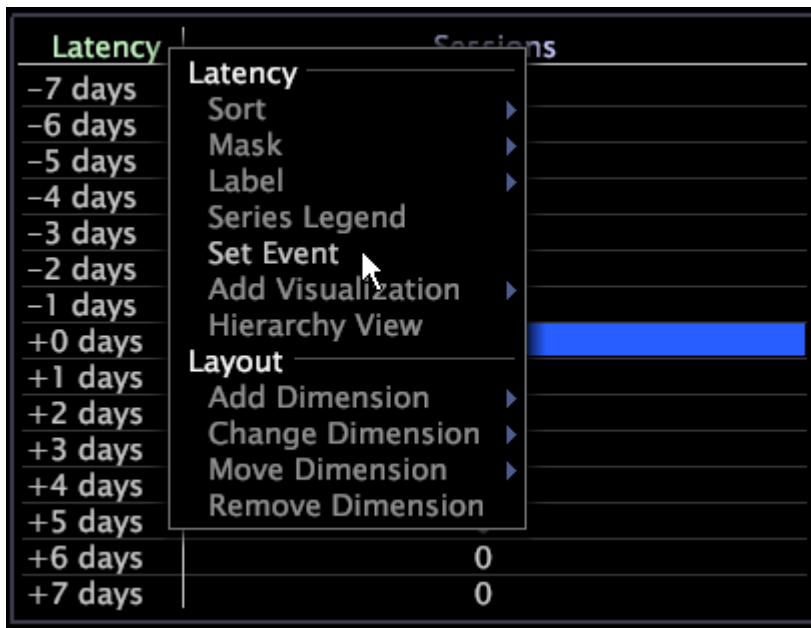
Create a latency table

To create a latency table, you begin by making a selection, then setting that selection as the event for which you want to track latency.

1. Right-click within a workspace and open the desired visualization(s), which must be based on the countable dimension used to configure your latency table.

For example, in **Site** the visualization(s) would need to be sessions-based.

2. Open a blank latency table.
3. Make a selection in your workspace.
4. Right-click within the latency table and click **Set Event**.



Note: Events that you select do not persist unless you save the selections as a latency dimension. For steps, see [Reuse a latency dimension](#) on page 378.

Reuse a latency table

If you want to use the same latency table again, you can save the latency table locally or if you have the proper permissions you can save it to the server for all users of a particular profile to access.

To save the latency table for use in other workspaces

1. Right-click the top border of the visualization and click **Save**. The **Save** window appears. The default save location is the `User\profile name\Work` folder.
2. In the **File name** field, type a descriptive name for the visualization and click **Save**.

To retrieve the saved latency table

1. Right-click within the workspace and click **Open > File**. The **Open Visualization** window appears.
2. Navigate to the latency table that you saved.
3. Select the latency table visualization file (*.vw) and click **Open**.

Reuse a latency dimension

If you want to use the same latency dimension again, you can save the latency dimension locally or if you have the proper permissions you can save it to the server for all users of a particular profile to access.

Any latency dimensions that you create are saved in the profile's Dimensions directory and are available in the **Change Dimension** drop-down list within Data Workbench.

To save the latency dimension for use in other workspaces

1. Right-click the **Latency** column label or one of its elements and click **Save Dimension**. The **Save Dimension As** window appears.
2. Select or create the appropriate sub-directory within the Dimensions directory.
3. In the **File name** field, type a descriptive name for the dimension (for example, `Latency for Campaign 11565.dim`) and click **Save**.

To retrieve the saved latency dimension

1. Right-click within the workspace and click **Open > File**. The **Open Visualization** window appears.
2. Navigate to the latency visualization that you saved in the `User\profile name\Dimensions` folder.
3. Select the latency dimension file (*.dim) and click **Open**.

Export to Microsoft Excel

For information about exporting windows, see [Export window data](#) on page 228.

Export to a TSV file

For information about exporting windows, see [Export window data](#) on page 228.

Legends

Navigation title:Legends

Legends provide explanatory details about the visualizations displayed within a workspace, enabling you to more fully explore the data represented in your visualizations.

For example, color legends enable you to assess performance measures such as conversion or retention rate by providing a color map for the visualization. Similarly, metric legends display the values for standard metrics based on the selections (queries) that you make within the workspace.

Color legends

Navigation title:

Color legends color-code visualizations by metrics such as Conversion and Retention and can be used in almost every workspace.

Linking business metrics to color makes it much easier to spot anomalies, exceptions, and trends. When a color legend is activated within a workspace, right-click within the legend to select the metric that you want to color-code. The visualizations within that workspace are then color-coded based on the metric displayed in the color legend. You can switch between metrics used for coloring all of the visualizations in the current workspace.



Note: If you do not select a metric in the color legend, all visualizations appear blue.

Activate or change a metric

- Right-click the metric name and click the desired metric.



Confidence legends

Navigation title:

Confidence legends help you to determine the likelihood that the numbers you are seeing are due to chance and to understand the possible deviations in the data.

Even if you are not sampling data, you cannot extrapolate the numbers from a specific time period or subset to other time periods or subsets with full confidence. The confidence legend lets you explore the probability that the numbers fall within a particular range.

If you think of real world data as being a big experiment, the real world still involves chance, even when working with exact numbers. For example, knowing the number of people who completed a transaction between 8 AM and 12 PM on one Tuesday does not mean that the exact same number will do so on the following Tuesday.

The following confidence legend displays confidence details about the Conversion metric, while the following table provides more information about what each data point means.

Metric or Formula	Conversion		
Measured value		2.3%	
Standard deviation		0.1%	
The "true" value...	... is at least	2.1%	... and at most
... with 90% probability		2.0%	2.4%
... with 95% probability		1.9%	2.5%
... with 99% probability			2.6%

Confidence Legend Fields

Field	Description
Metric or Formula	The metric name or metric expression for which you want to view confidence information. Any selections that you make in your workspace are reflected in the legend. This example displays details about the Conversion metric. For information about syntax rules for entering an expression, see Query language syntax on page 515.
Measured value	The value of the actual data collected. In this example, the conversion rate for the current selection is 2.3%.
Standard deviation	The standard deviation of the Measured Value. In this example, the standard deviation of the conversion rate for the current selection is 0.1%.
The "true" value	The likelihood that the Measured Value falls within the range listed for each probability. In this example, if this "real-world experiment" were repeated again and again, you could be 90% sure that the Measured Value would be between 2.1% and 2.4%.



Note: When analyzing the results of any calculations, you must take into account the following caveats:

- The numbers are estimates. If you repeated the same calculations with a different dataset you would get a different result. This is known as random variation.
- Extrapolations to higher probabilities depend on an assumption of normality that is not correct for all metrics. Therefore, the values for 99% probability are less reliable than the values for 90% probability.

If you need more exact numbers, you should consult an expert in statistics.

Change metrics or formulas

- In the confidence legend, click in the **Metric or Formula** field and type the desired metric or expression. For expression syntax rules, see [Query language syntax](#) on page 515.

Export to Microsoft Excel

For information about exporting windows, see [Export window data](#) on page 228.

Metric legends

Navigation title:

Metric legends enable you to see baseline metrics related to your profile and statistics related to the dataset (or to the current selection, if one has been made).

Metric legends in your Adobe application display the metrics relevant to your dataset, as in the following **Site** metric legend.

Visitors	22,658
Sessions	36,968
Page Views	115,350
Value	\$163,295
Value per Session	\$4.42
Conversion	2.3%
Retention	38.7%
Session Duration	02:27

Metric legends are actually a special case of a table visualization. As such, the same features available for tables are available for metric legends. See [Tables](#) on page 414.

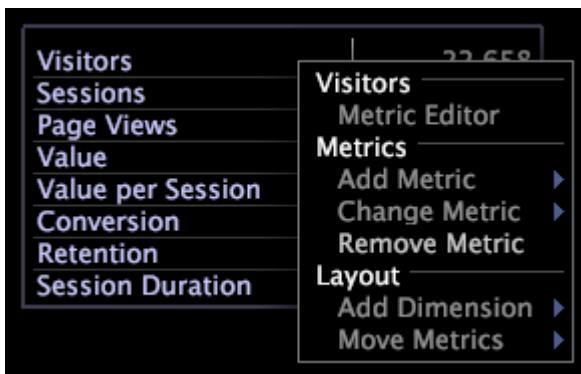
Add a metric

- In the metric legend, right-click a metric name and click **Add Metric > <metric name>**.



Remove a metric

- In the metric legend, right-click the name of the metric that you want to remove and click **Remove**.



Export to Microsoft Excel

For information about exporting windows, see [Export window data](#) on page 228.

Export to a TSV file

For information about exporting windows, see [Export window data](#) on page 228.

Value legends

Navigation title:

A value legend displays defined value events.

The value legend is configured only in the HBX and **Site** applications, but they can be configured for other applications. For more information, contact Adobe Consulting Services.

In HBX and **Site**, a value event is defined as a session that generated business value. For example, the event data records associated with particular page views (for example, an order thank-you page or an application completion page) might represent value events to a business organization.

With value events, you can measure and track the amount of value generated by the website. You can assess business value in dollars for each event and answer such questions as:

- What is the most profitable path through the website?
- What referrer or campaign has generated the most value?

For each event, the legend displays the event's unit value (value per event) and the total value generated by the event. You use the legend to define and change value events and to assign unit values to them.

The following table lists the metrics related to value events.

Metric	Description
Conversion	The percentage of sessions that generated business value
Value	The total business value generated, in dollars
Avg. Value	The average business value generated, in dollars, per session

You easily can define anything that visitors do in the website as a value event: posting a customer service request, completing an application, viewing a piece of content, or completing a purchase. Each value event corresponds to a user accessing a particular page or set of pages on the website, and is associated with a business value in dollars. For example, you might assume that every user reaching the “Thank you for purchasing” page generates a contribution margin of \$20 on average. You would define a value event for that page that has a value of \$20.

Define new value events

To define a new value event in HBX or Site

When you create a value event, you drag website pages that represent value from a visualization to a value legend.

1. Open a value legend.

Event	Value per Event	Value
Total		\$0.00

2. Add value events to the legend from process maps, URI page tables, or page hierarchy views:

- From a process map, drag nodes from the process map to the legend.
- From a URI page table, press Ctrl+Alt and drag a page from the table to the legend.
- From a page hierarchy view, click to the left of a node (folder, page, or group) and drag the it onto the legend.

URI	Sessions	No	Event	Value per Event	Value
/canvas.asp?idPage=hme	19,909	No	/apply/thankyou.appwiz	\$100.00	\$16,600
/loading.asp?idPage=hme_s1_red	12,689	No	Bank Account Application	\$200.00	\$0
/canvas.asp?idPage=pro_evr_t1	2,435	No	Credit Card Application	\$100.00	\$0

Your mouse pointer displays the word "No" until your mouse reaches the legend.

3. In the value legend, assign a business value to each session for which the event occurs:

- a. In the **Value per Event** column, click the cell that corresponds to the page you have added as a value event.
- b. Type the dollar amount to assign for the value of that event and press Enter.

Event	Value per Event	Value
/apply/login.appwiz	\$20.00	\$30,040.00
Total		\$30,040.00

By default, the URL of the page you defined as a value event appears in the value legend. If desired, you can double-click this URL in the legend to enter edit mode and rename the event. You also can edit the value of a particular event at any time. The Data Workbench server automatically recalculates value event-based metrics such as average value and conversion.

After you have defined at least one value event, the Value Segment dimension becomes available for use. This dimension represents the total value that a visitor has generated in all sessions.

Remove value events

- Right-click the desired event and click **Delete Event**.

Event	/apply/thankyou.appwiz: /apply/thankyou.appwiz	Are you sure?
/apply/thankyou.appwiz		0.00
Bank Account Application	Delete Event	Yes 0.00
Credit Card Application	\$100.00	\$0.00
Customer Service Request	\$5.00	\$405.00
General Refer a Friend	\$10.00	\$40.00
Home Equity Loan Application	\$200.00	\$0.00
Mortgage Application	\$250.00	\$146,500.00
Trading Account Application	\$100.00	\$0.00
World Currency Account Application	\$200.00	\$0.00
World Currency Newsletter	\$5.00	\$0.00
Total		\$163,545.00



Note: The Data Workbench server calculates metrics across the entire set of data accessible to the profile you are using. By default, the Data Workbench Server calculates such metrics as Value, Value Events, Average Value, and Conversion across all data in the analysis dataset, even if the data is not from the same logical source.

Export to Microsoft Excel

For information about exporting windows, see [Export window data](#) on page 228.

Path browsers

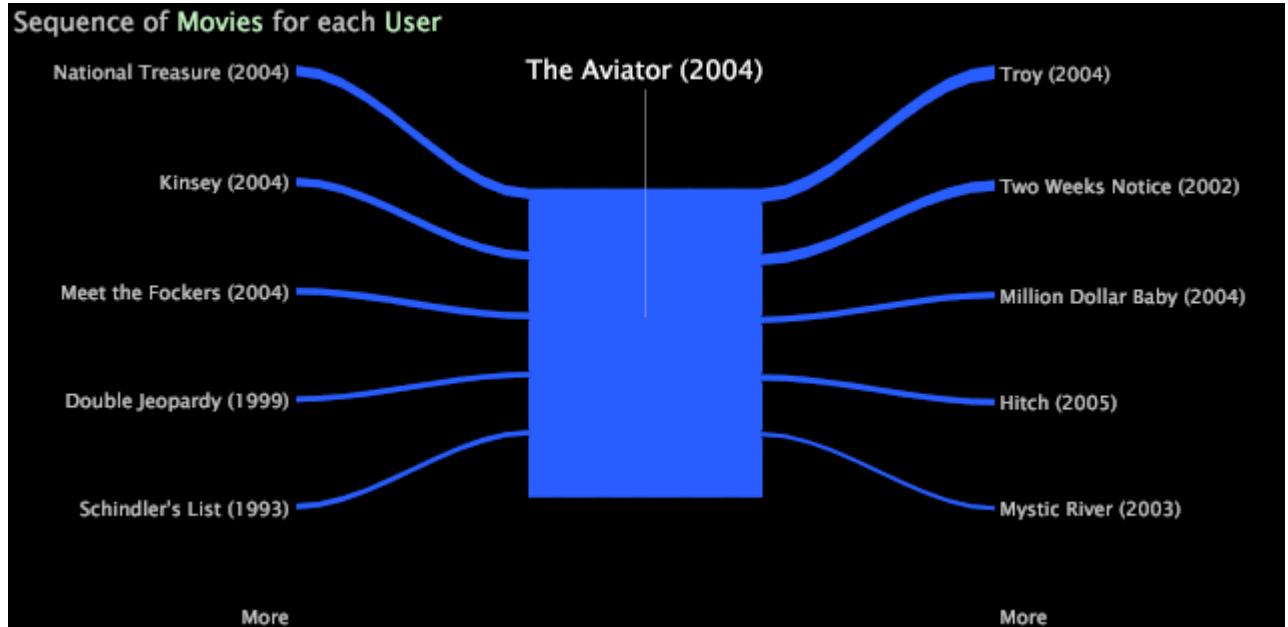
Navigation title:

A path browser enables you to analyze the sequence in which a particular dimension's elements were accessed.

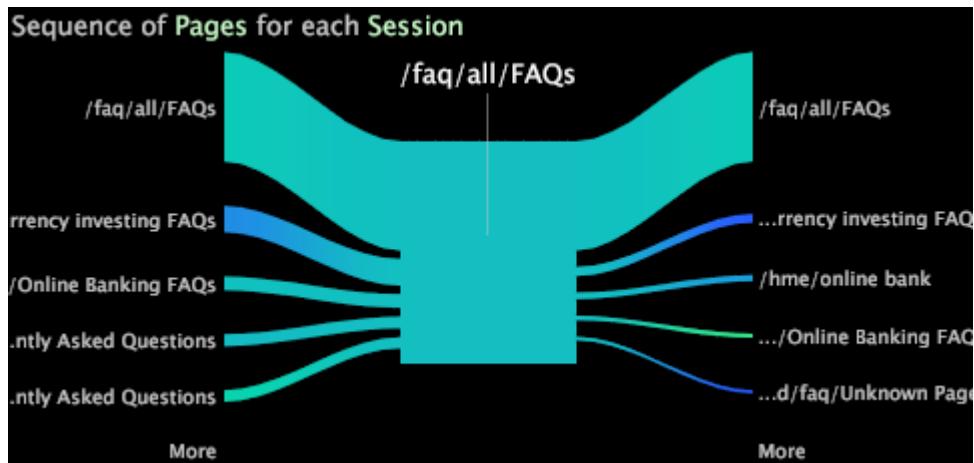
You create a path browser by dragging and dropping an element of a dimension onto a blank path browser visualization. The element that you drag and drop onto the path browser becomes the root of the path browser. The path browser displays paths that pass through the root, enabling you to see the sequence of elements that were accessed before and after the root.

The following path browser shows the sequence of movies that viewers rated before and after rating the movie *The Aviator*, which is the root of the path browser. Each movie name is an element of the Movie dimension,

which is defined in a dataset consisting of movie data that includes movies' names and viewers' ratings of those movies.



You can create path browsers to show the sequence in which elements of any dimension in your dataset were accessed. For example, if you are working with website data, you can create a path browser that shows the sequence of website pages that were accessed before and after the root for each session in which the root was viewed or for each site visitor that viewed the root.



Every path browser has an associated base dimension, group dimension, level dimension, and metric, which provide keys for interpreting the data shown in the path browser.

- **Base dimension:** When you drag and drop a root element onto the path browser, you are dragging and dropping an element of the base dimension. All of the other elements that appear in the paths are elements of the base dimension. You can change the base dimension by dragging and dropping an element of another dimension onto the path browser.
- **Level dimension:** Every dimension in your dataset has an associated level dimension (also referred to as a parent). The level dimension for your path browser should be the same as the level dimension (or parent) for your path browser's base dimension. A path browser's level dimension is important for two main reasons:

-
- As you follow a path from one base dimension element to the next, you move from one level dimension element to the next. For example, suppose that you have created a path browser showing pages of a website. Each page is an element of the Page dimension, and the level dimension for Page is Page View. As you move from one page to the next, you are moving from one page view to the next.
 - When you select a path of base dimension elements within a path browser, you are selecting data for the corresponding elements of the level dimension. The selection always includes the elements of the level dimension that relate to the root, and it is refined by adding more elements to the path. For example, when you select a path of pages, such as root > A > B, you are selecting data for the page views associated with the root where the next page is A and the next page is B.

For information about selecting a path in a path browser, see [Select a path](#) on page 389. For information about selections, see [Make selections in visualizations](#) on page 281.



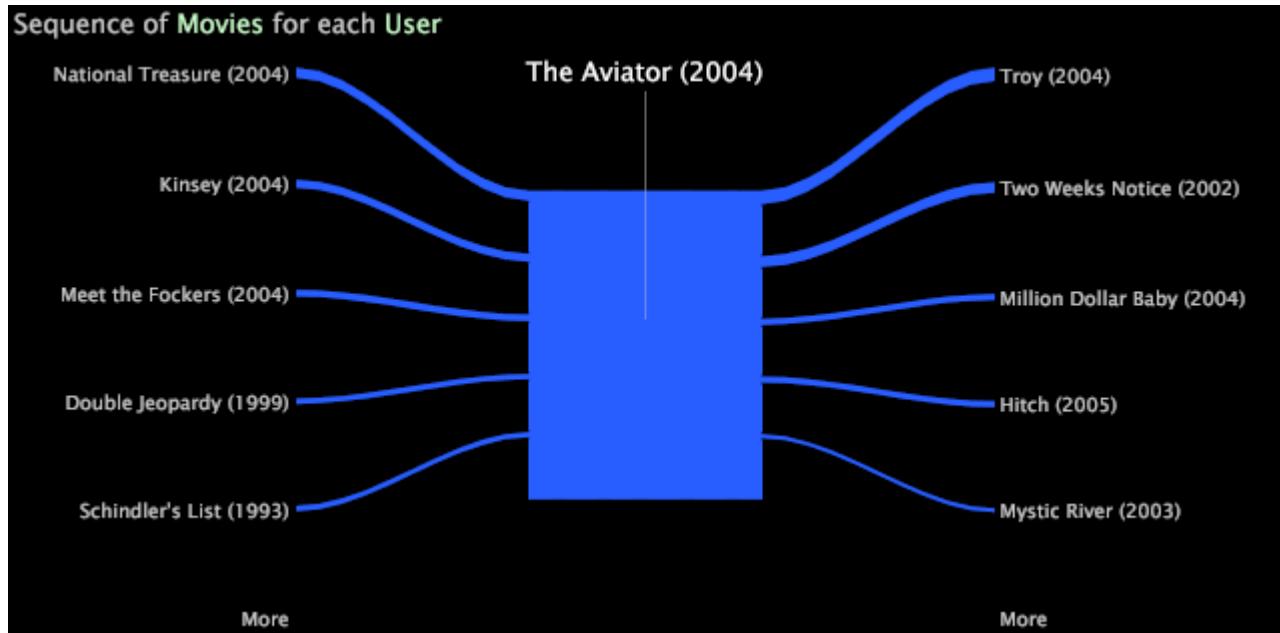
Note: The path browser ignores the elements of the level dimension with no associated base dimension elements. This situation can occur when you create a path browser from a process map. See [Creating path browsers](#) on page 386.

- **Group dimension:** The group dimension determines how the elements of the level dimension are grouped to form the paths of a path browser. Specifically, the level dimension elements associated with a single path in a path browser cannot span more than one element of a group dimension.

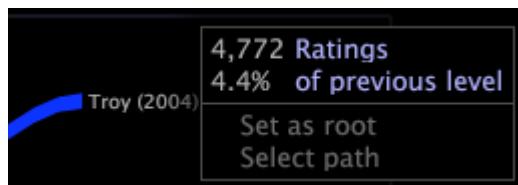
To understand this, consider an example using web data. Suppose that the base, level, and group dimensions for your path browser are Page, Page View, and Session, respectively. One path in your path browser shows the page sequence A > B > C. The group dimension (Session) tells you that the page views (elements of the Page View dimension) associated with the page sequence A > B > C occurred during any single session. It is important to note that there might be multiple sessions during which there were page views for the page sequence A > B > C. Therefore, the path showing the page sequence A > B > C represents all of the individual sessions in which the page views for that sequence occurred.

- **Metric:** The thickness of the path leading to a given element is proportional to the metric's value for that element. Thicker paths indicate greater metric values than thinner paths.

The label in the top left corner of the path browser names the base and group dimensions represented in the visualization. The name of the level dimension is not visible in the path browser visualization. The label takes the form “Sequence of *base dimension name*+s for each *group dimension name*. For example, the label Sequence of Movies for each User tells you that the base dimension is Movie and the group dimension is User.



By right-clicking any element in the path browser, you can see the name of the metric associated with the path browser and its value for that element.



Note: You can change the default dimensions and metric for a path browser. For instructions to configure a path browser visualization, see [Configure a path browser](#) on page 505.

Creating path browsers

Navigation title:

You can create a path browser from a graph, table, or process map.

To create a path browser from a graph or table

1. Add a path browser to your workspace. The visualization is blank except for the label (Sequence of...) in the top left corner.
2. Open a graph or table showing the dimension whose elements you want displayed on the path browser. For example, if you are working with website data, open a page graph or table and identify the page that you want to set as the root.
3. Press Ctrl+Alt and drag the desired element to the path browser. The label in the top left corner of the path browser changes to reflect the base dimension whose element you drag to the path browser.



Note: Dragging an element to a path browser may change the base dimension associated with the path browser, but it does not change the level dimension, group dimension, or metric. Therefore, you must exercise caution in choosing a base dimension that makes sense when used with the path browser's level dimension, group dimension, and metric. To change the level dimension, group dimension, or metric, you must edit the path browser's *.vw file in a text editor such as Notepad. See [Configure a path browser](#) on page 505.

To create a path browser from a process map



Note: A path browser created from a process map displays only the elements shown on the process map. Other elements of the base dimension are not displayed.

1. Create a process map. See [Create a process map](#) on page 395.
2. Drag the desired element from the process map to the path browser. The element becomes the root of the path browser.



Note: When you create a path browser from a process map, the path browser ignores the elements of the level dimension with no associated base dimension elements. When you drag a node from a process map onto a path browser, the path browser's base dimension (named Map) is defined by the process map, and its elements are limited to the elements on the process map. As a result, some elements of the path browser's level dimension do not have associated base dimension elements and are not represented in the path browser.

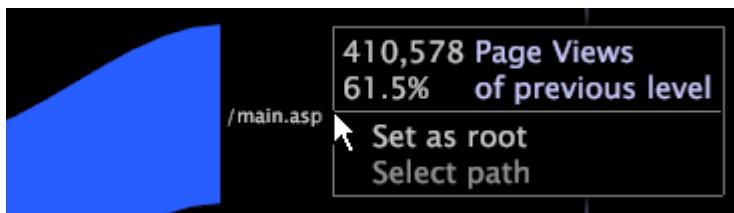
Display metric information

Navigation title:

Information on displaying metric information in a path browser.

Right-click an element name to view the number and percentage of the path browser's associated metric in the top section of the menu.

For example, if you are working with website data, right-click a page name to view the number and percentage of Page Views in the top section of the menu.

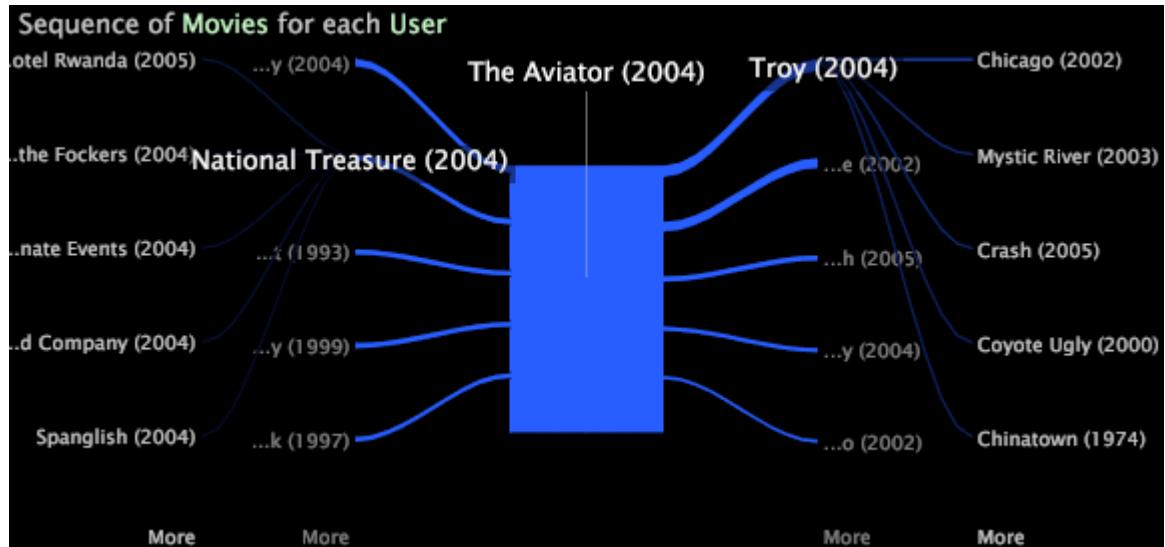


Extend a paths

Navigation title:

Information about extending a path in a path browser.

Click an element in the path browser to extend the displayed path to the left or right of the root.



Display the start and end of paths

Navigation title:

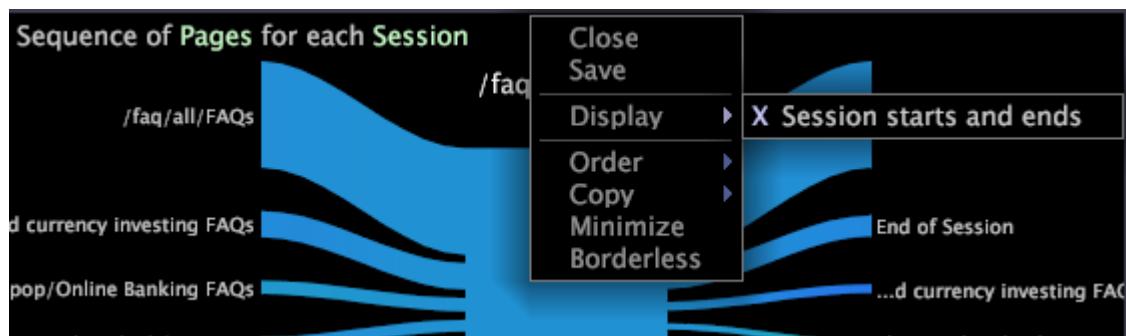
You can display the starts and ends of paths in the path browser.

For example, if your path browser shows the sequence of pages for each session, you might see Start of Session or End of Session nodes among the pages. The Start of Session node identifies the first page viewed during a particular session, which indicates that there were no previous page views in that session. The End of Session node identifies the last page viewed in a particular session, which indicates that there were no subsequent page views in that session.

To enable or disable the display of starts and ends

- Right-click the top border of the path browser and click **Display > <Group dimension name>starts and ends**. When enabled, an X appears to the left of *Group dimension name* starts and ends.

For example, if your path browser shows the sequence of pages for each session, you would choose to display Session starts and ends.



Note: You cannot set a start or end node as the root of a path browser or select a start or end node.

Change the path browser's root

Navigation title:

You can change a path browser's root by either designating a displayed element as the root or adding a new element to the visualization.



Note: You cannot set a start or end node as the root of a path browser.

To set the root of a path browser

- Right-click the desired element and click **Set as root**.

To add a new element to the path browser

1. Open a graph or table showing the dimension whose elements you want displayed on the path browser.

For example, if you are working with website data, open a page graph or table and identify the page that you want to set as the root.

2. Press Ctrl+Alt and drag the desired element to the root position on the path browser.



Note: You can change the base dimension associated with a path browser by dragging an element of the desired dimension to the path browser. This element becomes the new root of the path browser, and the label in the top left corner of the path browser changes to reflect the dimension for this element.

For example, suppose that you create a page path browser showing the Sequence of Pages for each Session. If you were to drag an element of the Search Term dimension to the path browser, the search term element would become the new root, and the label would now show the Sequence of Search Terms for each Session.



Note: Dragging an element to a path browser may change the base dimension associated with the path browser, but it does not change the level dimension, group dimension, or metric. Therefore, you must exercise caution in choosing a base dimension that makes sense when used with the path browser's level dimension, group dimension, and metric. To change the level dimension, group dimension, or metric, you must edit the path browser's *.vw file in a text editor such as Notepad. See [Configure a path browser](#) on page 505.

Select a path

Navigation title:

You can select a path within a path browser to create filters that include data associated with the elements in the path.

When you select a path of base dimension elements within a path browser, you are selecting data for the corresponding elements of the level dimension.

For example, suppose that you have created a path browser showing pages of a website. Each page is an element of the Page dimension, and the level dimension for Page is Page View. When you select a path of pages in a path browser, you are selecting data for the page views associated with those pages.



Note: You can change the default level dimension for a path browser. For instructions to configure a path browser, see [Configure a path browser](#) on page 505.

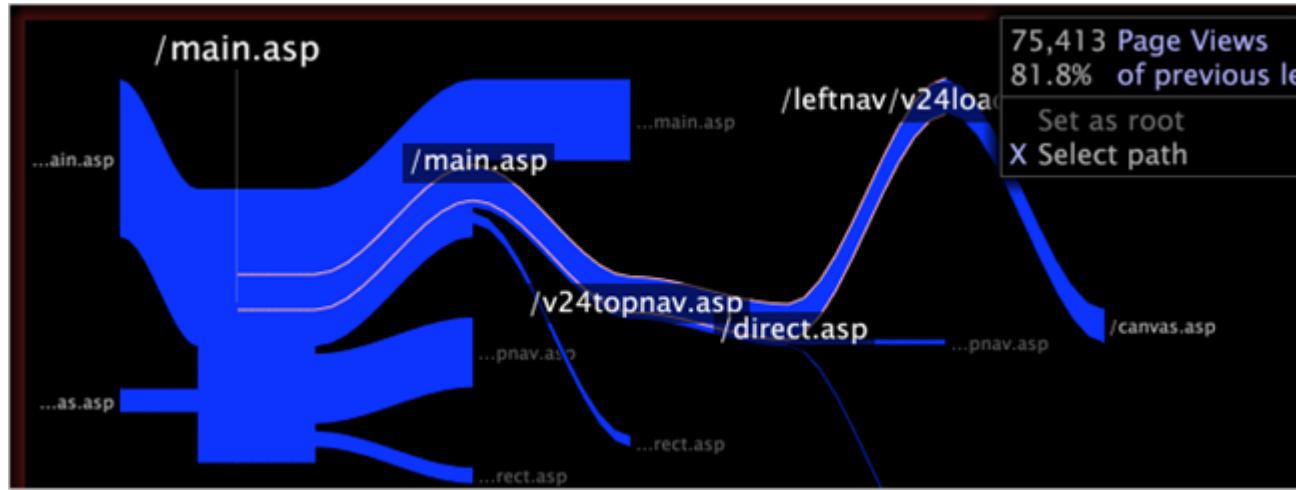
1. Click an element in the path browser to extend the displayed path to the left or right of the root.
2. Right-click the desired element and click **Select path**. The selected path is outlined in white.



Note: You cannot select a start or end node.

3. Repeat Step 1 for each element that you wish to add to the path.

For example, if you are working with website data, you could select a path of pages on your website.



This path constitutes a selection, and all other open visualizations in the workspace (including legends) update to display data associated with the path created by the selected elements. See [Make selections in visualizations](#) on page 281.

[View more elements in paths](#)

Navigation title:

You can choose to view more elements in the path.

In most cases, your path browser is not large enough to show all of the elements in the sequence.

- Click **More** in the lower left portion of the visualization to view the previous elements in the path.
- Click **More** in the lower right portion of the visualization to view the next elements in the path.

By clicking More at the lower left or lower right portion of the visualization, you open a table that lists the previous or next elements in the path. The table also shows the value of the associated metric for each element.

More Movies on path	Ratings
The Green Mile (1999)	
American Pie (1999)	
Swordfish (2001)	
The Good Girl (2002)	
The Day After Tomorrow (2004)	
Ransom (1996)	
Lemony Snicket's A Series of Unfortu	
Reservoir Dogs (1992)	
GoodFellas: Special Edition (1990)	
The Bourne Identity (2002)	
Entrapment (1999)	
The Usual Suspects (1995)	
Kiss the Girls (1997)	
Pearl Harbor (2001)	
Miss Congeniality (2000)	
The Italian Job (2003)	

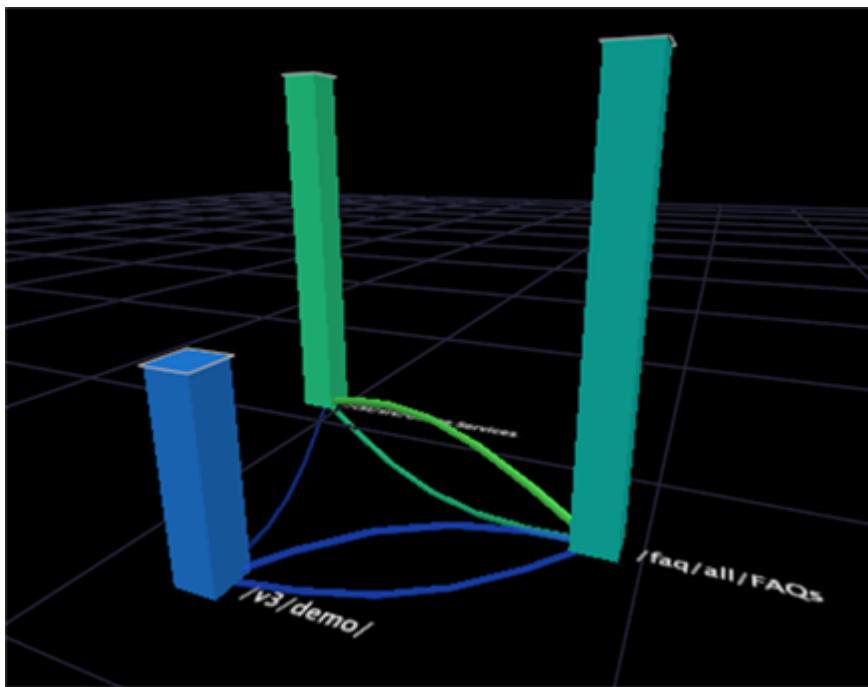
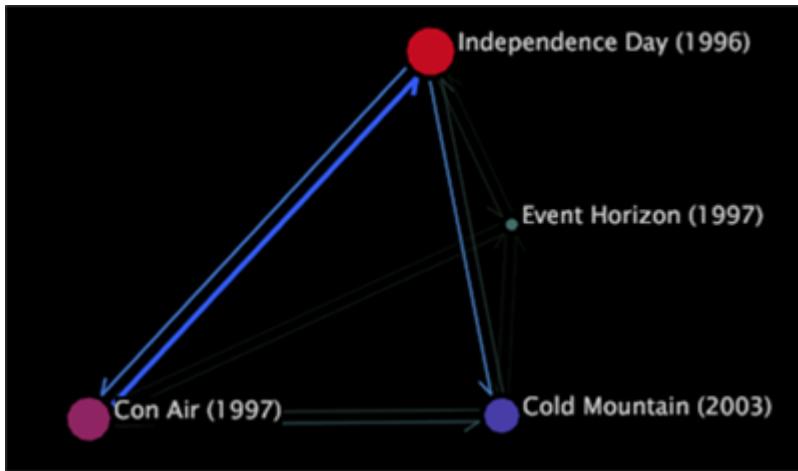
[Pie Chart](#)

[Process map](#)

Navigation title:

Process maps enable you to analyze the flow of activity between elements of a dimension.

You create process maps by dragging and dropping the elements of a dimension onto a blank two-dimensional (2D) or three-dimensional (3D) map. The elements become nodes on the map. Nodes are circles in 2D process maps and bars in 3D process maps.



 **Note:** A process map gets its name from its use in analyzing the flow of activity between the steps in a process. In this type of analysis, each element on the map represents a step in the process.

Unlike path browsers, process maps can show as few or as many elements as necessary for your analysis. You choose the elements of interest and drag and drop them onto the map. Also unlike path browsers, process maps depict the flow of activity in both directions between one element and one or more other elements.

 **Note:** For these maps to work most effectively, you should open a color legend in the workspace. For information about using color legends with process maps, see [Activate a color link](#) on page 400. For more information about color legends, see [Color legends](#) on page 378.

Every process map has an associated base dimension, group dimension, level dimension, and metric, which provide keys for interpreting the data shown in the process map.

The default settings for a process map's dimensions and metric depend on the Data Workbench application that you are using. For information about the dimensions and metrics available to you for your process maps, see the application guide for your Data Workbench application.

- **Base dimension:** When you drag and drop an element onto a process map, you are dragging and dropping an element of the base dimension.
- **Level dimension:** Every dimension in your dataset has an associated level dimension (also referred to as a parent). The level dimension for your process map should be the same as the level dimension (or parent) for your process map's base dimension. For example, if you drag a page (an element of the Page dimension) to the map, the corresponding level dimension would be Page View.
- **Group dimension:** The group dimension determines how the elements of the level dimension are grouped to form the connections between nodes. For process maps, the group dimension is important for three main reasons:
 - A connection between two nodes cannot span more than one element of a group dimension. To understand this, consider an example using web data. Suppose that the base, level, and group dimensions for your process map are Page, Page View, and Session, respectively. A connection from page A to page B tells you that, during any single session, a page view of page A occurred prior to a page view of page B with no intervening page views of other pages (nodes) on the map. Note that page views of other pages of the site might have occurred between page views for pages A and B during the same session, but these pages are not shown on this map.
 - A connection between two nodes can represent multiple elements of the group dimension. For example, there might be multiple sessions in which a page view of page A occurred prior to a page view of page B. Therefore, the connection between page A and page B represents all of the individual sessions in which a page view of page A occurred prior to a page view of page B with no intervening page views of other pages (nodes) on the map.
 - When you make a selection based on a node within a process map, you are selecting all of the elements of the group dimension that involved that node. See [Make selections in visualizations](#) on page 281. For information about selections, see [Make selections in visualizations](#) on page 281.
- **Metric:** The size of the node for a given element is proportional to the metric's value for that element. Larger nodes indicate greater metric values than smaller nodes.

For example, if you are using the Site or HBX application, you can drag, by default, elements of the Page dimension onto the process map. The size of each node is related to the quantity of sessions (defined by the Sessions metric) in which that page was viewed.



Note: You can change the default dimensions or metric for a process map. For steps to configure a process map, see [Configure a process map](#) on page 506.

Types of process maps

Navigation title:

Information about the different types of process maps.

2D process maps

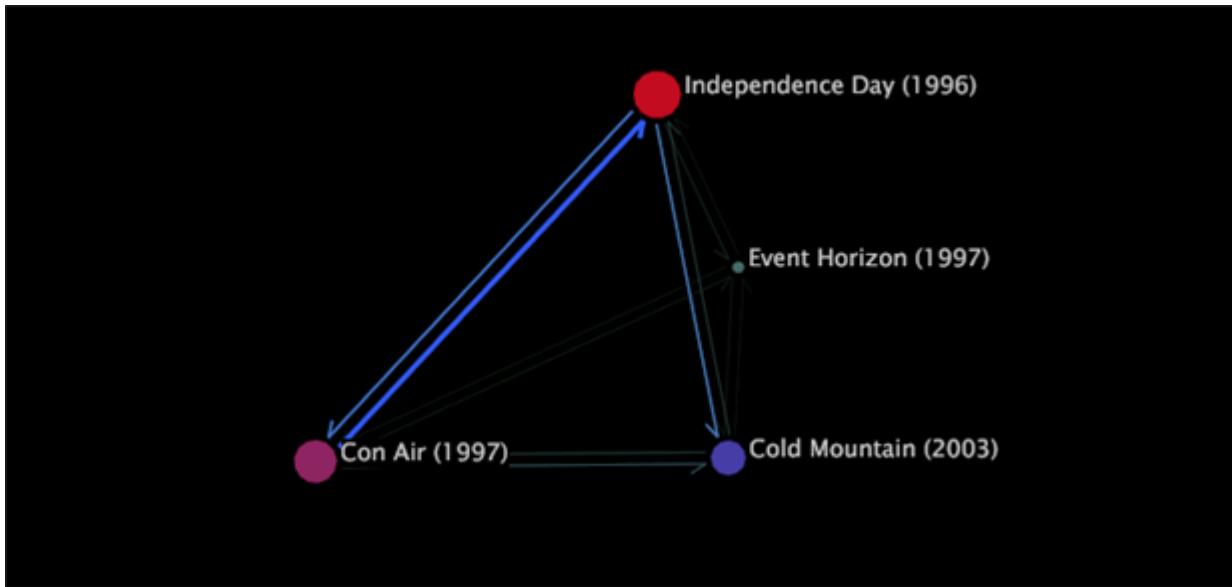
Two-dimensional process maps provide a two-dimensional view of the activity between dimension elements. The size of a node in a 2D process map is proportional to the value of the metric associated with that node. In addition, both the thickness and intensity of an arrow between two nodes are proportional to the average of the metric's values for those nodes.

Within a 2D process map, you can do any of the following tasks:

- Select, move, remove, and label nodes
- Make selections

- Save dimensions
- Create other visualizations
- Activate color links
- Display metric quantities
- Add callouts

The 2D process map in the following example shows nodes corresponding to the names of movies. Each movie name is an element of the Movie dimension, which is defined in a dataset consisting of movie data. The Movie dimension is the base dimension for this process map.



In the example, the size of each node and the thickness and intensity of each arrow are proportional to the Ratings metric, which is a count of ratings that a movie received. Therefore, a movie with a large node, such as *Independence Day*, has more ratings than a movie with a small node, such as *Event Horizon*. You also can see that more movie viewers rated *Independence Day* before *Cold Mountain* than rated the same movies in the opposite order. Note that the arrows do not indicate that viewers rated *Independence Day* and then rated *Cold Mountain* immediately afterward, or vice versa. Viewers might have rated other movies in between, but these movies are not shown on this map.

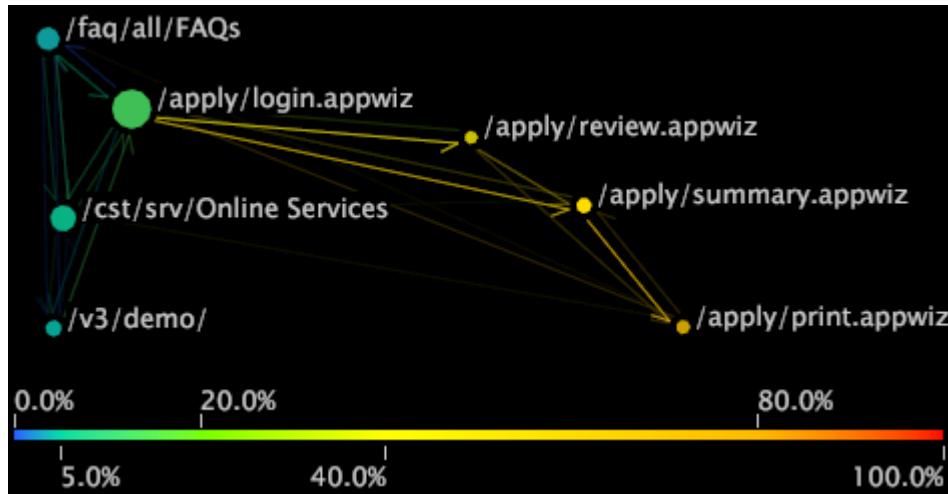
2D metric maps

Two-dimensional metric maps are a type of 2D process map that position nodes based on the value of a particular metric. In many cases, the metric used with the 2D metric map is either Conversion or Retention. Conversion and retention maps help you understand what steps in the processes of your customer-facing channels influence customer conversion and retention.

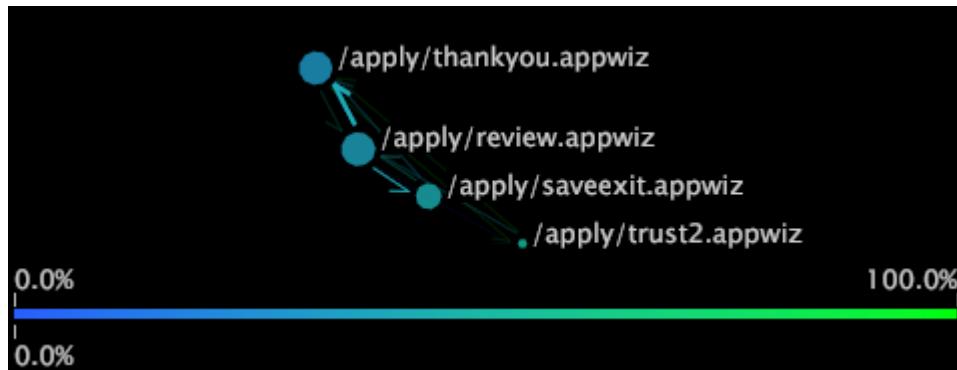


Note: The metric that you use with a 2D metric map must be expressed as a percentage.

In a conversion metric map, nodes with 0 percent conversion are plotted at the left of the graph, and pages with 100 percent conversion are plotted at the right. Activity between nodes is displayed, making it easy to see which steps in a process lead to increased or decreased conversion and which steps drive abandonment. A process-conversion analysis is an effective way to compare processes or compare different implementations of the same process.



Similarly, retention maps show elements with 0 percent retention at the left of the graph and elements with 100 percent retention at the right. You can see the retention rate for each node on the map, which helps you determine which elements influence customers to return.



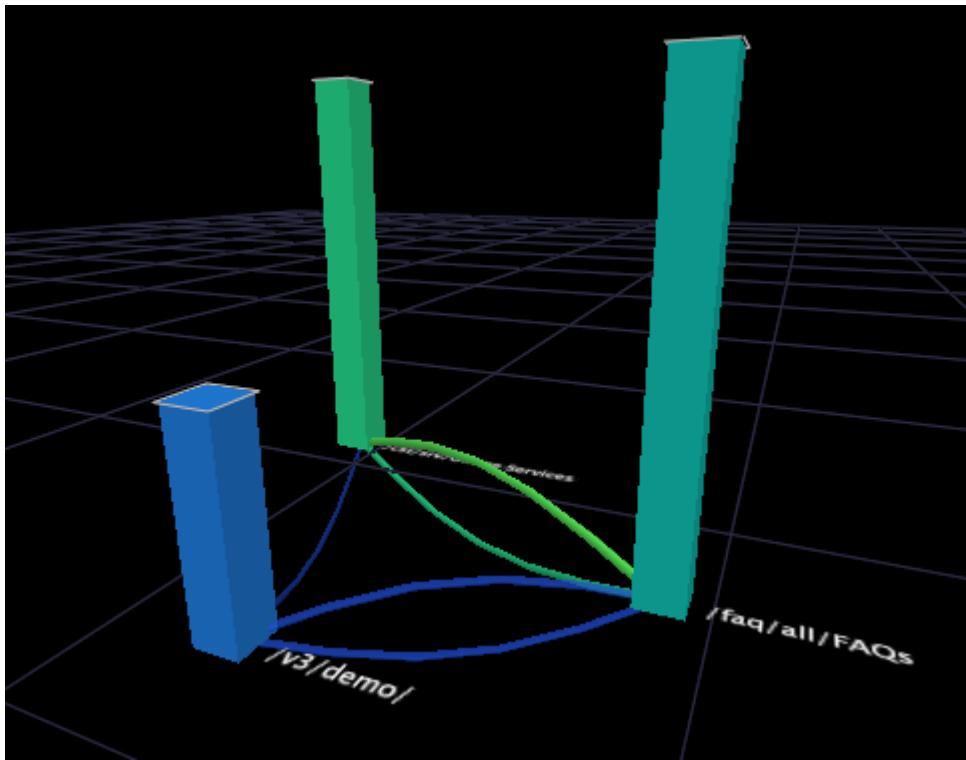
Note: You cannot move nodes on 2D metric maps horizontally. Metric maps are designed to position nodes left to right based on their metric values.

3D process maps

Three-dimensional process maps provide a three-dimensional view of the activity between dimension elements. The height of a bar in a 3D process map is proportional to the value of the metric associated with that node. As with 2D process maps, both the thickness and intensity of the connectors between two nodes are proportional to the average of the metric's values for those nodes. Within a 3D process map, you can do any of the following tasks:

- Select, move, remove, and label nodes
- Make selections
- Save dimensions
- Create other visualizations
- Activate color links

The 3D process map in the following example shows nodes corresponding to the pages of a website. Each page is an element of the Page dimension, which is defined in a dataset consisting of web traffic data. The Page dimension is the base dimension for this process map.



In the example, the height of each bar and the thickness and intensity of each connector are proportional to the Sessions metric, a count of sessions in which the pages were viewed. Therefore, a page with tall bar, such as /faq/all/FAQs, was viewed during more sessions than a page with a short bar, such as /vs/demo. Note that the connections between two pages do not indicate that one page was viewed immediately before or after another during a given session. Other pages might have been viewed during the same session, but these pages are not shown on this map.

Create a process map

Navigation title:

You can create 2D and 3D process maps by dragging and dropping elements from bar graphs, tables, and hierarchy views onto a blank map.

The elements that you can add must be elements of the process map's base dimension. You also can drag and drop nodes from one process map to another as long as the maps are using the same base dimension. In addition, the entire map can be zoomed or moved to focus on a particular node, or it can be changed to other visualization types. See [Zoom in or out in a visualization](#) on page 292.

To add elements to a process map using a table or bar graph

- From any table or bar graph with the same base dimension as the process map, press Ctrl+Alt while you click and drag individual elements to the process map. Your mouse cursor displays the word "No" until your mouse reaches the process map.



Note: The elements that you can add must be elements of the process map's base dimension.



To add elements to a process map using a hierarchy view



Note: Adobe recommends that you add nodes from the highest level of the hierarchy that you are analyzing.

1. From any table or bar graph with the same base dimension as the process map, right-click an element or the label of the base dimension and click **Hierarchy View**.
2. Press Ctrl+Alt while you click and drag elements to the process map. Your mouse cursor displays the word "No" until your mouse reaches the map.



Note: The elements that you can add must be elements of the process map's base dimension.

Dragging a single element onto a process map makes a map node for only that element, but if you select multiple elements (a group) or a folder containing multiple elements, dragging from the hierarchy creates a single node for that group or folder. For example, if you are working with website data, dragging a folder named `site.com/cgi-bin` onto a map makes a node called `site.com/cgi-bin/*`, which represents all pages and directories that are children of that folder.

For more information about page hierarchy views, see [Apply hierarchy views](#) on page 422.

To add nodes to a process map from another process map



Note: The process maps must have the same base dimension.

- Copy a node from the first to the second process map using the following methods:
 - To copy individual nodes, click and drag each node to the second process map.
 - To copy multiple nodes, Ctrl+click and drag to create a box around the nodes you want to copy, then click and drag the highlighted nodes to the second process map. All of the highlighted nodes are copied to the second process map.

Move and rotate a process map

Navigation title:

Information about moving a process map within a window and how to rotate a 3D process map.

To move a process map within a window

- Click and drag the map to the desired location.

To rotate a 3D process map

- Right-click and drag the map while moving the mouse in the desired rotational direction.

Select a node

Navigation title:

Selecting a node by clicking it dims the other nodes so that you can more easily see the arrows to and from the highlighted node.

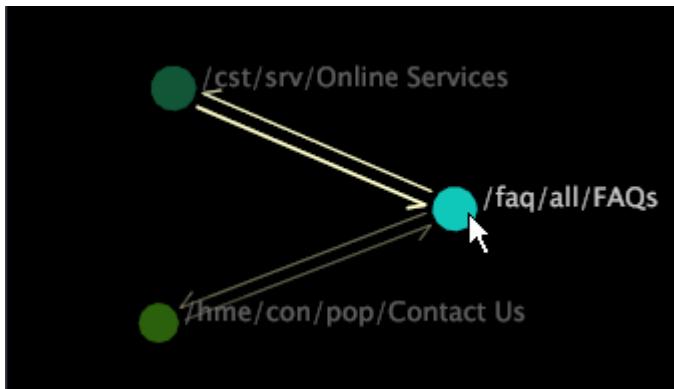
It also enables you to move the selected nodes as a group within the visualization.



Note: Selecting a node by clicking it does not filter the data in any way.

To select a single node

- Click the node to select it and highlight the arrows pointing to and away from the node.



To add another node or group of nodes to your selection

- Ctrl+click a node or Ctrl+drag across multiple nodes.

To remove a node or group of nodes from your selection

- Shift+click a node or Shift+drag across several nodes.

Move a node

Navigation title:

You can move nodes within a map or between process maps as long as the maps have the same base dimension.

For example, you can drag a node from a 3D process map to a 2D process map.

- Click and drag the node to the desired location in the same process map or in another process map in your workspace.

Remove a node

Navigation title:

Information about removing a node from a process map.

- Right-click the node and click **Remove<node name>from map**.

Label a node

Navigation title:

Information about displaying, hiding, and changing a node's label.

To display a node's label

- Right-click the node and click **Label<node name>**. The node's label displays. This menu option is available only if the node is not labeled.

To hide a node's label

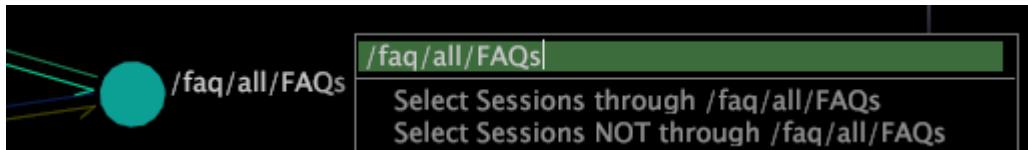
-
- Right-click the node and click **Unlabel<node name>**. The node's label is removed. This menu option is available only if the node is labeled already.

To change a node's label



Note: Changing a node's label affects only the label that is shown in the process map. It does not affect the name of the element that the node represents.

1. Right-click the node. The label of the node is shown at the top of the node menu.
2. Click the node label to enter edit mode and change the label of the node as desired.



Make a selection from a process map

Navigation title:

You can make selections within process maps to create filters that include or exclude data associated with a particular node.

Making a selection within a process map involves the map's group dimension, which determines how the elements of the base dimension (that is, the nodes in your map) are grouped to form the connections between nodes.



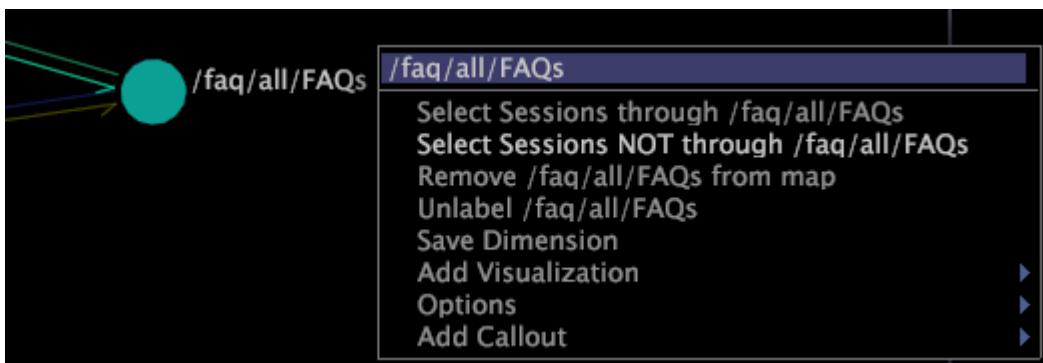
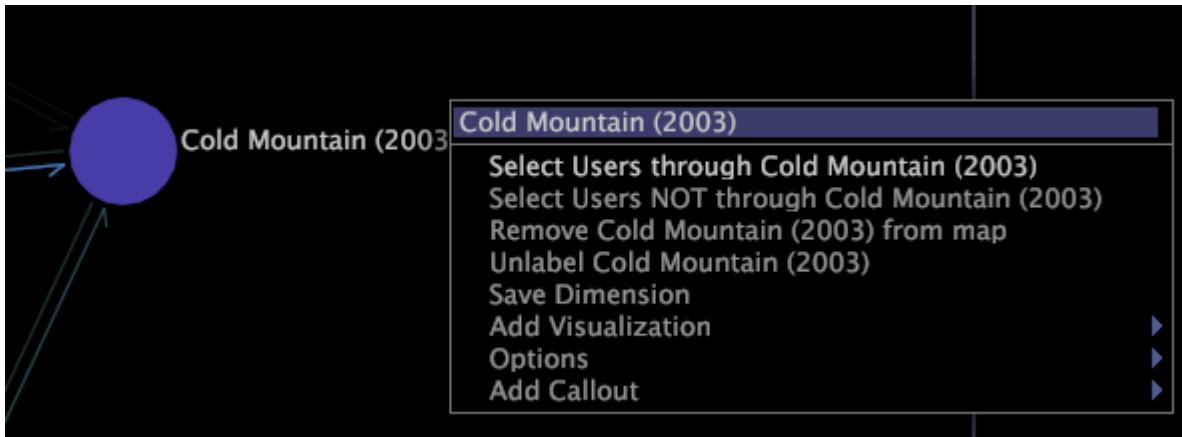
Note: You can change the default group dimension for a process map. See [Configure a process map](#) on page 506.

When you make a selection based on a node within a process map, you are selecting all of the elements of the group dimension that involved that node. To better understand the role of the group dimension, consider the following examples:

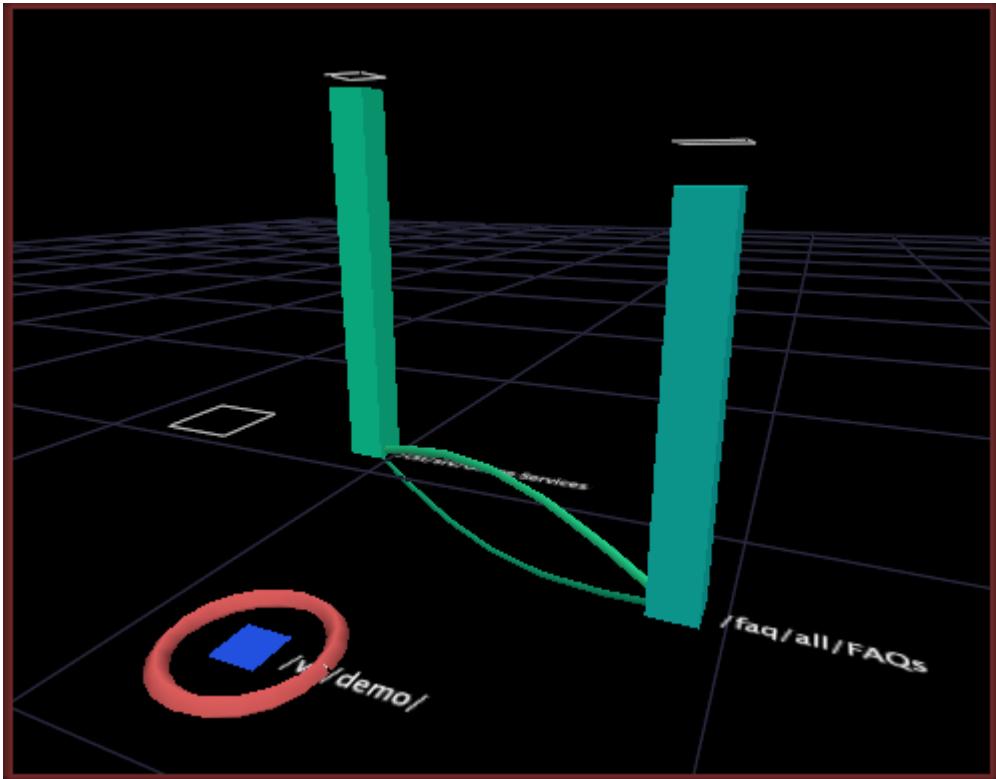
- Movies can be grouped by the viewers who rated them. Each viewer is an element of the User dimension, so the User dimension would be the group dimension for the process map. When you make a selection from a node for a particular movie, you create a filter that shows data for the users that did or did not rate that movie.
- Website pages can be grouped by the sessions in which they were viewed. Each session is an element of the Session dimension, so the Session dimension would be the group dimension for the process map. When you make a selection from a node for a particular page, you create a filter that shows data for the sessions during which that page was or was not viewed.

To make a selection

1. Right-click any node in a process map.
2. Click one of the following options to make a selection based on the node:
 - **Selectgroup dimension name +s through node name:** Filters the data to include all elements of the group dimension that passed through the node by filtering out all sessions that did not pass through the node.
 - **Selectgroup dimension name +sNOT through node name:** Filters the data to include all elements of the group dimension that did not pass through the node by filtering out all sessions that passed through the node.



When you make a selection in a 3D process map, the node for which the selection is made is circled. Benchmarks appear around each bar help you compare metric values with and without the selection. See [Benchmarks](#) on page 285.



Save a dimension from a process map

Navigation title:

Steps to save a dimension from a process map.

1. Right-click the selected nodes and click **Save Dimension**. The **Save Dimension As** window appears. The default save location is the `User\profile name\Dimensions` folder.

To create a new dimension based on multiple nodes, Ctrl+click and drag to create a box around the items that you want to select. Right-click the box and click **Collapse to a single node**. Right-click the collapsed node and click **Save Dimension**.

2. In the **File name** field, type a more descriptive name for the dimension and click **Save**. You now have access to this dimension from any dimension list.

If you would like all users of a profile to use the dimensions that you save, you must upload them to the appropriate profile using the **Profile Manager**. See [Publish files to your working profile](#) on page 459.

Create a visualization from a process map

Navigation title:

You can create other visualizations from a process map.

For example, you can create a graph, table, or scatter plot.

- Right-click any node and click **Add Visualization > <visualization type>**.

Each node on your process map becomes an element of the Map dimension. You can analyze the node data using the features of the selected visualization.

Activate a color link

Navigation title:

You can display the paths between nodes in a process map in color.

The color scheme depends upon the color legend that is open in the workspace. See [Color legends](#) on page 378.

Right-click any node in the map and click **Options > Color Links**.

An X to the left of Color Links indicates that colored links are active in the map. When Color Links is disabled, arrows between nodes appear gray.

Display a metric quantity

Navigation title:

A 2D process map can display the metric quantities representing the amount of activity to or from a node on the map.

The actual sequence of activity may have included nodes that were not added to this map. This quantity is expressed as a percentage, an absolute value, or no value at all.



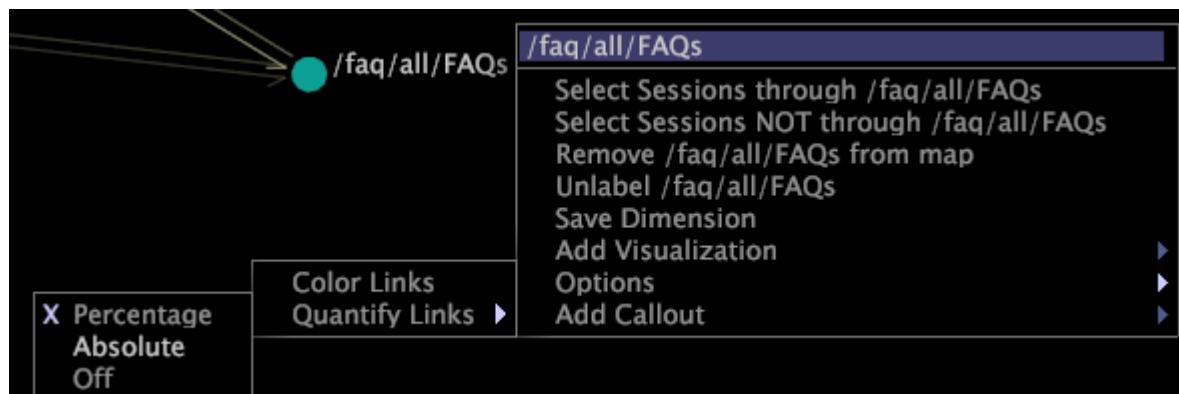
Note: You can display metric quantities for 2D process maps only.

To display metric quantities in a 2D process map

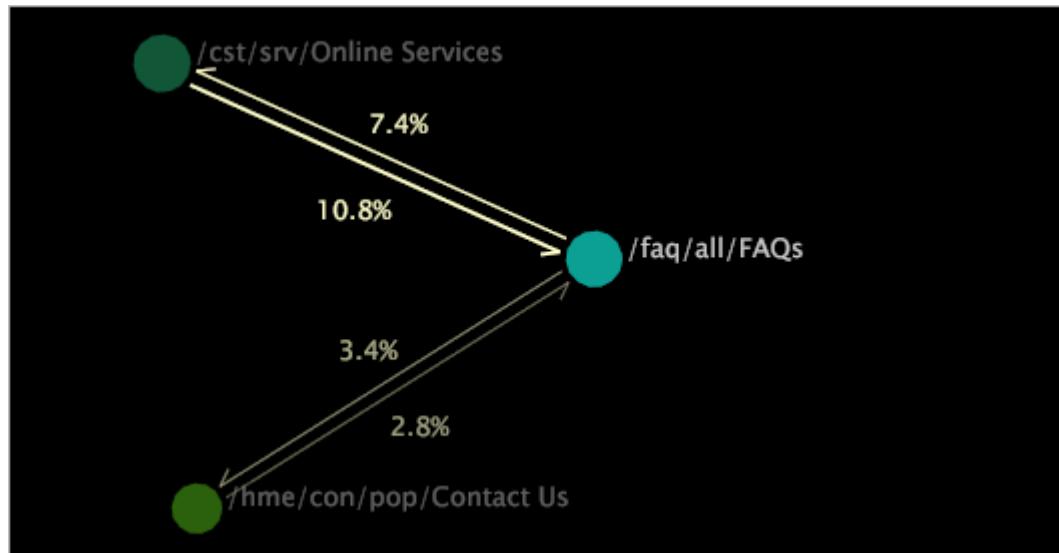
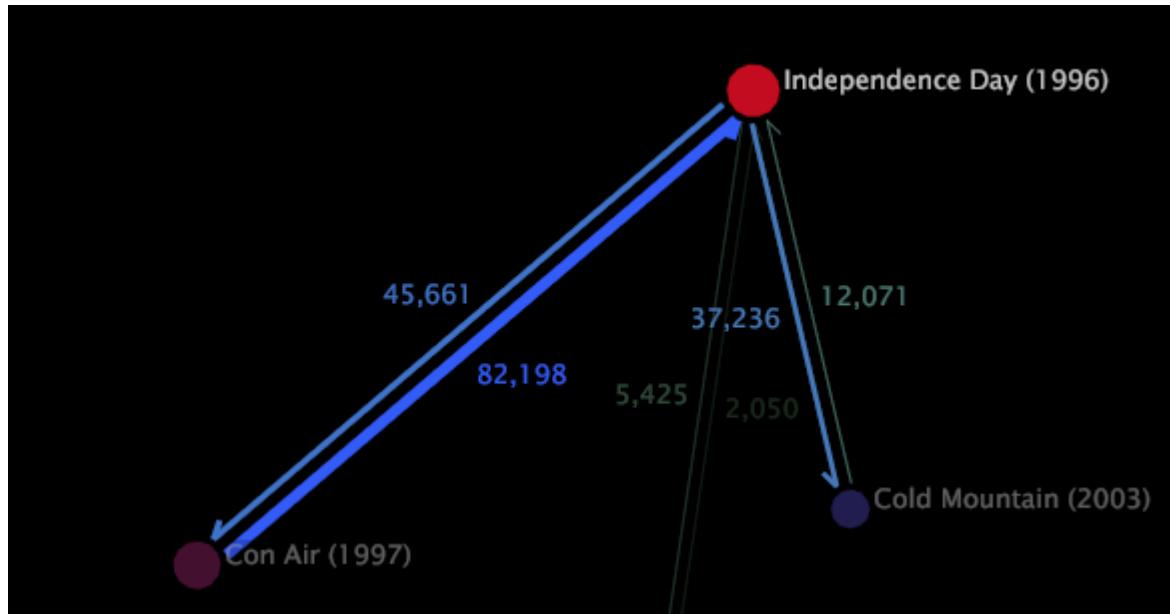
- Click the desired node. All of the other nodes in the process map dim so that you can more easily see the arrows to and from the highlighted node and the associated metric quantities.

To choose how the quantity is displayed

1. Right-click a node and click **Options > Quantify Links**.
2. Click **Percentage**, **Absolute Value**, or **Off**.



The following examples show metric quantities for two 2D process maps.



Add a callout

Navigation title:

Information about adding a callout to a 2D process map.

Right-click the node and click **Add Callout** > *<callout type>*.

For more information about callouts, see [Adding callouts to a workspace](#) on page 293.



Note: You can add callouts to 2D process maps only.

Propensity Scoring

Propensity scoring lets you define customers based on their possibility of a successful conversion or completion of a specified event. It allows you to maximize the potential impact of efforts before executing a process or directing a campaign.

The Value of Propensity Scoring

Propensity scoring lets you perform data discovery to identify hidden behaviors or patterns that exists across your data. Specifically, propensity scoring helps you identify clusters of similar customers using more focused and objective means rather than simple segmentation or filtering. In addition, propensity scoring lets you set up predictive capabilities to identify behavior for your company's high-value customer.

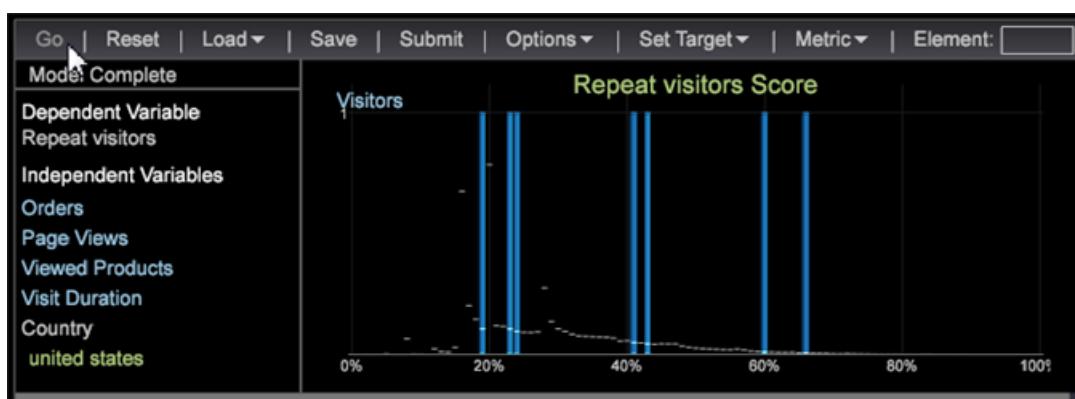
Once you have identified the high-value audience, you can then engage them for the greatest effect. For example, if you are Business to Business company, you may have sales call leads that allow you to then score the leads and identify their likelihood to convert offline. Because every lead increases costs, creating an incentive to identify prospective customers with the highest likelihood of converting a sale is the most effective and the least expensive way to focus your resources.

Propensity scoring provides the ability to identify those factors that are most predictive of a particular score or to increase the likelihood of an event taking place, but it can also be applied to answer specific questions: Will the customer convert? Will the customer respond to an email? Will the customer repurchase? Propensity scoring lets you answer these questions and identify visitors with an inclination for action that can then be set up and scored.

In addition, you can use filters to define a subset of visitors to be scored using the optional **Training Filter** feature. If no filter is applied, then all visitors are targeted for scoring.

Features of the Propensity Scoring Visualization

To open the Propensity Scoring Visualization, click **Add > Visualization > Predictive Analytics > Scoring > Propensity Score**.



The Propensity Scoring Visualization includes these features accessible from its toolbar:

Toolbar Feature	Description
Go	Click to run the scoring process after setting up parameters.
Reset	Clear all settings in the visualization.
Load	Loads a previously created ScoreDim that allows you to change and/or rebuild the scoring model.
Save	Save the Propensity scoring visualization as a Dim file to be accessed and opened as needed.
Submit	Submit scoring task for server-side processing.
Options	Set the Training Filter to limit the subset of visitors. The default filter is Train on Everyone , but you can change it by making workspace selections or building a filter using the Filter Editor .

Toolbar Feature	Description
Set Target	Set the Dependent Variable.
Metric	Add Metrics as Independent Variables.
Elements	Drag Dimension elements using the <Ctrl> + <Alt> keys from Dimension tables.

See also:

1. The [Propensity Gain and Lift Charts](#) on page 130. These views can be opened from a complete scoring model or from Add Visualization> Predictive Analytics > Scoring.
2. The [Model Viewer](#). These views can be opened from a complete scoring model or from Add Visualization> Predictive Analytics > Scoring.
3. The [Complex Filter Description](#) on page 131 feature.

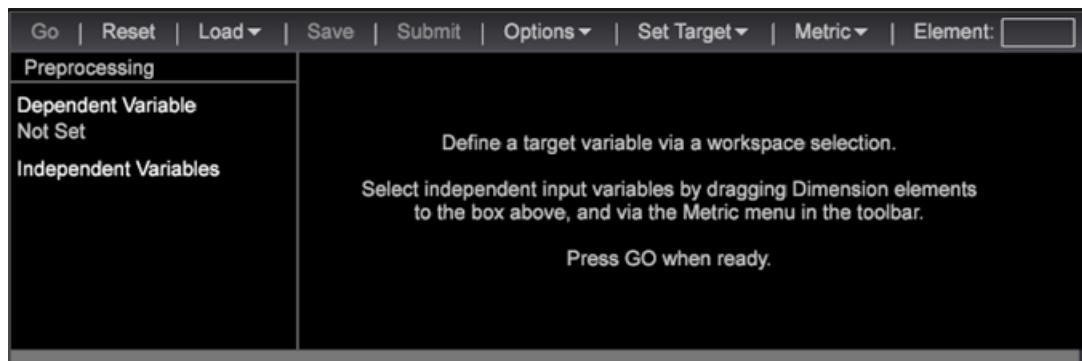
Using the Propensity Scoring Visualization

- **Define one or more filters to define the visitor population for scoring.** This optional **Training Filter** lets you target visitors based on selected criteria. If no training filter is applied, then all visitors are targeted for scoring. If the Training Filter is set, then the scoring result is meaningful to the defined visitor population, although each visitor will still be given a score.
- **Identify the positive visitors.** To define the dependent variable to specify a target filter identifying the positive visitors that match the desired outcome. This can be as simple as Revenue > \$10, or a much more complex filter.
- **The Target filter is not allowed to be the same as the Training filter.** Logically, the Target Filter should be an addition to the Training Filter, resulting in a positive subset of the visitor population to be scored.
- **Select variables of interest (independent variables) as inputs to the Propensity Scoring algorithm.** These can be Metrics or individual elements of a Dimension. Propensity Scoring will start preprocessing just as in [Visitor Clustering](#) on page 155. The system begins capturing a certain amount of samples that match the definition of the previously set training filter (if any). Currently, the sample size is set as 10% of scoring population (defined by training filter), with a minimum of 20,000 and maximum of 100,000, and is bound to the scoring population size.
- A Score Dimension has elements ranging from 0% to 100% that determines the likelihood of the visitors matching the Target variable.

Setting up Propensity Scoring

Follow these steps to use the Propensity Scoring visualization.

1. Open a new workspace and click **Add > Visualization > Predictive Analytics > Scoring > Propensity Score.**



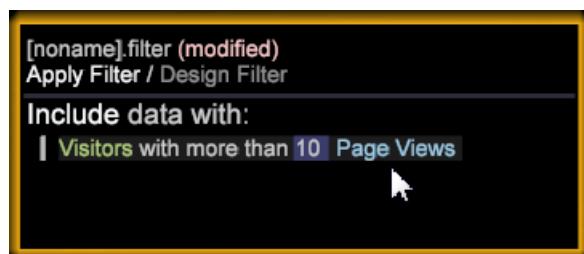
2. Set the **Target** (the dependent variable).

Set the dependent variable by selecting:

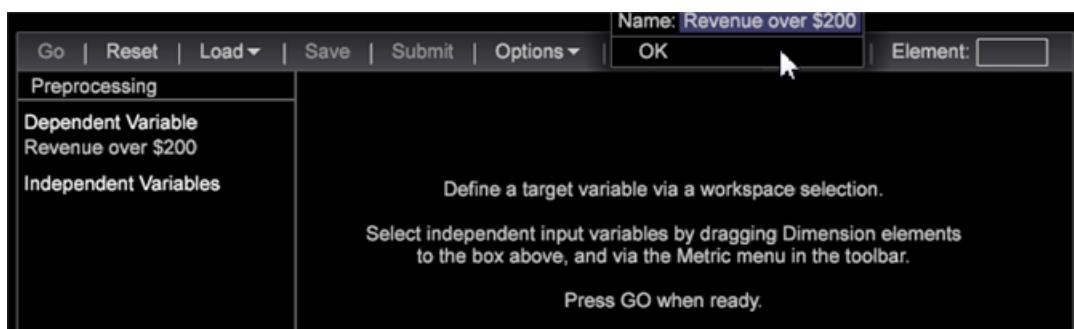
- **Dimension elements:** Right-click in the workspace and select **Table**. Then select a Dimension elements as your dependent variable.

OR

- **Filter Editor.** Click **Add > Visualization > Filter Editor** to open the Filter Editor visualization.



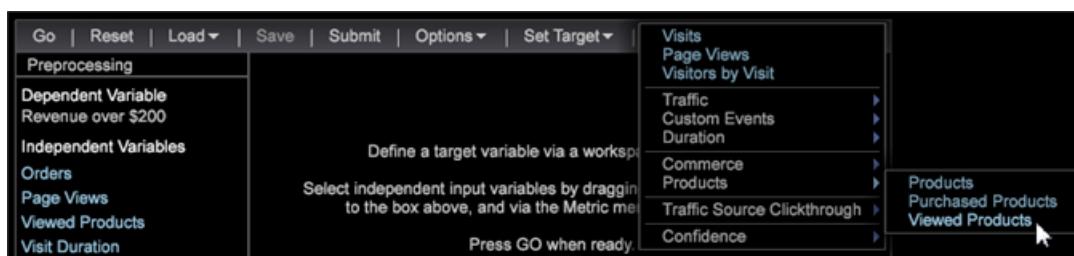
After selecting a Dimension element or Filter as the dependent variable, click **Set Target**, enter a name to describe the dependent variable. Then click **OK** (and make sure the filter box is highlighted) to set the Target.



The name you give the target is the dependent variable that will appear in the left pane.

3. Add independent variables.

Add the independent variables using Metrics or Dimension Elements.



- **Metrics.** From the Propensity Scoring toolbar, select a metric from the **Metrics** menu.

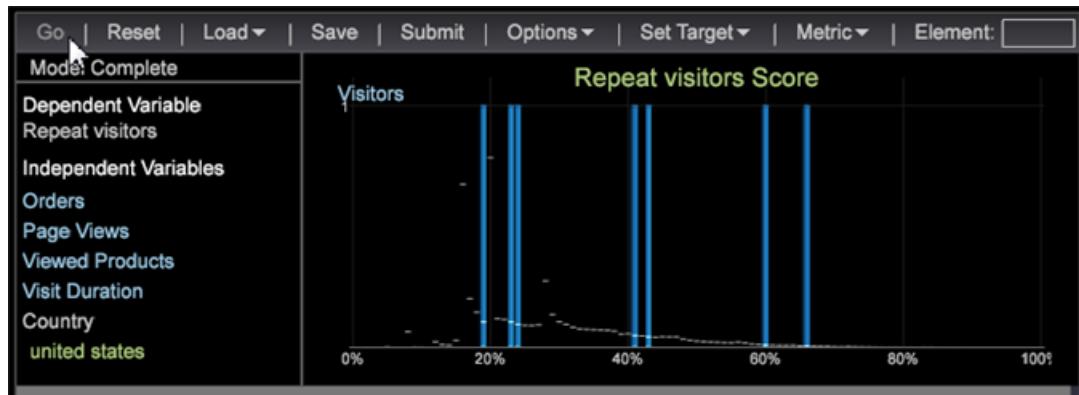
- **Dimension elements:** Right-click in the workspace and select **Table**. Select one or more Dimension elements and drag to the left column under **Independent Variables** or to the **Element** box using the <Ctrl> + <Alt> keys.

-
4. Set **Training Filter**. You can define the set of visitors that you want to score by clicking **Options > Set Training Filter** from the Propensity Scoring toolbar. This will provide a subset of data built using only the visitors that you want to score. For example, who visited in the last month, visitors who reside in Australia, or visitors who viewed specific products.

The default filter is **Train on Everyone**, but you can change it by activating **Dimension Elements** in a table or building a filter using the **Filter Editor**.

After selecting a Dimension element or building a filter and while activated, click **Options > Set Training Filter**, enter a name to describe the filter, and then click **OK**.

5. Once you have identified all your inputs, press **Go**.



The scoring process will begin by passing over the data multiple times. It will then display the results as bar charts over a percentage line.

6. Save Propensity Score.

Starting with 6.1, you now have an option when using the Save Propensity Score:

- Dimension
- Dimension and Metric

You can end up with two saved files, both a dimension and a defined metric.



Note: If you submit the Propensity Score for processing you will get a dimension only.

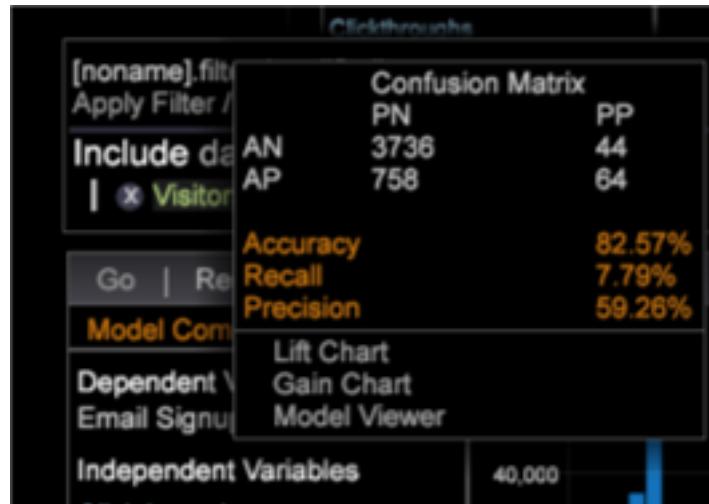
The derived metric is the associated average score metric.

7. Check for accuracy.

The system will display **Model Complete** and generate a scoring model when the process is complete.

Right-clicking on **Model Complete** will identify the accuracy of the scoring model as defined by the system. Values ranging from 0 percent to 100 percent will identify the likelihood of the visitors matching the **Target** variable.

The Confusion Matrix gives four counts by the combination of Actual Positive (AP), Actual Negative (AN), Predicted Positive (PP), and Predicted Negative (PN). These numbers are obtained by applying the resulted scoring model to the 20% withheld testing data of which we know the true answer. If the score is greater than 50%, it is predicted as a positive case (matching the defined event).



Accuracy	Indicates how accurate the model is by identifying the correct predictions over all predictions. $(TP + TN)/(TP + FP + TN + FN)$
Recall	Identifies the ability to re-identify the scoring model. $TP / (TP + FN)$
Precision	Identifies the level of discrepancy. $TP / (TP + FP)$

8. Open a *Propensity Gain and Lift Charts* on page 130, or the *Model Viewer* on page 132.

Right-click on the **Model Complete** visualization and select **Lift Chart**, **Gain Chart**, or **Model Viewer**.

Calculating Propensity Scoring

The statistical calculations for Propensity Scoring are defined.

Conceptually, the score calculated for each visitor is an estimated probability that the specified event (defined by the target filter) might happen, resulting in a score value range from 0 to 100 percent. The scoring procedure uses existing samples as training data to find the relationship between the event probability and the selected independent variables of interest.

Mathematically, such relationships are reflected in each quantitative value associated for each independent variable. Those values are called model coefficients. ScoreDim currently uses the Iteratively Reweighted Least Squares (IRLS) algorithm to estimate the model coefficients. IRLS goes through the samples multiple times until the difference of coefficients between current pass and the previous pass is less than 1.0e-6, at which point it is called **converged**. However, depending on the data, IRLS may not be able to reach convergence.

In such case, the model training iteration will terminate when

- the coefficient difference gets larger instead of smaller,
- 1,000 passes have been reached, or
- a mathematical error prevents continuing iteration.

If IRLS doesn't converge, a backup algorithm called Stochastic Gradient Decent (SGD) will be used. SGD will also go through the training samples multiple times. But unlike IRLS, the SGD model coefficients are controlled so that the difference between iteration will always decrease in an exponential manner. Similarly, SGD will terminate when the coefficient difference falls below 1.0e-6 or 100,000 passes have been reached. The failure of IRLS and engagement of SGD will be recorded in trace log.

For both algorithms, not all samples go into model training. Eighty percent are currently used to train the model. After the model is trained, the remaining twenty percent samples will be used to assess the model strength in terms of Accuracy, Recall, and Precision that is calculated from the confusion matrix. The closer to 100%, the better the scoring model.

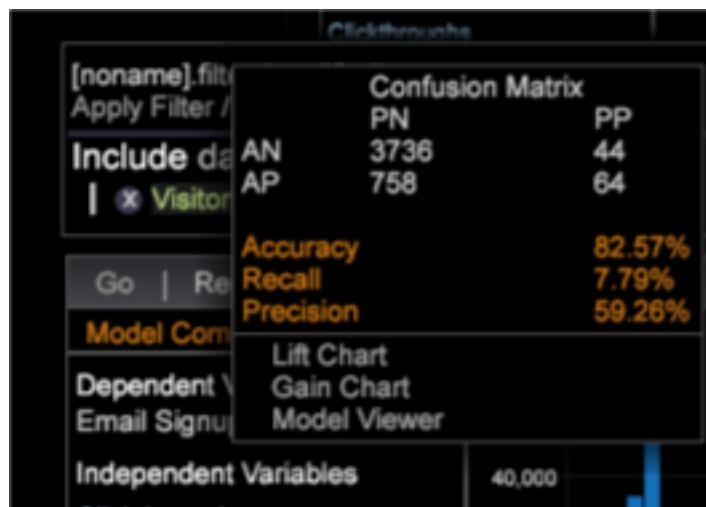
Propensity Gain and Lift Charts

The Lift and Gain charts offer visualizations for evaluating the potential performance of a scored model to evaluate performance over defined portions of the audience.

Gain and lift charts are visualizations built to evaluate the potential performance of the scored model. These charts evaluate performance over each portion of the population.

To Open a Lift or Gain chart

1. Select Add Visualization > Predictive Analytics > Scoring.
2. Hover over **Model Complete** of a saved score.



About Lift and Gain Charts

The Lift and Gain Charts are useful visual tools for measuring the value of a predictive model. Both charts consist of a lift curve (green) and a baseline (pink). For the **Gain Chart**, the distance between the lift curve and the baseline represents how much you can improve performance in responses (or the "gain") from using the predictive mode. The gain is realized by prioritizing and targeting the prospects (customers/visitors) who are most likely to convert, rather than marketing to customers/visitors at random. In this way, you can quantify the expected value of using the predictive model to choose which prospects to contact.

Similar to the Gain Chart, the **Lift Chart** shows how much more likely you are to receive positive responses than if you contacted prospects at random. You want the distance between the lift curve and the baseline to be as large as possible, representing larger expected gains from using the predictive model to contact customers. Mathematically, the gain and lift charts are defined as follows:

- **Gain** = (Expected Response using Predictive Model to Contact Prospects) / (Expected Response from Randomly Contacting Prospects)
- **Lift** = (Expected Response among a Specific Group Size of Prospects identified using the Predictive Model) / (Expected Response among the same Specific Group Size of Prospects identified Randomly)

Example of Lift and Gain Charts

For example, consider the example of a retailer who wants to launch an email re-marketing campaign to sell yoga pants. Historically, the analyst expects an average response rate of 20 percent based upon past email

re-marketing campaigns similar to this one. While the analyst has nearly 5 million customers in its email database, the business only wants to market to those customers that are most likely to respond to the email and purchase. In this way, the company will maximize the ROI of the campaign while ensuring that they don't unnecessarily send emails to uninterested customers. Given an expected response rate of 20 percent, the marketer and analyst expect that approximately 1 million customers are likely to respond and purchase. Rather than randomly guessing which of those customers will be among the 20 percent responses, the analyst wants to be smart about predicting which of the one million prospects (among the database of 5 million customers) are most likely to respond.

Using Adobe's Audience Scoring capability, the analyst defines success as a prospect clicks on an email and purchases yoga pants (the dependent variable). After selecting the independent variables (based upon experience and knowledge gained from analyzing data correlations and audience clustering among other analyses), each prospect is scored on their likelihood of positively responding the email re-marketing campaign (clicking on the email and purchasing yoga pants). The analyst opens the resulting Gain and Lift charts based upon the predictive model.

The y-axis shows the percentage of the cumulative expected positive responses. In our example, we expect a total of 1 million positive responses. A value of 20% on the y-axis corresponds to 20% of the 1 million expected positive responses, or 200,000 positive responses. The x-axis shows the percentage of prospective customers contacted. In our example, the x-axis represents a fraction of the 5 million customers in the email database. The baseline (Pink) is the overall response rate - if you contact X% of prospects then you will receive X% of the total positive responses. Using the predictive model, the lift curve (green) shows the percentage of positive responses obtained (y-axis) by contacting a given percentage of prospects (x-axis).

The Lift chart plots the expected lift as a result of using the predictive model to determine the top one million prospects most likely to purchase yoga pants after receiving and clicking on the email. For contacting 20 percent of randomly selected prospects using no predictive model, you should expect to get 20 percent of responders. However, using the predictive model to identify the top 20 percent of prospects most likely to respond, you expect to get 50% of responders. The y-value of the lift curve at 20 percent is $50/20 = 2.5$. The lift chart shows how much more likely you are to receive respondents than if you contact a random sample of prospects. For example, by contacting only 20 percent of prospects based on the predictive model you will reach 2.5 times as many respondents as compared to having not used any predictive model.

Model Viewer

A model viewer lets you generate a logistic regression model using the Propensity Scoring feature.

The Model Viewer displays the coefficient weights of each input variable (including the constant term) and their statistical error range. Input variables showing a high absolute coefficient and small margin of error are the most significant predictors in the model.

To Open a Model Viewer chart

1. Select Add Visualization > Predictive Analytics > Scoring .
2. Hover over Model Complete of a saved score.

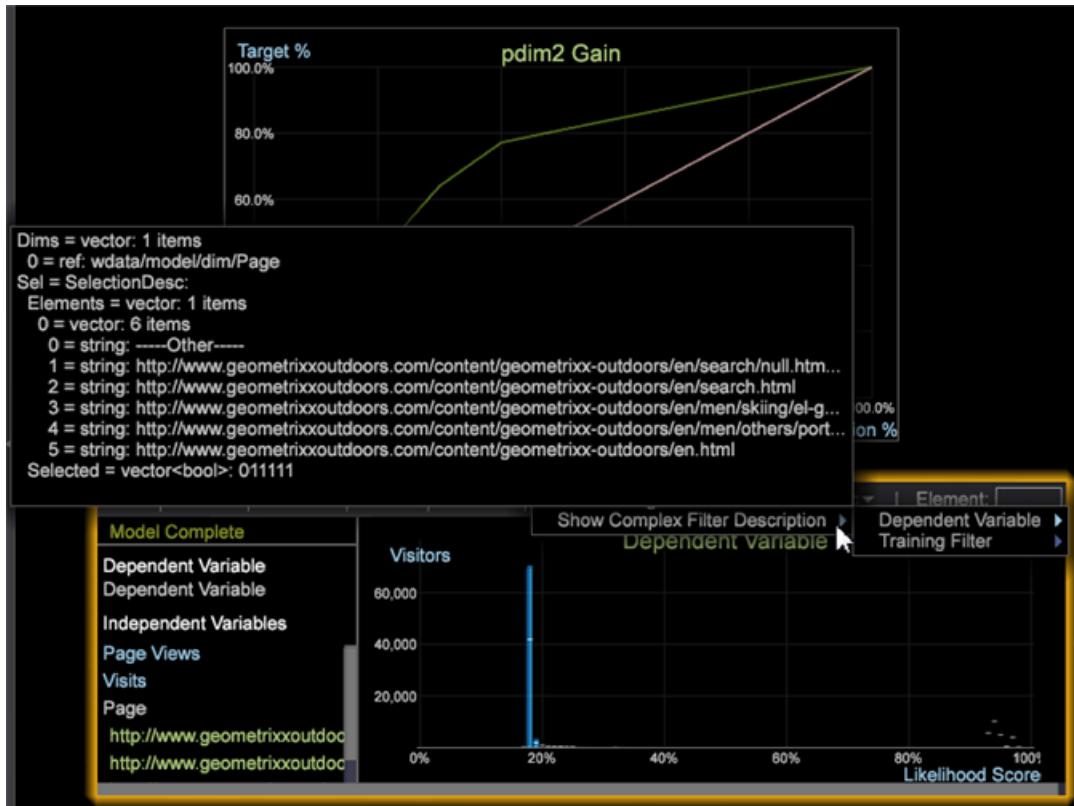


The input variables with a coefficient ≥ 1 are positive influences on the propensity model. The coefficients that are < 1 are negative influences on the propensity model. The range defined within the parentheses is the error, and indicates the consistency of the input variable across the successful population.

Complex Filter Description

You can add a Show Complex Filter Description for Dependent Variables and for the Training Filter.

To **Show the Complex Filter Description**, from the scoring dialog select Options > Show Complex Filter Description.



Radar visualization

Navigation title:

Radar charts provide quick focus on the areas that are in most need of attention, by providing a visual view of a set of metrics, and how they relate or differ.

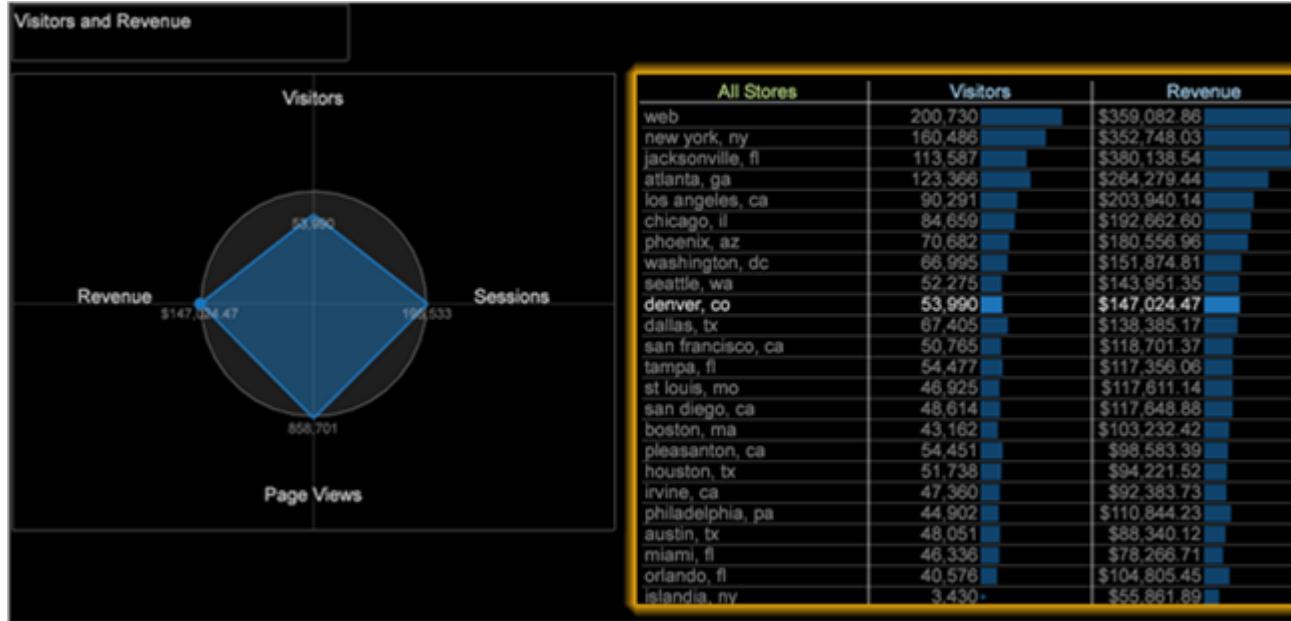
Many times you need more than one metric to understand and evaluate selected observations in a workspace.

This visualization is useful for comparisons or benchmarks among the table selections. For example, you might add a workspace table that lists stores, then add a radar visualization with metrics like Revenue, Visitors, and Page Views. (As shown in the screen in the following procedure.) As you make store selections in the table, the radar chart footprint shifts, identifying weaknesses or strengths in the metrics for the selected store.

Each radial of a radar chart is a metric, and a minimum of three metrics are required. The metric data is plotted in relation to one anchored metric. The anchored metric and the Scale to Anchor parameter for each metric determine the scaling of the metrics with respect to benchmarks.

To create a radar visualization

1. Right click in the Workspace, then click **Visualization > Radar**.



2. To add metrics, right-click in the visualization and select **Add Metric**.
3. To anchor a metric to the chart, right click on a metric and choose the following option:

Anchor to this metric: Uses this metric as the benchmark to which other metrics are drawn. You can anchor one metric at a time. Each metric on the chart is filtered by the active workspace selection, or by no filter. The benchmark ratio between these two values is plotted on the axis between the center of the graph and metric name on the radar. Zero is plotted in the center.

4. To scale a metric with the anchored metric, right-click the metric and choose the following option:

Scale with anchor: When enabled, this metric's axis is scaled so that the benchmark ratio for the selected anchor metric is plotted in the circle, with zero at the center. When not selected, the circle represents a benchmark ratio of 1. Typically, you turn on Scale with Anchor for countable metrics, such as Visitors or Page Views, and turn it off for ratio metrics, such as Conversion, Average Session Duration, or Page Views per Session.

Statistical display in tables

Navigation title: Statistical display in tables

The statistical display lets you analyze how data relates to each other in a multi-dimensional table.

For example, you might want to know whether purchases of a red shoe are statistically related to purchases of a red belt or a different color belt. The statistic shows you whether the relationship is more than you expected, or less than you expected.

On the top border of a table, right-click **Display As > Stats**.

Segments

Navigation title:Segments

Segments enable you to build dimensions on-the-fly from your selections.

Segments are most useful when working with disjointed, non-overlapping groups that you want to use as a dimension, such as when comparing two types of shoppers based on their previous purchases.

Using a segment affects only the workspace in which you are using the segment dimension.

Create a segment dimensions

Navigation title:

To create a segment dimension, you begin by making a selection within a workspace and then adding the segment to a visualization.

To create a segment dimension

1. Add a segment visualization to the workspace. For example:

Segments	Visitors	Pct of Visitors
Other	31,867	100.0%

2. Add visualizations to your workspace that you want to use to define your segment, then make the desired selections to define your segment.
3. In the segment visualization, right-click the label of the segment after which you want the new segment to be added and click **Add Segment**.



Note: To create a new first segment, right-click the **Segments** label and click **Add Segment**.

Segments	Visitors	Pct of Visitors
Other	1,514	69.1%
New Segment	676	30.9%

A new segment (named New Segment) appears in the visualization. The Other segment represents all of the data not included in your defined segments: it is effectively the difference between your dataset data and your segment data.

4. Right-click the newly created segment and click **Rename Segment**.
5. Type a descriptive name for your new segment in the name field.



Note: If a metric value, such as a particular visitor in **Site**, meets the criteria of multiple segments, the metric value is included in only the first listed segment that it matches.

To save the segment dimension

1. Right-click the Segments label and click **Save Dimension**. The **Save Dimension As** window appears. The default save location is the `User\profile name\Dimensions` folder.
2. In the **File name** field, type a descriptive name for the segments that you are saving as a dimension and click **Save**.

You can access the segment dimension whenever you are working with a visualization. You also can export data associated with the elements in your saved dimension using the segment export feature.

For more information about the segment export feature and instructions to configure it for your needs, see [Configure segments for export](#) on page 234.

Reusing a segment visualization

Navigation title:

If you want to use the same segment information again, you can save the segment visualization, the workspace containing the segment visualization, or save the segments themselves as a dimension.

If you would like all users of a profile to use the segments and dimensions that you create or modify, a user with the appropriate permissions must upload them to the appropriate profile using the **Profile Manager**. See [Publish files to your working profile](#) on page 459.

To save the segment visualization for use in other workspaces

1. Right-click the top border of the segment visualization and click **Save**. The **Save** window appears. The default save location is the User\profile name\Work folder.
2. In the **File name** field, type a more descriptive name for the visualization and click **Save**.

To retrieve the saved segment visualization

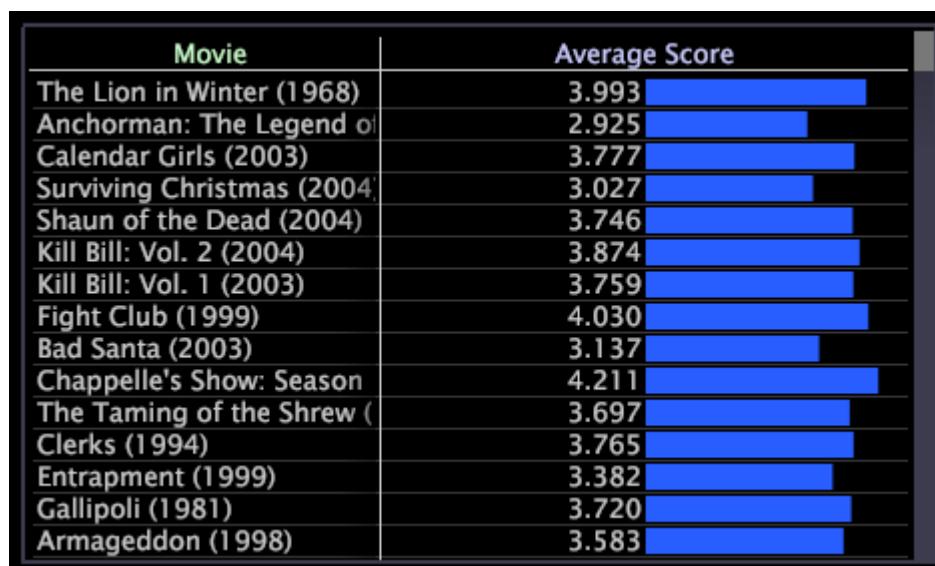
1. Right-click within the workspace and click **Open > File**. The **Open Visualization** window appears.
2. Navigate to the segment visualization that you saved in the User\profile name\Work folder.
3. Select the segment visualization file (*.vw) and click **Open**.

Tables

Navigation title: Tables

A table can display one or more metrics across one or more dimensions of data.

The following example shows you a table in which the Average Score for each Movie is displayed as a value but also is depicted by the length of its bar.



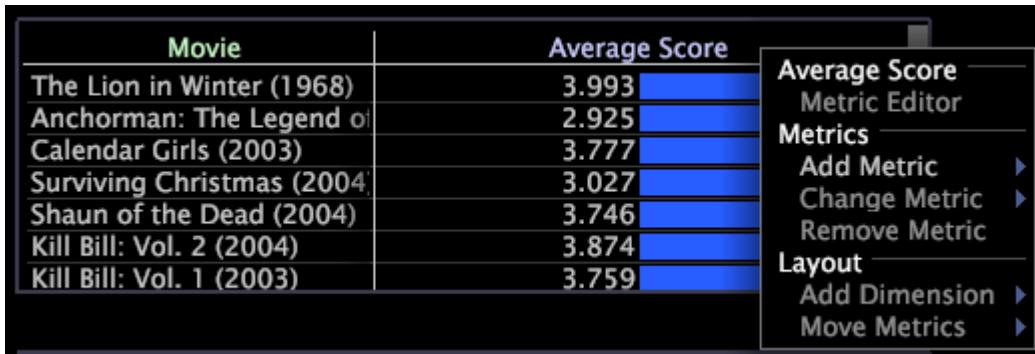
Add and change a metric

Navigation title:

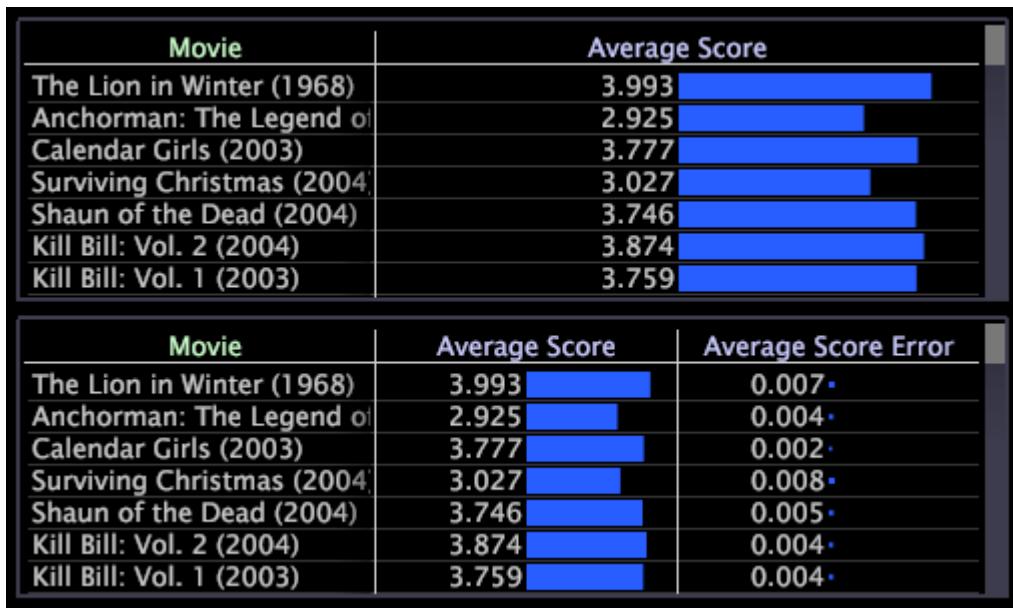
Information about adding and changing metrics.

To add a new metric

- Right-click the label or an element of the metric, click **Add Metric > <metric name>**. The metric is added to the right of the metric that you selected originally.



The following example shows the Average Score metric by the Movie dimension (in the top table) and the same table with a second metric, Average Score Error, added to the table.



To change a metric

- Right-click the label or an element of the metric that you want to change, click **Change Metric > <metric name>**. The new metric replaces the original metric.

Add, change, and move a dimension

Navigation title:

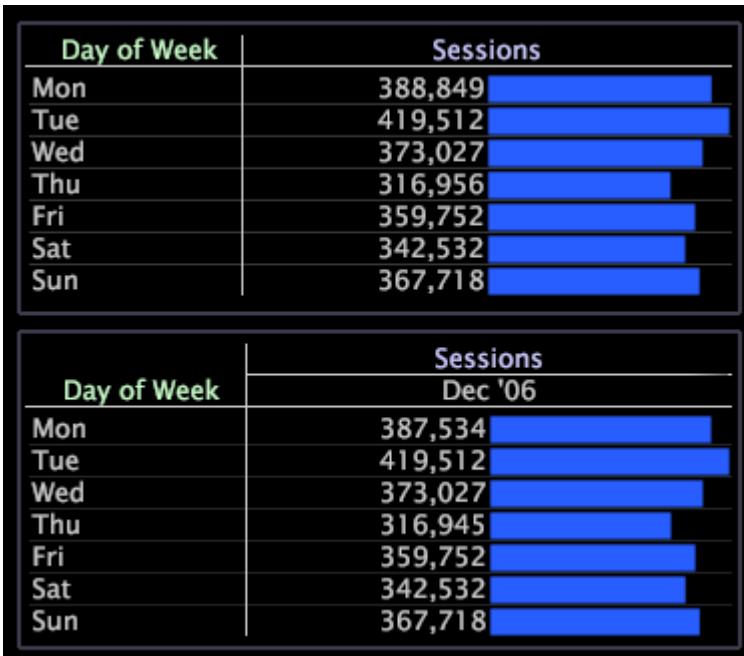
You can add multiple dimensions to a table to attain a more detailed cross-tabulation of the data.

You can add dimensions to either axis of the table.

To add a new dimension

- Right-click an element or the label of any dimension or the label of any metric and click **Add Dimension > <dimension name>**. The dimension is added to the visualization on the chosen axis.

The following example shows the Sessions metric graphed over the Day of Week dimension (in the top table) and the same table with a second dimension, This Month, added to the top axis (in the bottom table).



To change a dimension

- Right-click an element or the label of the dimension you want to change and click **Change Dimension > <dimension name>**.

To move a dimension to another location

You can move a dimension to a different position on the same axis or to the opposing axis.

- Right-click an element or the label of the dimension that you want to move, click **Move<dimension name>**, then complete the appropriate step:
 - To move a dimension to the opposing axis, click to top axis or to left axis.
 - To swap locations with another dimension on the same axis, click **(move here)**.

Fit data to a window

Navigation title:

The Fit to Window option allows the data columns in a table to automatically expand or shrink to fit the window each time you resize it.

When **Fit to Window** is disabled, the data columns must be sized manually.

To enable or disable the fit to window option

- Right-click the top border of the visualization and click **Fit to Window**.



When enabled, an X appears to the left of **Fit to Window**.

To resize data columns manually

- Click the line dividing any two data columns (in the top-axis region) and drag to the desired width.

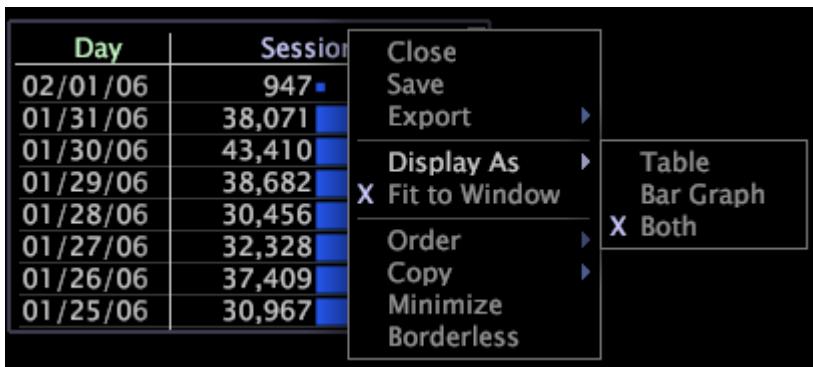


Change the table display

Navigation title:

Information to select the display type.

- Right-click the top border of the visualization and click one of the following options:
 - Display As > Table**
 - Display As > Bar Graph**
 - Display As > Both** (default)



An X appears to the left of the active display mode.

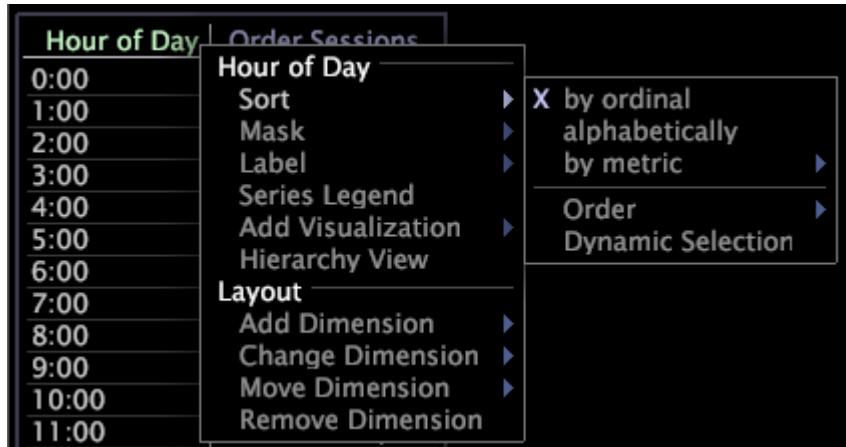
Sort data in a table

Navigation title:

Steps to sort data.

If the table has only one dimension, you can simply click the label of the metric on which you want the data sorted.

1. Right-click an element or the label of the dimension that you want to sort and click **Sort**.



2. Click one of the following options:

- **By ordinal** to sort the elements according to the natural order of the elements. For example, the elements of the Hour dimension displays in chronological order. If the dimension has no natural order, such as with referrer or URI, the sort order is not useful, so you should select to sort either alphabetically or by metric.
- **Alphabetically** to sort the dimension alphabetically by element name.
- **By metric** to select a metric by which you want the data sorted. For example, you can sort the Referrer dimension by the Sessions metric to see which referrers contribute the most sessions on your site.

When you sort by a metric, by default the order in the table corresponds to the values of the metric as affected by the selection at that time. If you later change the selection, the sorted order does not change from the original order unless the dimension is resorted or you enable Dynamic Selection. When you click **Sort > Dynamic Selection**, the table is resorted each time you change the selection.

To sort by an existing metric in the table, click the metric label.

3. (Optional) To choose whether to sort in ascending or descending order, right-click an element or the label of the dimension that you want to sort and click **Sort > Order > Ascending** or **Sort > Order > Descending**.

If the table has only one dimension, you can simply click the label of the metric to reverse the sort order. Clicking the label again reverses the sort order.

Mask data

Navigation title:

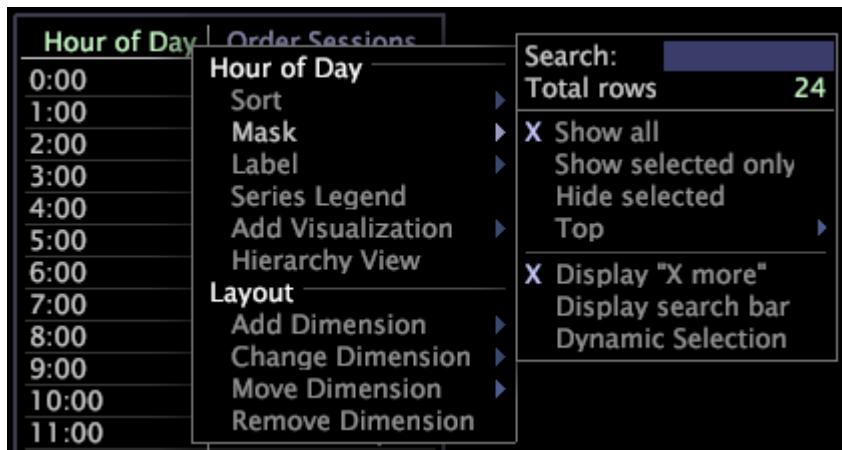
Masking refers to selecting a subset of your data or a subset of the elements in a dimension.

You mask or hide those elements that you do not want included in the analysis.

Data Workbench provides two methods for masking dimension elements. The first method employs the options available in the **Mask** menu. Using the **Mask** menu options, you can use your mouse to select those elements to show or to mask, or you can show top-ranked elements when you sort the data by metric. The second method for masking dimension elements employs a search.

To mask data

1. Right-click an element or the label of the desired dimension and click **Mask**.



2. Click one of the following options:

- **Show all**
- **Show selected only**
- **Hide selected**
- **Show top > 5, 10, 25, 50, 100, or 500** of the displayed elements sorted by metric
- **Show top > All Positive** to show only values greater than zero (0)
- **Display “X more”** to show the number of currently masked elements
- **At least one ><countable dimension name>** (available only when working with a denormal dimension)

When working with a denormal dimension, this option enables you to mask a dimension by a countable dimension. When selected, the table shows only those elements that have at least one element of the countable dimension that you selected. The table displays up to 1,023 elements.

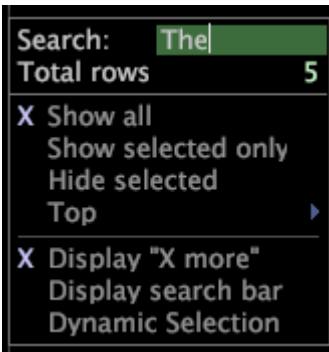


Note: Because the Adobe Platform processes data in random order, when At least one masking results in more than 1,023 elements, the more common and larger elements have a greater chance of being included in the table.

When you mask by Show top or At least one, by default the order in the table corresponds to the values as affected by the selection at that time. If you later change the selection, the order does not change from the original order unless the table is resorted or you enable Dynamic Selection. When you click **Mask > Dynamic Selection**, the table is resorted each time you change the selection.

To mask data using a search

- You can mask data using either of the following search options:
- Right-click an element or the label of the desired dimension, click **Mask**, then in the **Search** box type the phrase for which you want to search.



- Right-click an element or the label of the desired dimension, click **Mask > Display search bar**, then in the search box that displays in the dimension label cell, type the phrase for which you want to search.

The	Average Score
The Green Mile (1999)	4.303
The Good Girl (2002)	3.098
The Day After Tomorrow (2004)	3.446
The Bourne Identity (2002)	3.968
The Usual Suspects (1995)	4.372
Kiss the Girls (1997)	3.864
The Italian Job (2003)	3.747
The General's Daughter (1999)	3.347
The Sopranos: Season 1 (1999)	4.412
The Firm (1993)	3.644

As you type a search phrase, Data Workbench updates the dimension to reflect matches.

To further constrain the masking during a search, you can use any of the following methods:

- You can type “re:” in the **search** box or bar to have the search phrase interpreted as a regular expression. You can use any of the syntax associated with regular expressions in your search phrase. For more information about regular expressions, see the Regular Expression appendix in the *Dataset Configuration Guide*.
- You can type the \$ symbol as the first character in your search string to find phrases that begin with the string you entered, or as the last character to find phrases that end with the string you entered.
- You can type a space as the first character in your search string to find any words within a phrase that begin with the string you entered, or as the last character to find any words within a phrase that end with the string you entered.

Following are examples of different ways to mask a table using the string “on” in a search:

- Typing “on” displays every phrase that contains the string “on” anywhere in the phrase: “**online** banking,” “contact buyers,” “bullion coins,” “bank **online**,” “gold options,” and “silver **bullion**.”
- Typing “\$on” displays every phrase that begins with the string “on”:
“**online** banking” and “on-line payment.”
- Typing “on\$” displays every phrase that ends with the string “on”:
“silver **bullion**” and “gold option.”
- Typing “on” displays every phrase that contains a word that begins with the string “on”:
“**online** banking” and “bank **online**.”

- Typing “on” displays every phrase that contains a word that ends with the string “on”: “bullion coins” and “silver bullion.”
- Using “on” displays every phrase that contains the string “on” as a word: “on line banking” and “bank on line.”

Enable a series legend

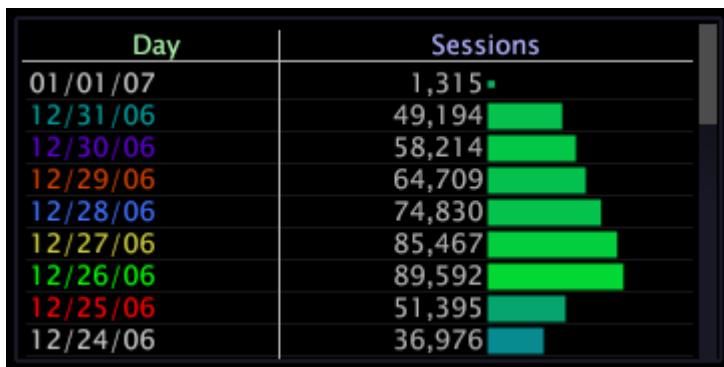
Navigation title:

Elements within a dimension can be color-coded as a series so that each element can be tracked more easily across other visualizations within the workspace.

To enable or disable a series legend

- Right-click an element or the label of the dimension that you want to color-code and click **Series Legend**.

When the series legend is active, an X appears next to the menu option and the elements within the dimension appear in alternating colors.



Open a URI from a table

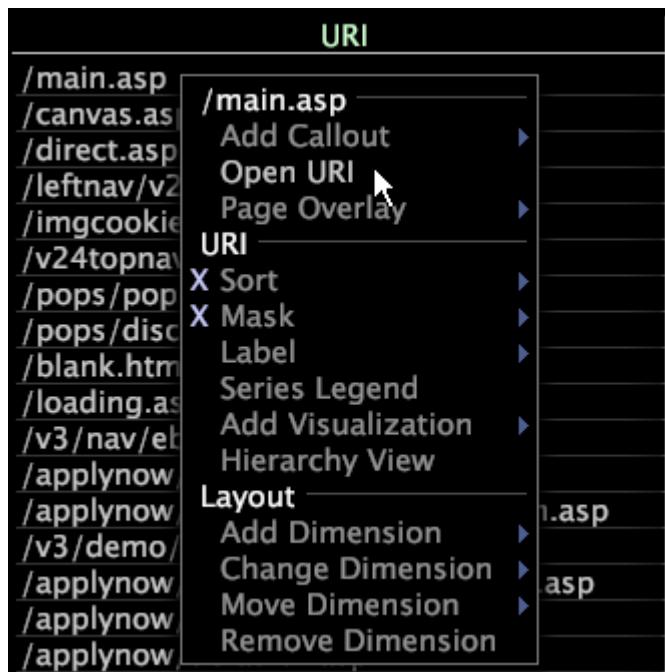
Navigation title:

Open functionality is currently configured only in the **Site** application and only for opening URIs.

See [Configure open functionality](#) on page 499.

In **Site**, from a URI table, you can right-click a URI to display a web page in a web browser. To view a URI in its native format (such as HTML), Data Workbench must have access to the referenced location and the application needed to open that item. For example, to view a web page, Data Workbench must have access to the Internet as well as have a web browser installed.

- Right-click an element of the dimension and click **Open URI**.



The URI opens in its native application.

Apply hierarchy views

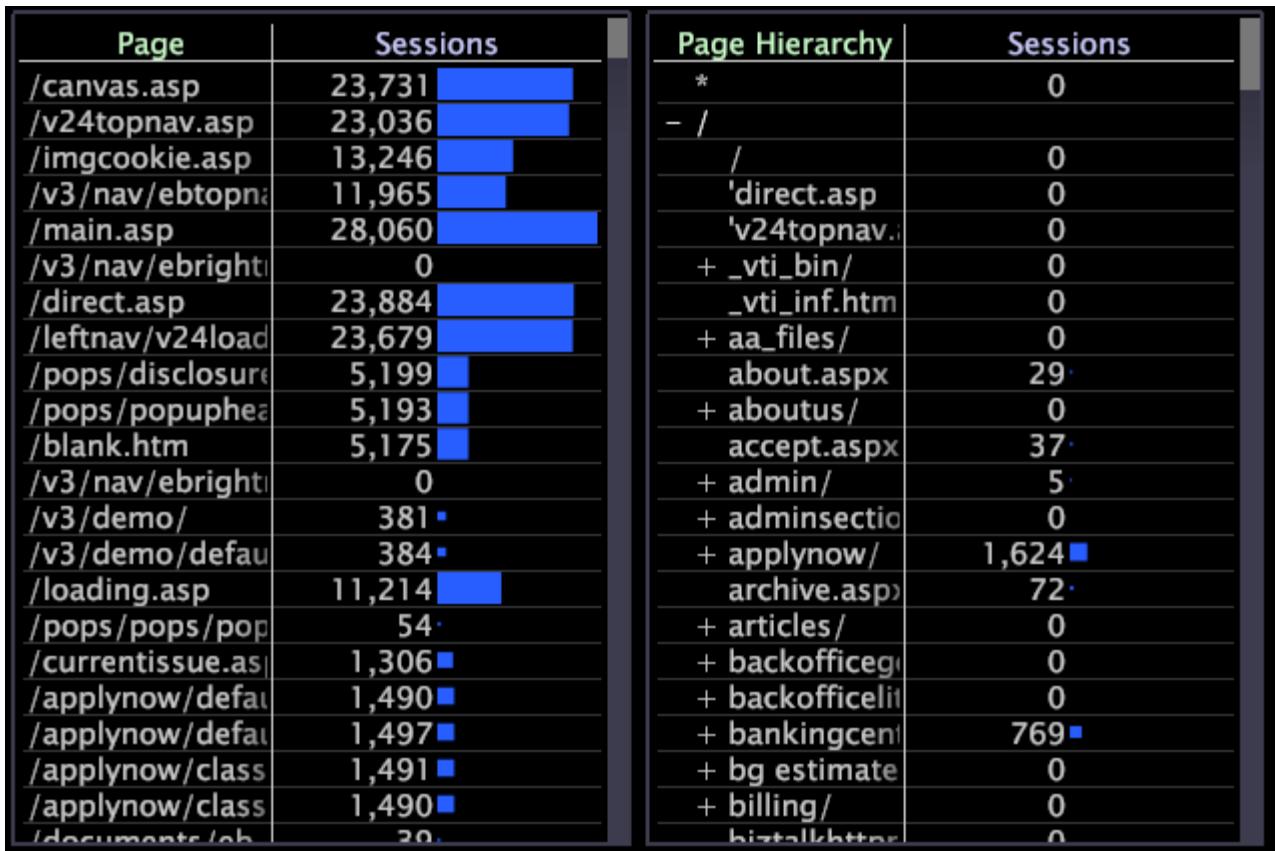
Navigation title:

Hierarchy views are available only when using the Site or HBX application.

The hierarchy view displays the pages in a website organized hierarchically by file name and sorted alphabetically. While useful for analysis itself, the hierarchy view also can be used to set up such advanced visualizations as process maps. For more information about process maps, see [Process map](#) on page 390.



Note: If your dataset has been configured to run on multiple servers in a cluster, for this feature to work properly, your system administrator must designate which machine functions as your Central Normalization Server. For steps to do so, see the Log Processing Configuration File chapter of the *Dataset Configuration Guide*.



To enable or disable the hierarchy view

- From any page or URI visualization, right-click an element or the label of the page dimension and click **Hierarchy View**.

URI	Page
/leftnav/v24loadlnav.asp	482,080
/main.asp	
/v24topnav.asp	
/v3/nav/ebrightnavbranch	
/imgcookie.asp	
/direct.asp	
/canvas.asp	
/loading.asp	
/v3/nav/ebtopnavbranch	
/documents/wm_risk.asp	
/pops/disclosure_pop.asp	
/pops/popupheader.asp	
/blank.htm	
/currentissue.aspx	
/applynow/default.asp	
/applynow/default2.asp	
/applynow/classes/produc	

/leftnav/v24loadlnav.asp

- Add Callout
- Open URI
- Page Overlay
- URI
- X Sort
- X Mask
- Label
- Series Legend
- Add Visualization
- Hierarchy View**
- Layout
- Add Dimension
- Change Dimension
- Move Dimension
- Remove Dimension

An X is shown next to the option when the **hierarchy view** is active.

The hierarchy is organized into website sections and pages using a tree structure. Sections (nodes) can be expanded or condensed using the + or – symbol next to the section name. Individual pages do not have a + or – symbol next to them.

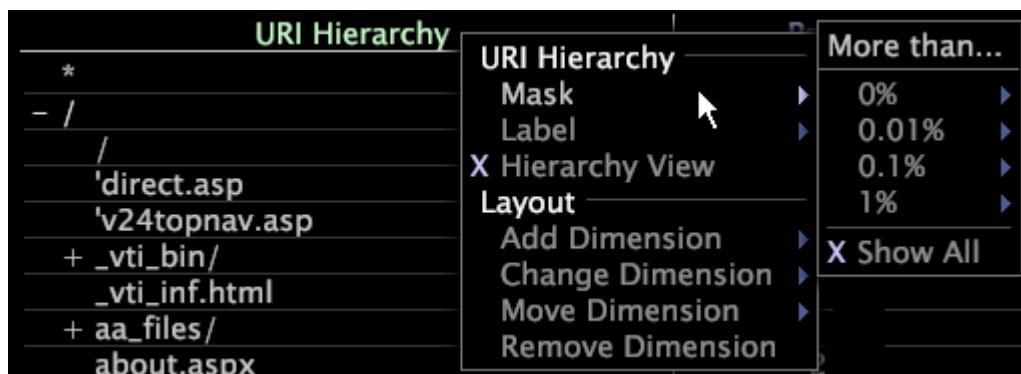


Masking Dimension Elements in a Hierarchy View

Masking refers to selecting a subset of your data or a subset of the elements in a dimension. You mask or hide those elements that you do not want included in the analysis. Using the **Mask** menu options for hierarchy views, you select the minimum percentage of a metric that an element must have to be displayed in the visualization.

To mask data using the Mask menu option

1. Right-click an element or the label of the dimension and click **Mask**.



2. Under More than, click the appropriate percentage, then click the metric that you want to mask.

For example, if you click 0.1%, and then click Page Views, you are masking (hiding) any element that has less than 0.1% of the total number of pages views and displaying any element that has more than 0.1% of the total number of pages views. If you click 0%, you are masking all elements with a value of 0 (zero) for the selected metric.

Export to Microsoft Excel

Navigation title:

Link to information on exporting Microsoft Excel.

For information about exporting windows, see [Export window data](#) on page 228.

Export to a TSV file

Navigation title:

Link to information on exporting to a TSV file.

For information about exporting windows, see [Export window data](#) on page 228.

Work with data in worksheets

Navigation title:

Text or expressions can be entered into any cell of a worksheet.

All expressions in a worksheet are preceded by an equal sign (=) unless using `eval()`, which treats the text in the referenced cell as a expression.

For a full list of metric, dimension, and filter syntax rules, see [Query language syntax](#) on page 515.

To type data into a worksheet

1. Click twice on a cell in the spreadsheet to enter edit mode. The selected cell is highlighted.
2. Type or paste the desired data into the cell.

To copy and paste from one cell to another

1. Right-click the cell containing the data that you want to copy and click **Copy**.
2. Right-click the cell into which you want to paste the copied data and click **Paste**.

Data Workbench automatically updates the references in the new cell to refer to the appropriate columns and rows.

To copy and paste from a group of cells to another

1. Select the cells containing the data that you want to copy.
2. Right-click the cells containing the data that you want to copy and click **Copy**.
3. Right-click the first cell into which you want to start pasting the copied data and click **Paste**. The data is pasted into the first cell and below it.

Data Workbench automatically updates the references in the new cell to refer to the appropriate columns and rows.

To insert a column

- Right-click a column and click **Insert Column**. The new column is inserted to the left of the selected column.

To delete a column

- Right-click the column that you want to delete and click **Delete Column**. The column is removed.

To insert a row

- Right-click a row and click **Insert Row**. The new row is inserted above the selected row.

To delete a row

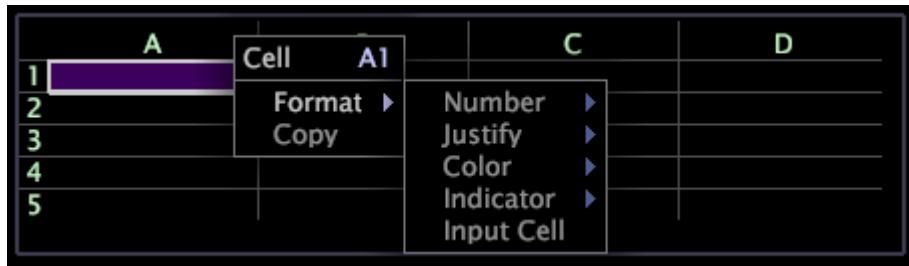
- Right-click the row that you want to delete and click **Delete Row**. The row is removed.

To resize a column

1. In the column header row, place your cursor over the dividing line to the right of the column whose size you want to change.
2. Click and drag to the right to increase the width of the column, or to the left to reduce the width of the column.

To format a cell

-
1. Right-click the cell and click **Format**.



2. Click the desired format in the menu of available options:

Menu Option	Description
Number	Applies the selected numeric format to your data, such as time, date, percentage, or decimal. Click Default to remove the selected formatting.
Justify	Justifies the data within the cell to the left, center, or right. The default justification is left. Click Default to remove the selected formatting.
Color	Applies the selected font color to the data within the cell. The default font color is white. Click Default to remove the selected formatting.
Indicator	Creates a metric indicator using this cell. For more information, see Create a metric indicator on page 429. Click Default to remove the selected formatting.
Input Cell	Makes the selected cell an input cell. For more information, see Create an input cell on page 431. Click Default to remove the selected formatting.

Keyboard shortcuts

Within worksheets you can use many of the basic editing keyboard shortcuts that you can use in any text editor, such as Notepad or Microsoft Word.

The following table lists the basic keyboard shortcuts that you can use when entering data into a worksheet.

Shortcut	Description
Arrow Keys	Move from cell to cell in your worksheet using the up, down, left and right arrow keys.
F2	Edit the cell by placing your cursor in the cell that you have selected.
Enter	Completes the editing of the cell that you have selected. Your cursor is removed from the cell and the cell contents reflect your editing.
Esc	Cancel the editing of the cell that you have selected. Your cursor is removed from the cell and the cell contents revert to what they were before you began editing.
Delete	Delete the contents of the cell(s).
Ctrl+A	Select the contents of the cell.

Shortcut	Description
Ctrl+c	Copy the contents of the cell(s).
Ctrl+x	Copy and remove the contents of the cell(s).
Shift+Delete	
Ctrl+v	Paste the contents of cell(s) that you have copied into the selected cell(s).
Shift+Insert	
Ctrl+z	Undo typing.
Ctrl+Shift+z	Redo typing.

Worksheets

Navigation title: Worksheets

Data Workbench provides worksheets for creating ad hoc filters and metrics for the current visualization.

A worksheet is similar to an ordinary spreadsheet, but unlike a spreadsheet, every cell containing a formula can define a metric.

	A	B	C	D
1	=Ratings/Users			
2				
3				
4				
5				
6				
7				

Unlike metrics created using the **Metric Editor**, metrics defined using worksheets do not appear in the metric menus, but can be saved and included in any of your workspaces.

For information about the **Metric Editor**, see [Work with derived metrics](#) on page 461.

Worksheet expressions

Navigation title:

Conceptual information about worksheet expressions and using cell references.

The following worksheet provides details about the visitors who view the Application Wizard page provided on the online application form of a bank's website.

	A	B	C	D
1		Apply Now	Apply Now with Wizard	Percentage of Wizard Viewers
2	Visitors	86,235	49,441	57.3
3	Referred Visitors	40,491	24,642	60.9
4	Referred Visitors from Ref A	5,964	3,294	55.2

- Column A shows a list of the categories of visitors being evaluated: visitors, referred visitors, and referred visitors from Referrer A.

-
- Column B shows the number of visitors in each category who viewed the Apply Now page.
 - Column C shows those visitors that viewed both the Apply Now and the Application Wizard pages.
 - Column D contains the percentages of Apply Now viewers in the three categories who also viewed the Application Wizard page.

The worksheet shows that approximately 55 percent of the visitors referred from Referrer A that viewed the Apply Now page also viewed the Application Wizard page.

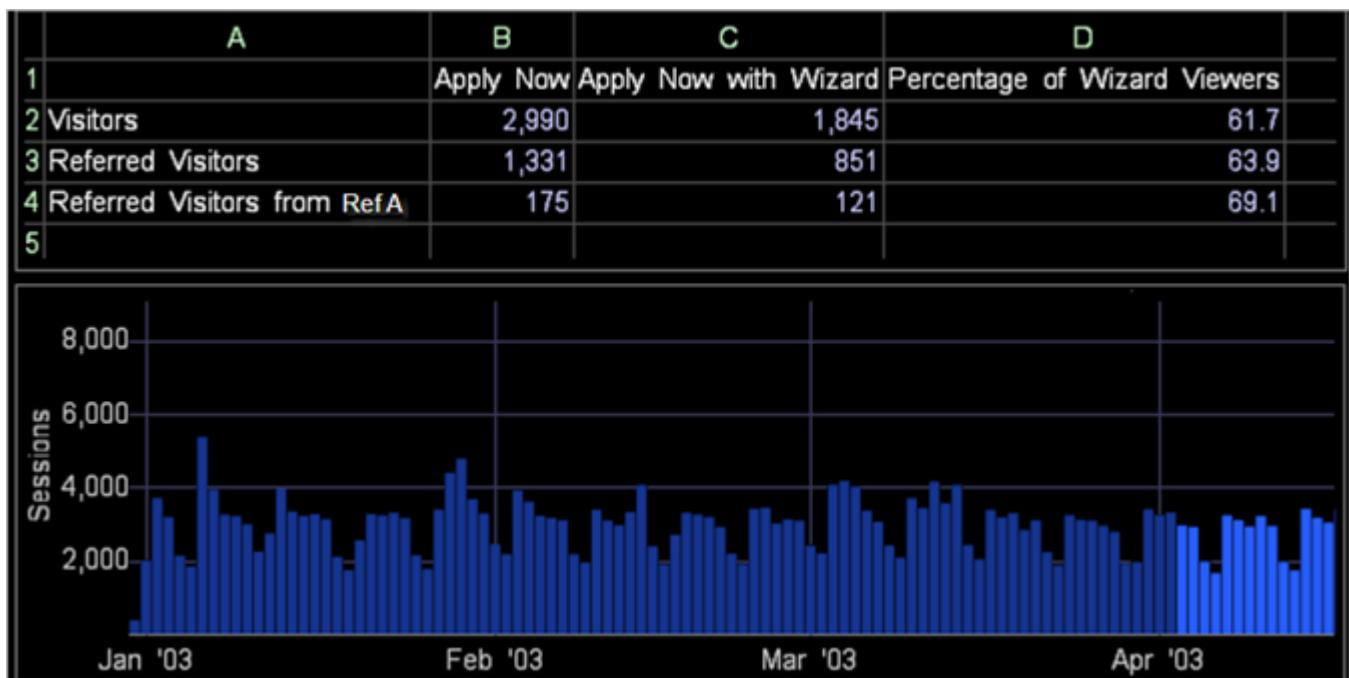
The following table provides sample formulas for the worksheet in the previous example:

Worksheet Cell	Formula
B2 Visitors who viewed the Apply Now page	=Visitors[Page="/applynow/default.asp"]
B3 Referred Visitors who viewed the Apply Now page	=Referred_Visitors[Page="/applynow/default.asp"]
B4 Referred Visitors from Referrer A who viewed the Apply Now page	=Referred_Visitors[Page="/applynow/default.asp" AND Referrer="Ref A"]
C2 Visitors who viewed the Apply Now page and the Application Wizard page	=Visitors[Page="/applynow/default.asp" AND Page="/applynow/appwizard.asp"]
C3 Referred Visitors who viewed the Apply Now page and the Application Wizard page	=Referred_Visitors[Page="/applynow/default.asp" AND Page="/applynow/appwizard.asp"]
C4 Referred Visitors from Referrer A who viewed the Apply Now page and the Application Wizard page	=Referred_Visitors[Page="/applynow/default.asp" AND Page="/applynow/appwizard.asp" AND Referrer="Ref A"]
D2 Percentage of Visitors who viewed the Apply Now page and the Application Wizard page	=C2/B2*100
D3 Percentage of Referred Visitors who viewed the Apply Now page and the Application Wizard page	=C3/B3*100
D4 Percentage of Referred Visitors from Referrer A who viewed the Apply Now page and the Application Wizard page	=C4/B4*100

As with other visualizations, worksheets update automatically when you make a selection in another visualization in the workspace. For more information about making selections, see [Make selections in visualizations](#) on page 281.

In the following web data example, several days of session data has been selected in the Sessions by Day visualization. The worksheet shows that during the selected timeframe, approximately 69 percent of the visitors from Referrer A who viewed the Apply Now page also viewed the Application Wizard page. Without this

selection (as shown in the example above), approximately 55 percent of the visitors from Referrer A viewed the Apply Now page as well as the Application Wizard page.



Using Cell References

You can substitute any string, whether on its own or within another expression in the worksheet, with a cell reference.

- **Simple cell reference:** Cell A2 contains the text `Visitors`, which is used as a heading. Cell B2 contains `=eval(A1)`, which evaluates to `=Visitors`.
- **Filter cell reference:** Cell A5 contains yesterday's date. Cell B5 contains `Visitors[Day=A5]`, which evaluates to the number of `Visitors` yesterday.
- **Concatenated cell reference:** Cell A5 contains today's date and Cell A6 contains the 08:00 to 08:59 one-hour period of time. Cell B6 contains `Visitors[Hour=A5+" "+A6]`, which evaluates to the number of `Visitors` today between 8:00 AM and 9:00 AM.

Create a metric indicator

Navigation title:

You can use worksheets to indicate that a metric has reached a defined threshold.

In addition, you can use Report to automatically generate and distribute a report when a metric reaches a defined threshold within a specified time frame.

For more information about Report, see the *Data Workbench Report Guide*.

- [Up or down indicator](#) on page 430
- [Check indicator](#) on page 431

To create a metric indicator using a worksheet

1. Define the contents of the cells of the worksheet.
 - a. In Column A, enter the name of the desired metric (for example, `Visitors`).
 - b. In Column B, enter the value of the desired metric (for example, `=Visitors`).

-
- c. In Column C, enter the low threshold of the metric.
 - d. In Column D, enter the high threshold of the metric.
 - e. In Column E, enter an appropriate formula. For examples, see [Up or down indicator](#) on page 430 or [Check indicator](#) on page 431.
 - f. In the formula cell (Column E), right-click and click **Format > Indicator**, then click one of the following:
 - **None**: Lists the exact calculation instead of an indicator.
 - **Check**: Uses a check mark or an X to indicate that the value is either above or below the threshold you set, depending on your formula. See [Check indicator](#) on page 431.
 - **Up or Down**: Uses an up or down arrow to indicate whether the value is below the low threshold (down arrow), above the high threshold (up arrow), or between the low and high thresholds (blank). See [Up or down indicator](#) on page 430.
2. Repeat Step 1 for other metrics for which you want to create indicators.

The resulting worksheet would look something like the following example:

A	B	C	D	E	F
1	Metric Value	Low Threshold	High Threshold	Up/Down Indicator	
2 Page Views	679,701	400,000	600,000	↑	
3 Visitors	43,080	50,000	75,000	↓	
4					
5					

Up or down indicator

For the **Up or Down indicator**, use the following formula:

$(\text{metric value} - \text{low threshold}) / (\text{high threshold} - \text{low threshold}) * 2 - 1$

For example: $=(\text{b2}-\text{c2}) / (\text{d2}-\text{c2}) * 2 - 1$

Three outcomes are possible for each metric when using this formula with the **Up or Down indicator**:

- If the metric value is between the low and high thresholds, the formula evaluates to a number between -1 and 1 (exclusively). The up or down arrow does not display in the worksheet.
- If the metric value is less than or equal to the low threshold, the formula evaluates to a value less than or equal to -1. The metric indicator changes to a down arrow.
- If the metric value is greater than or equal to the high threshold, the formula evaluates to a number greater than or equal to 1. The metric indicator changes to an up arrow.

The following worksheet illustrates what the example formula $=(\text{b2}-\text{c2}) / (\text{d2}-\text{c2}) * 2 - 1$ would display:

Alert Sheet					
A	B	C	D	E	F
1 Metric	Value	Low Threshold	High Threshold	Alert	
2 Visitors	631,762	200,000	600,000	↑	
3 Sessions	1,990,462	2,000,000	3,000,000	↓	
4					
5					

Check indicator

For the **Check indicator**, you use a formula that indicates whether you want to be notified when the metric value is above or below the threshold you specify. For example:

- If you want to be notified when the value is below the threshold you set, you could use the following format:

- $\text{threshold} - \text{metric}$

For example: $= (c2 - b2)$

- If you want to be notified when the value is above the threshold you set, you could use the following formula:

- $\text{metric} - \text{threshold}$

For example: $= (b3 - c3)$

When a check mark displays, the formula evaluated to a positive number. When an X displays, the formula evaluated to a negative number.

There are two possible outcomes for each metric when using the **Check indicator**:

- If the formula indicates that keeping the metric value above the threshold is desirable, a check mark displays when the metric value is greater than or equal to the threshold, and an X displays when the value is less than the threshold.
- If the formula indicates that keeping the metric value below the threshold is desirable, a check mark displays when the metric value is less than or equal to the threshold, and an X displays when the value is greater than the threshold.

The following worksheet illustrates what the example formulas $= (c2 - b2)$ and $= (b3 - c3)$ would display:

Alert Sheet					
	A	B	C	D	E
1	Metric	Value	Low Threshold	High Threshold	Alert
2	Visitors	631,762	1,000,000		✓
3	Sessions	1,990,462	2,000,000		✗
4					
5					

Create an input cell

Navigation title:

How to create an input cell.

When cell B1 is formatted as an input cell (by right-clicking the cell and clicking **Format > Input Cell**) and you lock the visualization (by right-clicking its top border and clicking **Locked**), you can change the value of the input cell to view results “on-the-fly.”

The following example shows a worksheet displaying the formula for an input cell.

	A	B
1	Referrer	Referrer A
2	Referred Visitors	=Referred_Visitors[Referrer=B1]
3		

This example shows the same worksheet displaying the result of the input cell formula.

	A	B	C
1	Referrer	Referrer A	
2	Referred Visitors		5,956
3			

In both examples, Row 1 shows the referrer currently being considered, Referrer A, while Row 2 shows the number of visitors who were referred from Referrer A.

After you format B1 and lock the visualization, you can change the value of the Referrer listed in B1 by simply typing the new value:



Use callouts with worksheets

Navigation title:

To make the information in a worksheet more clear, you can add a callout to bring attention to a particular metric.

By default, Data Workbench provides the following types of metric callouts:

- Annotation
- Blank Bars
- Blank Line Graph
- Blank Scatter Plot

To add a new visualization to your workspace that includes a metric that you calculated in your worksheet, you add a metric callout as a bar graph, line graph, or scatter plot callout. Metric callouts function as selections and, therefore, affect any other visualizations within the workspace.

You can change the default dimensions for metric callouts as well as add new types of callouts by configuring the callout files stored in the *profile name\Context\Metric Callout* folder of the Server installation folder. See [Configure a callout](#) on page 500.

Export to Microsoft Excel

Navigation title:

Link to information about exporting workspaces to Microsoft Excel.

For information about exporting workspaces and individual windows to Microsoft Excel, see [Export a workspace](#) on page 220.

Imagery layers

Navigation title:

Information about creating and implementing layers.

About imagery layers

Navigation title:

Conceptual information about imagery layers.

Types of imagery layers

You can view the following types of imagery layers in Data Workbench:

- **Terrain image layer:** This type of layer displays terrain imagery of the Earth, over which geographical data can be displayed. The globe visualization in is an example of a terrain image layer. See [Terrain image layers](#) on page 434.
- **Element point layer:** This type of layer displays one point on the globe for each element of a dimension. See [Element point layers](#) on page 439.
- **Vector layer:** This type of layer displays vector data (line art) on the globe. See [Working with vector layers](#) on page 443.

In Data Workbench, you can select which of the these layers you want to display for a particular analysis task.

Geography profile layers

The **Geography** profile provides you with a set of default imagery layers, which are stored in the Profiles\Geography\Maps folder within the Data Workbench server installation directory:

- **Blue Marble 2km:** This terrain image layer creates a 3-D map of the world, which is what displays when you add the globe visualization to a workspace. When this layer is not selected, the globe is not visible but the other layers still display. The Blue Marble 2km.layer file references the Blue Marble 2km.tsi file.
- **Zip Points:** This element point layer enables you to map locations in your dataset using a United States ZIP Code. The Zip Points.txt lookup file (provided by Adobe) contains a list of all United States ZIP Codes and each ZIP Code's latitude and longitude. The Zip Points.layer file references the Zip Points.txt file and the Zipcode.dim file and contains the configuration parameters needed to display the locations on the globe. Each element of the ZIP Code dimension (Zipcode.dim) that you define within your dataset is mapped on the globe using the latitude and longitude listed for that ZIP Code in the Zip Points.txt lookup file.

For information about defining dimensions, see the *Dataset Configuration Guide*.

- **Boundaries:** This vector layer provides the major world political boundaries, such as countries, as well as the boundaries of natural physical features of the Earth, such as lakes and islands. The Boundaries.layer file references one or more of the mwcoast.vec, mwisland.vec, mwlake.vec, mwnation.vec, mwriver.vec, mwstate.vec, US states.vec, and world boundaries.vec files.
- **IP Coordinates:** This element point layer uses dynamic points to enable you to map locations in your dataset using IP addresses. The IP Coordinates.layer file references the Coordinates dimension (Coordinates.dim) and specifies the Visitors metric as the metric to use to determine the size of the points on the globe for each coordinate.

Your **Geography** profile or other profiles in your installation may contain additional imagery layers that Adobe provided or your company created.

Create a new layer

You can create new imagery layers by copying the appropriate type of layer file included in the **Geography** profile into any Profiles\profile name\Maps folder, then renaming and editing the file as appropriate. All new layers must meet the following requirements:

- The .layer file must adhere to the format of one of the supported layer types.
- The .layer file must reference the appropriate lookup and dimension files, if necessary.
- The referenced lookup file also must be stored within the Data Workbench server installation directory, and its path must be specified accurately in the .layer file.

For more information about the format and parameters for each type of layer file and its associated files, see the section in this chapter for the appropriate layer type.

Terrain image layers

Navigation title:

A terrain image layer displays terrain imagery of the Earth.

Terrain image layers are stored in the **Geography** profile in a custom format. These image layers can be generated by Adobe, or the Data Workbench server can transform your user-supplied terrain imagery into terrain layers suitable for use on the globe visualization.



Note: To work with **terrain image layers**, you must install the `Terrain_Images.cfg` file provided by Adobe.

To define a terrain image layer, you must have the following:

- **One or more terrain image files** containing the images to be displayed on the globe.
- A **Terrain_Images.cfg** file that specifies the terrain image file(s) to be used for the layer(s). The `Terrain_Images.cfg` file enables you to add one or more sources to create a **terrain image layer**. The format of your terrain image file determines the type of source that you should add. The following table provides descriptions of the available terrain image layer sources, including the supported terrain image file formats:

Type	Description
Raw unprojected bitmap	<p>Creates terrain image layers from 24-bit headerless RGB files that are latitude-longitude aligned (unprojected), where north is the top of the image, and east is the right.</p> <p>Supported image format(s): RAW</p> <p> Note: This source requires projection information. For information about projection formats, see Specify projection information for terrain images on page 437.</p>
General image, unprojected	<p>Creates terrain image layers from 24-bit, latitude-longitude aligned (unprojected) image formats, where north is the top of the image, and east is the right.</p> <p>Supported image format(s): BMP, JPG, PNG, TIFF</p> <p> Note: This source requires projection information. For information about projection formats, see Specify projection information for terrain images on page 437.</p>
Image with embedded projection	<p>Creates terrain image layers from image formats that embed geodetic data in the image file. The projection information is extracted from the image.</p> <p>Supported image format(s): Erdas (IMG), GeoTIFF</p> <p> Note: This source usually does not require projection information but supports the addition of such information if needed. For information about projection formats, see Specify projection information for terrain images on page 437.</p>

To define a terrain image layer

1. In Data Workbench, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the **Servers Manager** workspace.
2. Within the **Servers Manager** window, right-click the icon of the desired Data Workbench server and click **Server Files**.
3. In the **Server Files Manager**, click **Components** to view its contents. The `Terrain_Images.cfg` file is located within this directory.
4. Right-click the check mark in the server name column for `Terrain_Images.cfg`, then click **Make Local**. A check mark appears in the **Temp** column for `Terrain_Images.cfg`.

-
5. Right-click the newly created check mark in the **Temp** column and click **Open > from the workbench**. The **Terrain Images.cfg** window appears.
 6. In the **Terrain Images** window, click **component** to view its contents.
 7. Right-click **Sources** > **Add new** and choose one of the following source types:
 - **Raw unprojected bitmap**. (Once added, this source type is labeled **RawTerrainSource** in the **Terrain Images** window.)
 - **General image, unprojected**. (Once added, this source type is labeled **GDALTerrainSource** in the **Terrain Images** window.)
 - **Image with embedded projection**. (Once added, this source type is labeled **GDALTerrainSource** in the **Terrain Images** window.)
 8. Edit the parameters for the source as necessary using the following sample file and table of parameters as guides.

-Terrain Images.cfg (modified)	
	TerrainImageBuilder
-component	
-Sources	
-0	RawTerrainSource
Projection Info	LatLonProjection
Lat0	90
Lat1	-90
Lon0	-180
Lon1	180
Gamma	0.75
Height	10800
Source Image	earth.raw
Tile Compression Quality	80
Tile Compressor	JPEG
Width	21600
-1	GDALTerrainSource
Projection Info	LatLonProjection
Lat0	90
Lat1	-90
Lon0	-180
Lon1	180
Gamma	1.0
Source Image	earth.png
Tile Compression Quality	100
Tile Compressor	
-2	GDALTerrainSource
Gamma	1.6
Source Image	2007-*.*.img
Tile Compression Quality	75
Tile Compressor	JPEG
Source Image Location	MapSources\
Temp Image Storage	Temp\
Write Layers To	Profiles\Geography\Maps\

Terrain Images.cfg: Sources

Parameter	Description
Gamma	Optional for all sources. Specifies gamma correction to be applied to the source image. This may be desirable due to the fact that Data Workbench normally runs with a high gamma setting. The default value is 1.
Height	Required for raw unprojected bitmap images. The height of the source image in pixels.
Projection Info	Required for raw unprojected bitmap images and general images, unprojected, but supported for images with embedded projection. Data Workbench supports latitude-longitude projections and Transverse Mercator (TM) projections for terrain image layers. The default projection format is the latitude-longitude projection (LatLonProjection). For information about projection formats, see Specify projection information for terrain images on page 437.
Source Image	Required for all sources. The name of the source image file. This may be a file name or a wildcard pattern. Using a pattern may be useful if, for instance, images for the same region at different dates are uploaded, with no change in the associated metadata. Therefore a pattern like <code>Tysons Corner * .raw</code> would create layers from <code>Tysons Corner 050211.raw</code> , <code>Tysons Corner 050218.raw</code> , and so on as new images are added, with no additional configuration necessary if the parameters for the files are otherwise identical.
Tile Compression Quality	Optional for all sources. For JPEG compression, an integer from 0 to 100 specifying how to balance image size and quality. (The default value is zero.) A higher number results in better image quality, but produces larger images and longer download times for Data Workbench users.  Note: Compressing images below 70 may result in image degradation.
Tile Compressor	Optional for all sources. Specifies what compression method is used to write output files. The only currently supported methods are RAWRGB (the default, resulting in no compression) and JPEG. Use JPEG compression to reduce the size of layers that are transmitted during profile synchronization.
Width	Required for raw unprojected bitmap images. The width of the source image in pixels.

9. Edit the Source Image Location, Temp Image Storage, and Write Layers To parameters using the following table as a guide. These parameters apply to all of the terrain image sources that you define in the **Sources** section of this file.

Terrain Images.cfg: Additional Parameters

Parameter	Description
Source Image Location	Required. The directory that is scanned for images to translate into terrain layers. If it is not an absolute path, it is interpreted relative to the Data Workbench server installation directory.
Temp Image Storage	Optional. The name of a directory that is used for storage of temporary files used in the translation of source images to terrain layers. If it is not an absolute path, it is interpreted relative to the Data Workbench server installation directory. The default location is the Temp directory.
Write Layers To	Required. The directory to which terrain layers is output. Ordinarily, this is the Maps subdirectory of a profile directory, so that the Globe visualization can find the layers.

10. Save the file by right-clicking (**modified**) at the top of the window and clicking **Save**.
11. To save an updated file to the Data Workbench server computer, in the **Server Files Manager**, right-click the check mark for **Terrain Images.cfg** in the **Temp** column, then click **Save to > <server name>**.

Specify projection information for terrain images

Navigation title:

Data workbench supports both latitude-longitude projections and Universal Transverse Mercator (UTM) projections for all terrain image layer sources.

Projection information is required for raw unprojected bitmaps and general images, unprojected. You can specify projection information for images with embedded projection information, though it is usually not required because the parameters of the projection are determined automatically from geodetic data embedded in the image itself. The following sections provide details about specifying these projection formats in the `Terrain_Images.cfg` file.

Latitude-Longitude Projections

The latitude-longitude projection format (`LatLonProjection`) in the `Terrain_Images.cfg` file is defined by four parameters for latitude and longitude.

To specify a `LatLonProjection` for unprojected images (raw unprojected bitmaps and general images, unprojected), you can enter settings for the `LatLonProjection` within the `Terrain_Images.cfg` window in Data Workbench.

To specify a `LatLonProjection` for images with embedded projection information, you must open the `Terrain_Images.cfg` file in a text editor such as Notepad, set the `Projection Info` parameter to `LatLonProjection`, and add settings for the `LatLonProjection`.

To specify a `LatLonProjection` for unprojected images

1. Open the `Terrain_Images.cfg` file in Data Workbench and add a terrain image layer source as described in [Terrain image layers](#) on page 434.
2. Edit the `Projection Info` parameters using the following parameters table as a guide:

LatLonProjection Parameters

Parameter	Description
Lat0	The latitude of the top edge of the image, in degrees, where 90 is the North Pole and -90 is the South Pole.
Lat1	The latitude of the bottom edge of the image.
Lon0	The longitude of the left-hand edge of the image, in degrees, where positive numbers are east and negative numbers are west longitudes.
Lon1	The longitude of the right-hand edge of the image.

3. Save the file by right-clicking (**modified**) at the top of the window and clicking **Save**.
4. To save the locally made changes to the Data Workbench server computer, in the **Server Files Manager**, right-click the check mark for `Terrain_Images.cfg` in the **Temp** column, then click **Save to > <server name>**.

To specify a `LatLonProjection` for images within embedded projection information

In the **Server Files Manager**, click **Components** to view its contents. The `Terrain_Images.cfg` file is located within this directory.

Right-click the check mark in the server name column for `Terrain_Images.cfg`, then click **Make Local**. A check mark appears in the **Temp** column for `Terrain_Images.cfg`.

Right-click the newly created check mark in the **Temp** column and click **Open > in Notepad**. The `Terrain_Images.cfg` file appears in a Notepad window.

Edit the Projection Info parameters using the following sample file fragment as a guide. Be sure to specify the projection type as highlighted below. For descriptions of the parameters, see the LatLonProjection Parameters table in the previous procedure.

```
Projection Info = LatLonProjection:  
    Lat0 = double: 90  
    Lat1 = double: -90  
    Lon0 = double: -180  
    Lon1 = double: 180
```

Universal Transverse Mercator Projections

The Universal Transverse Mercator (UTM) projection is defined by eight parameters. When specifying a Universal Transverse Mercator projection for a terrain image layer, your terrain image files must be aligned with false (projected) north towards the top of the image, and false east to the right of the image.

To specify a UTM projection for any terrain image source, you must open the `Terrain_Images.cfg` file in a text editor such as Notepad, set the `Projection Info` parameter to “`TransverseMercatorProjection`”, and add settings for the UTM projection.

To specify a Universal Transverse Mercator projection

1. In the **Server Files Manager**, click **Components** to view its contents. The `Terrain_Images.cfg` file is located within this directory.
2. Right-click the check mark in the server name column for `Terrain_Images.cfg`, then click **Make Local**. A check mark appears in the **Temp** column for `Terrain_Images.cfg`.
3. Right-click the newly created check mark in the **Temp** column and click **Open > in Notepad**. The `Terrain_Images.cfg` file appears in a Notepad window.
4. Edit the `Projection Info` parameters using the following sample file fragment and parameters table as guides. Be sure to specify the projection type as highlighted below.

```
Projection Info = TransverseMercatorProjection:  
    Ellipsoid Inverse Flattening = double: 294.9786982139006  
    Ellipsoid Semimajor Axis = double: 6378206.4000000004  
    False Easting = double: 500000  
    False Northing = double: 0  
    Northwest Corner Coordinates = v3d: (550339, 5.42059e+006, 0)  
    Prime Meridian = double: -123  
    Scale Factor = double: 0.9996  
    Southeast Corner Coordinates = v3d: (555099, 5.41356e+006, 0)
```

TransverseMercatorProjection Parameters

Parameter	Description
Ellipsoid Inverse Flattening, Ellipsoid Semimajor Axis	The parameters of the ellipsoid used for the projection. The semimajor axis is specified in meters.
False Easting	The false easting of the central meridian of the projection, in meters. For UTM, this is always 500,000.
False Northing	The false northing of the equator in the projection, in meters. For UTM, this is 0 for northern hemisphere zones and 10,000 for southern hemisphere zones.
Northwest Corner Coordinates, Southeast Corner Coordinates	The coordinates (in projected meters) of the top left and bottom right corners of the image.

Parameter	Description
Prime Meridian	The longitude of the central meridian of the projection, specified in degrees east of Greenwich. Negative numbers may be used to specify degrees west.
Scale Factor	The ratio of the radius of the projection cylinder to the semimajor axis of the ellipsoid. For Universal Transverse Mercator (UTM) projections, this is always 0.9996.

Make a new terrain image layer available

Navigation title:

Steps to make any terrain layer available to display on the globe visualization.

1. In the Profiles\profile name\Maps folder within the Data Workbench server installation directory, place the layer file and the supporting image files.
2. Edit the order .txt file in the Profiles\profile name\Maps folder to reflect the order in which you want the layers to display. By default, layers appear in lexicographic order by their names.



Note: When editing the order .txt file, take care not to cover up map layers that you want to show.

For more information about using order .txt files, see the Configuring Interface and Analysis Features chapter of the *Data Workbench User Guide*.

3. In Data Workbench, select the desired profile by right-clicking the workspace title bar and clicking **Switch Profile > <profile name>**.
4. Right-click the workspace title bar and click **Work Online**. An X appears next to Work Online.
5. Open a workspace and on a globe visualization, right-click and select the new layer. An X appears next to the layer name.

Element point layers

Navigation title:

In Data Workbench, an element point layer displays one point on the globe for each element of a dimension.

The size of the point displayed on the globe is determined by evaluating a specified metric over the dimension whose elements you want to plot. Therefore, larger metric values correspond to larger points on the globe.

Define element point layers referencing lookup files

Navigation title:

When creating an element point layer that references a lookup file to obtain latitude and longitude data, the location of the point is obtained by retrieving each element and its associated latitude and longitude from the lookup file.



Note: Instead of using a lookup file, you can use the Dynamic Points functionality, which embeds the latitude and longitude of a location in the name of each element of a dimension. See [Define element point layers using dynamic points](#) on page 441.

To define an element point layer that references a lookup file, you must create or already have available the following:

- A **dimension** defined in the Transformation.cfg file or a **transformation dataset include** file. For information about transformation configuration files, see the *Dataset Configuration Guide*.
- A **lookup file** containing the data used to plot each data point. This file must contain at least three columns of data for each data point: the key, the longitude, and the latitude. For more information about the required format of the lookup file, see [Element point layer file format](#) on page 440.

-
- A **layer file** that specifies the location of the lookup file and identifies the related dimension and metric as well as the key, longitude, and latitude column names in the lookup file. For more information about the required format of the layer file, see [Element point layer file format](#) on page 440.



Note: The Zip Points.layer file, provided with the **Geography** profile, is an element point layer that identifies the Zipcode.dim file, the Sessions.metric file, the Zip Points.txt lookup file, and the names of the key, longitude, latitude, and name columns in the lookup file.

Element point lookup file format

The element point layer lookup file must contain at least the following three columns:

- **Key column:** This column should contain common key data, which enables the Data Workbench server to connect the data in the lookup file to that in the dataset. The **key** column must be the first column in the lookup file. Each row in this column identifies an element of the dimension.
- **Longitude column:** This column should contain the longitude for each data point in the **Key** column.
- **Latitude column:** This column should contain the latitude for each data point in the **Key** column.
- **Name column (Optional):** If you want to specify a name to be displayed on the map for each data point, you can include a **Name** column in the lookup file.

Each row in the Zip Points.txt lookup file contains a ZIP Code in the first column followed by the longitude, latitude, and associated city name.

```
tude, and associated city name.  
ZIP_CODE LATITUDE LONGITUDE NAME  
00210 +43.005895 -071.013202 PORTSMOUTH, NH  
00211 +43.005895 -071.013202 PORTSMOUTH, NH  
...
```

Element point layer file format

Each element point layer .layer file that references a lookup file must be formatted using the following template:

```
Layer = ElementPointLayer:  
  Data Paths = vector: 1 items  
    0 = Path: Maps\\Lookup File Name.txt  
  Longitude Column = string: Longitude Column Name  
  Latitude Column = string: Latitude Column Name  
  Name Column = string: Location Column Name  
  Key Column = string: Key Column Name  
  Dimension = ref: wdata/model/dim/Dimension Name  
  Metric = ref: wdata/model/metric/Metric Name  
  Scale = double: Scale  
  Color = v3d: RGB Color Vector  
  Rendering Mode = int: Mode Number
```

Element point layer parameters: lookup files

Parameter	Description
Data Paths	Path to the lookup file containing latitude and longitude data.
Longitude Column	The name of the column in the lookup file containing the longitude data.
Latitude Column	The name of the column in the lookup file containing the latitude data.
Name Column	Optional. The name of the column in the lookup file containing the names of the locations represented by the latitude and longitude data.

Parameter	Description
Key Column	The name of the column in the lookup file containing the common key data, which enables the Data Workbench server to integrate the data in the lookup file into the dataset. This must be the first column in the lookup file. Each row in this column is an element of a dimension. This dimension must be defined in the Transformation.cfg file or a transformation dataset include file and specified in the Dimension parameter of this file. For more information about transformation configuration files, see the <i>Dataset Configuration Guide</i> .
Dimension	The name of the dimension (defined in a transformation configuration file) containing elements that correspond to the data rows in the Key column.
Metric	The name of the metric that is evaluated over the dimension specified in the Dimension parameter.
Scale	Optional. Value used to size the points in the layer. The default value is 100. Larger values make the points bigger, and smaller values make them smaller.
Color	Optional. The RGB color vector, which is expressed as (red,green,blue). For each color in the vector, you can enter a value from 0.0 to 1.0. For example, (1.0, 0.0, 0.0) is bright red, and (0.5, 0.5, 0.5) is gray.
Rendering Mode	Optional. Integer value representing the rendering mode to use for the layer. The three available modes are as follows: <ul style="list-style-type: none"> • Rendering Mode 1. Points size is defined in screen space (points stay a constant size relative to the computer screen). Points are rendered using polygons, so there is no upper limit on point size. This is the default rendering mode. • Rendering Mode 2. Point size is defined in world space (points stay a constant size relative to the globe). Points are rendered using polygons, so there is no upper limit on point size. • Rendering Mode 3. Point size is defined in screen space. Points are rendered using OpenGL smooth points.

The `Zip Points.layer` file is formatted as follows:

```
Layer = ElementPointLayer:
  Data Paths = vector<string>: 1 items
    0 = Path: Maps\\Zip Points.txt
  Longitude Column = string: LONGITUDE
  Latitude Column = string: LATITUDE
  Name Column = string: NAME
  Key Column = string: ZIP_CODE
  Dimension = ref: wdata/model/dim/Zipcode
  Metric = ref: wdata/model/metric/Sessions
```

Define element point layers using dynamic points

Navigation title:

When creating an element point layer using dynamic points, the latitude and longitude data is embedded in each element of the dimension.

To define an element point layer using dynamic points, you must create or already have available the following:

- A dimension, defined in the **Transformation.cfg** file or a **transformation dataset include** file, in which each element contains the string “latitude,longitude” or “latitude,longitude,name.”

For steps to create a dimension, see the *Dataset Configuration Guide*.

- A layer file that specifies the related dimension.

For more information about the required format of the layer file, see [Element point layer file format](#) on page 442.



Note: When using **Dynamic Points**, it is essential to ensure that the cardinality of the dimension specified in the layer file is reasonable. If every row of a dataset has a different latitude and longitude, the dimension quickly fills up and most rows fall into a Small Elements element. Because the Small Elements element does not have a latitude and longitude, it does not appear on the globe.

Element point layer file format

Each element point layer file using dynamic points must be formatted using the following template:

```
Layer = ElementPointLayer:  
    Dimension = ref: wdata/model/dim/Dimension Name  
    Metric = ref: wdata/model/metric/Metric Name  
    Dynamic Points = bool: true  
    Scale = double: Scale  
    Color = v3d: RGB Color Vector  
    Rendering Mode = int: Mode Number
```

Element point layer parameters: dynamic points

Parameter	Description
Dimension	The name of the dimension (defined in a transformation configuration file), which must contain elements with the string “latitude,longitude” or “latitude,longitude,name” as shown in the following examples: <ul style="list-style-type: none">• 37.5181,-77.1903• 35.3317,-77.8126,Somewhere
Metric	The name of the metric that is evaluated over the dimension specified in the Dimension parameter.
Dynamic Points	Enables Dynamic Points. Set to true.
Scale	Optional. Value used to size the points in the layer. The default value is 100. Larger values make the points bigger, and smaller values make them smaller.
Color	Optional. The RGB color vector, which is expressed as (red,green,blue). For each color in the vector, you can enter a value from 0.0 to 1.0. For example, (1.0, 0.0, 0.0) is bright red, and (0.5, 0.5, 0.5) is gray.
Rendering Mode	Optional. Integer value representing the rendering mode to use for the layer. The three available modes are as follows: <ul style="list-style-type: none">• Rendering Mode 1. Points size is defined in screen space (points stay a constant size relative to the computer screen). Points are rendered using polygons, so there is no upper limit on point size. This is the default rendering mode.• Rendering Mode 2. Point size is defined in world space (points stay a constant size relative to the globe). Points are rendered using polygons, so there is no upper limit on point size.• Rendering Mode 3. Point size is defined in screen space. Points are rendered using OpenGL smooth points.

The IP Coordinates.layer file is formatted as follows:

```
Layer = ElementPointLayer:  
    Dimension = ref: wdata/model/dim/Coordinates  
    Metric = ref: wdata/model/metric/Visitors  
    Dynamic Points = bool: true
```

Make a new element point layer available

Navigation title:

Steps to make any element point layer available to display on the globe visualization.

-
1. In the Profiles\profile name\Maps folder within the Data Workbench server installation directory, place the layer file and its related lookup file.
 2. If you defined a new dimension for your element point layer and have not yet retransformed your dataset, retransform your dataset now.
 3. Edit the order.txt file in the Profiles\profile name\Maps folder to reflect the order in which you want the layers to display. By default, layers appear in lexicographic order by their names.



Note: When editing the order.txt file, take care not to cover up map layers that you want to show.

For more information about using order.txt files, see the Configuring Interface and Analysis Features chapter of the *Data Workbench User Guide*.

4. In Data Workbench, select the desired profile by right-clicking the workspace title bar and clicking **Switch Profile > <profile name>**.
5. Right-click the workspace title bar and click **Work Online**. An X appears next to Work Online.
6. Open a workspace and on a globe visualization, right-click and select the new layer. An X appears next to the layer name.

Working with vector layers

Navigation title:

In Data Workbench, a vector layer displays static vector data (line art) on the globe.

You can use vector layers to display outlines of geographic regions, such as states, countries, islands, or bodies of water.

Define vector layers referencing vector files

Navigation title:

You can create a vector layer that references one or more vector (.vec) file, which contains the data that defines the vectors to be drawn on the globe.

To define a vector layer that references one or more .vec files, you must have the following:

- **One or more .vec files** that contain the data used to draw the vectors on the globe.



Note: To obtain .vec files to use with your vector layers, contact Adobe.

- **A layer file** that specifies the location of the .vec files. For more information about the required format of the layer file, see [Vector layer file format](#) on page 443.



Note: The Boundaries.layer file, provided with the **Geography** profile, is a vector layer that references the mwnation.vec, mwstate.vec, mwcoast.vec, mwlake.vec, and mwisland.vec files.

Vector layer file format

Each vector layer file referencing .vec files must be formatted using the following template:

```
Layer = VectorLayer:  
  Vec Files = vector: n items  
    0 = string: Maps\\.vec file 1  
    1 = string: Maps\\.vec file 2  
    . . .  
    n-1 = string: Maps\\.vec file n  
  Color = v3d: color vector  
  Alpha = double: alpha
```

```

Width = double: width
Error Factor = double: error factor

```

Vector layer parameters: .vec files

Parameter	Description
Vec Files	Path(s) to the .vec file(s) containing the vector data.
Color	The RGB color vector, which is expressed as (red,green,blue). For each color in the vector, you can enter a value from 0.0 to 1.0. For example, (1.0, 0.0, 0.0) is bright red, and (0.5, 0.5, 0.5) is gray.
Alpha	Controls the transparency of the vectors shown on the globe. The range is 0 to 1, with 0 being the most transparent.
Width	Optional. Sets the width of the data in pixels. The recommended range is 1 to 4.
Error Factor	Controls how accurately the vectors are drawn. For larger values, the vectors are drawn less accurately but faster. The default value is 5.

The `Boundaries.layer` file is formatted as follows:

```

Boundaries.layer file is formatted as follows:
Layer = VectorLayer:
  Vec Files = vector: 5 items
    0 = string: Maps\\mwnation.vec
    1 = string: Maps\\mwstate.vec
    2 = string: Maps\\mwcoast.vec
    3 = string: Maps\\mwlake.vec
    4 = string: Maps\\mwisland.vec
  Color = v3d: (.5,.5,1)
  Alpha = double: .5
  Error Factor = double: 4

```

Vector layers referencing tab separated values files

Navigation title:

When creating a vector layer that references a tab separated values (.tsv) file, the vector data is obtained by retrieving drawing instructions as well as longitude and latitude data from the .tsv file.

To define a vector layer that references a .tsv files, you must have the following:

- A **.tsv file** that contains the data used to draw the vectors on the globe, including longitude and latitude data. For more information about the required format of the .tsv file, see [Vector TSV file format](#) on page 444.
- A **layer file** that specifies the location of the .tsv file. For more information about the required format of the layer file, see [Vector layer file format](#) on page 445.

Vector TSV file format

The .tsv file must contain the following three tab separated columns:

- **Begin:** This column should indicate whether to begin a new line. Values in this column can be either 0 (do not begin a new line) or 1 (begin a new line).
- **Longitude:** This column should contain longitude values.
- **Latitude:** This column should contain latitude values.



Note: Any additional columns are ignored.

Following is a sample .tsv file that contains data for a vector layer:

```

Begin  Longitude      Latitude
1      -77.256903    38.98191
0      -77.262703    38.985309
0      -77.259803    38.989409
0      -77.255903    38.991009
0      -77.255217    38.988962
0      -77.256458    38.985688
0      -77.256903    38.98191
1      -77.23540928  38.97669798
0      -77.234803    38.97631
0      -77.232268    38.979502
0      -77.231601    38.979917
0      -77.229992    38.979858

```

Vector layer file format

Each vector layer file referencing .tsv files must be formatted using the following template:

```

Layer = VectorLayer:
  TSV Files = vector: n items
    0 = string: Maps\\File Name.tsv
    1 = string: Maps\\File Name.tsv
    .
    .
    .
    n-1 = string: Maps\\File Name.tsv
  Color = v3d: color vector
  Alpha = double: alpha
  Width = double: width
  Error Factor = double: error factor

```

Vector layer parameters: .vec files

Parameter	Description
TSV Files	Path(s) to the .tsv file(s) containing the vector data. Example: Maps\\USVectorData.tsv
Color	The RGB color vector, which is expressed as (red,green,blue). For each color in the vector, you can enter a value from 0.0 to 1.0. For example, (1.0, 0.0, 0.0) is bright red, and (0.5, 0.5, 0.5) is gray.
Alpha	Controls the transparency of the vectors shown on the globe. The range is 0 to 1, with 0 being the most transparent.
Width	Optional. Sets the width of the data in pixels. The recommended range is 1 to 4.
Error Factor	Controls how accurately the vectors are drawn. For larger values, the vectors are drawn less accurately but faster. The default value is 5.

Make a new vector layer available

Navigation title:

Steps to make any vector layer available to display on the globe visualization.

1. In the Profiles\profile name\Maps folder within the Data Workbench server installation directory, place the layer file and the .vec or .tsv files.
2. Edit the order.txt file in the Profiles\profile name\Maps folder to reflect the order in which you want the layers to display. By default, layers appear in lexicographic order by their names.



Note: When editing the order.txt file, take care not to cover up map layers that you want to show.

For more information about using order.txt files, see [Customize a menu using order.txt files](#) on page 486.

-
3. In Insight, select the desired profile by right-clicking the workspace title bar and clicking **Switch Profile** > *<profile name>*.
 4. Right-click the workspace title bar and click **Work Online**. An X appears next to Work Online.
 5. Open a workspace and on a globe visualization, right-click and select the new layer. An X appears next to the layer name.

Administrative interfaces

Navigation title:

Information about the administrative interfaces available to manage and monitor your implementation of Data Workbench. Also information about opening administrative interfaces as well as information about several other interfaces.

Open administrative interfaces

Navigation title:

Because your implementation of Data Workbench can be fully customized and, therefore, may differ from what is documented in this guide, exact paths to each visualization are not provided in this guide.

All of the administrative interfaces can be opened using either or both of the following options:

- **Admin menu:** Right-click within a workspace, click **Admin**, then click the appropriate menu option.
- **Admin tab:** Click to open the appropriate workspace thumbnail on the **Admin** tab.

For information about administering the servers on which Adobe software applications are installed, see the *Server Products Installation and Administration Guide*.

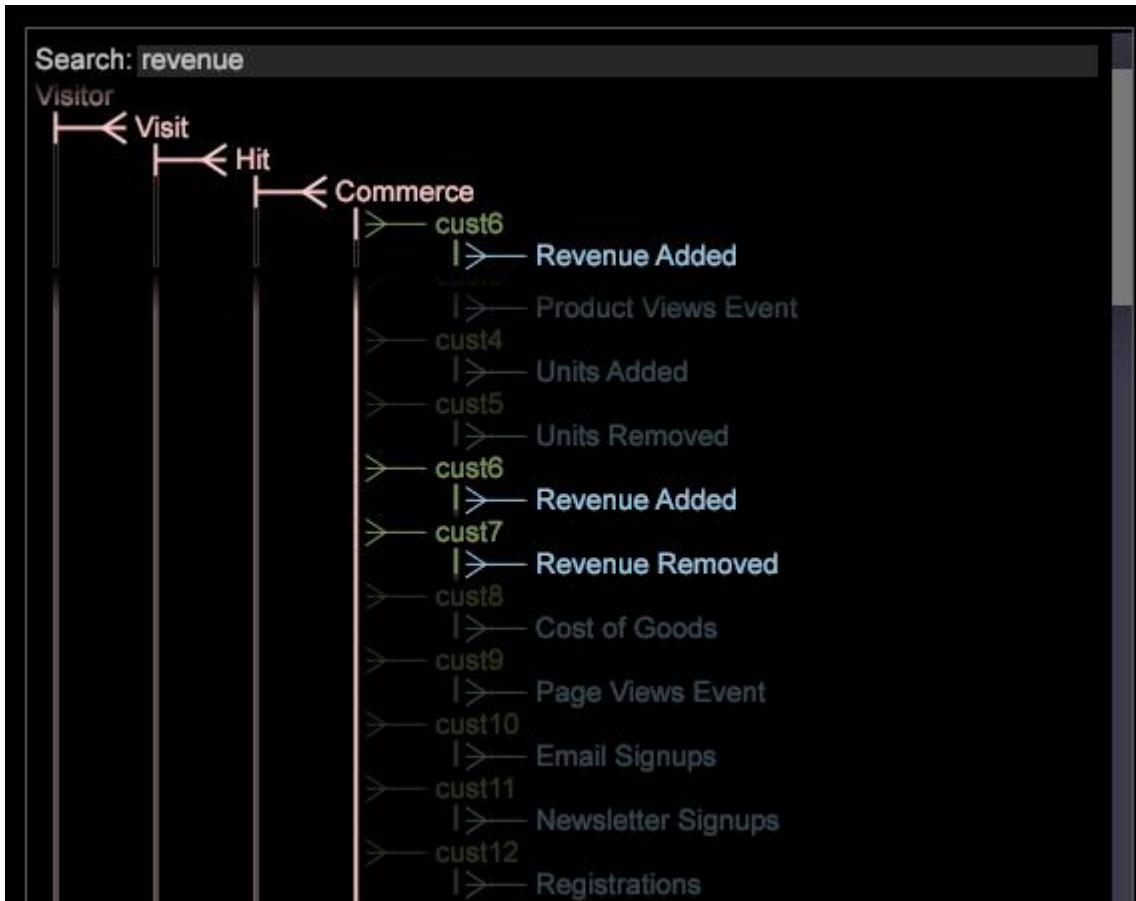
For information about administering specific products, see the appropriate product guide.

Dataset Schema interface

Navigation title:

The **Dataset Schema** interface displays the extended dimensions (countable, simple, many-to-many, numeric, denormal, and time dimensions) defined in any transformation dataset configuration file and provides a view of the relationships between those dimensions.

In addition, the **Dataset Schema** interface shows any derived dimensions that you have defined, as well as any extended dimensions that are configured to be hidden.



Note: You can search for dimensions from within the schema diagram. The name of the dimensions found by the search string are highlighted, and the lines of the parent class change color for any hits found in subordinate child dimensions. The Countable dimensions stay visible as you scroll to provide viewable hierarchy and context.

To interpret a dimension type using the Dataset Schema interface

The following table lists the dimension types and the colors in which their names appear in the **Dataset Schema** interface. Parents for the sample dimensions (from the example above) are noted as well.

Dataset Schema Interface

Dimension Type	Color	Sample Dimension and Parent
Countable	Pink	Visitor - In this schema, Visitor is a root countable dimension. Session - parent is Visitor
Denormal	Yellow	DenormalPage - parent is Page View
Derived	Blue	Next Page - parent is Page View
Many-to-Many	Pink and Green (The stem from the parent is pink, while the dimension name is green.)	Search Term - parent is Session
Numeric	Green	Exact Page Duration - parent is Page View. In this example, Exact Page Duration is a hidden numeric dimension. See the Hidden dimension type in this table.
Simple	Green	Page - parent is Page View

Dimension Type	Color	Sample Dimension and Parent
Time	Green	Hour - parent is Session
Hidden	Hidden dimensions are a darker version of the appropriate dimension type color. For example, a hidden numeric dimension is a darker, less bright green.	Exact Page Duration - parent is Page View

For more information about these dimension types, see the *Dataset Configuration Guide*.

To display the default visualization for a dimension

- In the **Dataset Schema** interface, click the desired dimension. The default visualization displays. For example, if the default visualization is a table displaying Sessions and the selected dimension and you click the URI dimension, Data Workbench displays a table with URI by Sessions.



Note: If you want to change the default visualization that displays, see *Dataset Schema interface* on page 446.

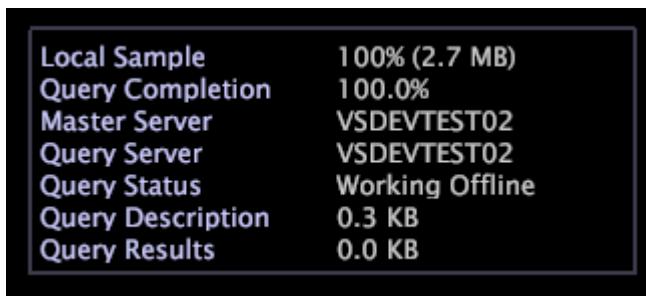
To display a specific visualization for a dimension

- In the **Dataset Schema** interface, right-click the desired dimension and click **Add Visualization > <visualization type>**.

Query Status Legend

Navigation title:

The **Query Status Legend** provides detailed information about the queries you are performing.



The following table lists the tasks that can be completed using the **Query Status Legend**.

Query status legend

To perform this task...	Do this...
To see the ratio of the local sample size to the overall dataset size followed by the size in MB of the local sample	View the Local Sample field.
To see the minimum percent complete for all current queries	View the Query Completion field.
To identify the master server for a client in a cluster	View the Master Server field.
To determine the size of a query	View the Query Results field. This field shows you the size of the current query and provides you with a way to see how much of an impact is made on your query size by adding (or removing) one metric, making a selection, or making any other change to your workspace.

Query Queue

Navigation title:

Normally, the Data Workbench server answers incoming user queries as they are received, and continues to provide results and real-time updates until the user is no longer requesting them.

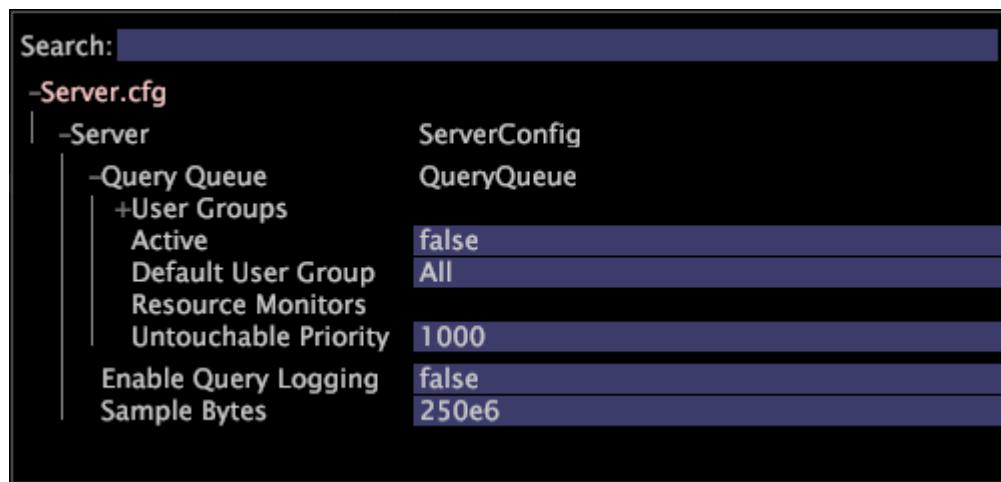
At times, particularly on systems with many Data Workbench users, the number of active queries require more system resources than are available from the server. **Query Queue** allows the server to place some queries temporarily on hold until the resources necessary to provide answers become available. The **Query Queue** also provides features to prioritize queries based on a variety of parameters, so that in case of resource contention, higher-priority queries are answered first.

Queries from a single client or report server are placed in a bunch and scheduled as a unit. You can configure resource monitors to limit the amount of certain system resources that are used by queries. When the monitored resources permit the scheduling of another query bunch, the highest-priority bunch is scheduled. Users whose queries are not scheduled yet, due to resource limitations, do not receive an error but are notified that their queries are queued, and the user can continue to work on the local sample.

The default configuration includes a simple configuration for the **Query Queue**, but leaves it disabled. Administrators can enable or disable the **Query Queue**, configure resource monitors to determine how much of various resources are used for querying, and configure complex prioritization policies for different users.

To configure the Server.cfg file for Query Queuing

1. Open Server.cfg by clicking **Admin > Profile Manager > Dataset**.
2. Right-click **Server.cfg** and make it local for editing.
3. Expand **Query Queue**.



4. Configure the following parameters:
 - **User Groups:** Lets you configure policies, users, and the queue priority. See [Query Queue User Groups](#) on page 450 for definitions.
 - **Active:** (Vector) Enables or disables the **Query Queue**. Valid values are true or false. The default setting is false.
 - **Default User Group:** (String) Type a name of the user group to which users are added, if they are not listed in any user group.
 - **Resource Monitors:** (Vector) Right-click to add a resource monitor. You can specify whether the **Query Queue** monitors memory or the number of queries. Right-click **Resource Monitor** to choose Memory Budget Monitor or Number of Queries Monitor. See [Query Queue Resource Monitors](#) on page 451 for more information.

- **Untouchable Priority:** (Int) Specifies that bunches with a priority greater than or equal to this value are never preempted to scheduling of higher priority bunches. Used in conjunction with the **Memory Budget Monitor** described in the [Query Queue User Groups](#) on page 450.

Query Queue User Groups

Navigation title:

Table that defines the User Group parameters.

Field	Type	Description
Name	string	A user-defined name of the user group, such as Analysts.
Policies	vector	Specifies a policy type. Right-click to choose Standard Policy or Daily Schedule.
Standard Policy		<p>A Standard Policy ensures that users with a low priority are incrementally moved up the queue and scheduled, even if higher priority users enter the queue. You can add multiple policies of the same type in a group, and their effect is cumulative.</p> <ul style="list-style-type: none"> • Priority Limit: The limit above which the priority is not incremented. The maximum priority value. You can use this value to keep the priorities generated by this policy in a specific range (for instance, so that priorities for some other group of users are always higher, or so they do not rise above the Untouchable Priority). <p>Standard Policy Increments</p> <p>The increment settings for the Standard Policy increase the priority of a query bunch as time passes. This does not force the bunches to be scheduled, but you can use these settings to prioritize users who have been waiting for a long time. The queued parameters affect queries that are currently queued (such as on hold due to insufficient resources to complete them). The scheduled parameters affect queries that are being answered. The priority of a query rises by the number specified in the appropriate increment and increment interval fields:</p> <ul style="list-style-type: none"> • Queued Increment: Sets the priority increment per update while queued. This setting ensures that low priority users are moved up the scheduling queue. • Queued Increment Interval: Sets the number of seconds between updates while queued. • Scheduled Increment: Sets the priority increment per update while scheduled. • Scheduled Increment Interval: Sets the number of seconds between updates while scheduled. <p> Note: Setting the increment and interval update rates higher for queued bunches than for scheduled bunches can cause oscillation. (For example, suppose you set the Queued Increment value to 100 and the Scheduled Increment to 0, and set the Queued Increment Interval value to 1 and the Untouchable Priority to be high. If two query bunches come in with a base priority of 0, and there are not enough resources to run both queries at the same time, then one of them is scheduled. After one second, the query that was not scheduled has a priority of 100, and preempts the one that was scheduled. After two more seconds, the one that was preempted now has a priority of 200, and the two switch places again. Neither query finishes, because every two seconds the query that is being computed is preempted so the other query can run.)</p>
Daily Schedule Policy		Lets you change the priority at specific times of the day. This schedule is useful for automated clients, such as Report Server , and when users of the system live in different time zones.
Changes	int	Right-click to add a scheduled priority change. The Change Time is the time of day at which the change occurs. The format is hour:minutes AM/PM. If AM or PM is not entered, the system uses military time.
Priority Limit	int	The maximum priority value resulting from a change. The Priority Change is the amount added to the priority. For example, a value of 0 returns to a default priority. Any other value results in a priority of the default priority plus this number.
Users	vector	<p>Lists the users that are members of the group.</p> <p>Name: The user's name as it appears in the Common Name field in the user's certificate.</p>

Field	Type	Description
		Extra Priority: Provides additional priority to the user group's base priority to determine the starting priority for that user.

Query Queue Resource Monitors

Navigation title:

The resource monitor vector contains the **Memory Budget Monitor** and the **Number of Queries Monitor**.

The following table describes the resource monitor fields used for query queuing.

Resource Monitors Parameters Table

Field	Type	Description
Memory Budget Monitor		Monitors the query memory used by the current user group. If the current usage is between the Low Threshold and the High Threshold, no new bunches are scheduled until memory usage returns to below the Low Threshold value, for example, as a result of users closing their workspaces. Scheduled bunches are allowed to grow.
High Threshold	double	The high threshold for memory usage (bytes). If the memory usage is above this value, no scheduling occurs, and the lowest priority-scheduled bunches are unscheduled one at a time, over a period of time, until the memory usage is brought to below this value.
Low Threshold	double	The low threshold for memory usage (bytes). If Memory Budget Monitor value is below this value, new bunches are allowed to be scheduled, and scheduled bunches are allowed to grow. For example bunches grow when a user adds a visualization to a workspace.
Reaction Time	double	The time constant for the smoothing of the memory usage estimate. Smoothing values avoids reaction to usage spikes.
Number of Queries Monitor		Monitors the total number of queries that are currently scheduled for the profile. This resource monitor lets you schedule bunches if the total number of queries remains below the value in the Low Threshold field. This monitor allows currently-scheduled bunches to grow if the total number of queries stays below the value in the High Threshold field. Additionally, this monitor removes bunches of a low priority in order to allow higher priority bunches to be scheduled or grow. However, this setting does not preempt bunches with a priority greater than specified in the Untouchable Priority field.
High Threshold		The high threshold for memory usage (bytes). If the memory usage is above this value, no scheduling occurs, and the lowest priority scheduled bunches are unscheduled one at a time, over a period of time, until the memory usage is brought to below this value.
Low Threshold		The low threshold for memory usage (bytes). If Memory Budget Monitor value is below this value, new bunches can be scheduled, and scheduled bunches can grow.

Processing Legend

Navigation title:

The **Processing Legend** provides detailed information about a particular server's data processing and transformation, enabling you to track the progress of data that is being reprocessed and retransformed.

Total Log Entries	928,319
Log Bytes Total	249,123,945
Total Decoded Log Entries	928,319
Log Bytes Read	249,123,945
Log Processing Progress	100.0%
Total Filtered Log Entries	418,218
Transformation Progress	100.0%
Total Processed Log Entries	418,218

The following table lists the tasks that can be completed using the **Processing Legend**.

To perform this task...	Do this...
To view the total size of all of your data	Review the values in the Total Log Entries and Log Bytes Total fields.
To check whether filtering is working	Review the values in the Total Filtered Log Entries fields. If the value is 0, filtering is not working and you must check your configuration to address the problem.
To check the progress of log processing	Review the value in the Log Processing Progress field. This percentage indicates how much of the reprocessing is complete. When reprocessing to refine your dataset, you may want to keep an eye on the number of Total Decoded Log Entries versus the number of Total Filtered Log Entries .
To check the progress of transformation	Review the value in the Transformation Progress field. This percentage indicates how much of the transformation is complete.

Detailed Status interface

Navigation title:

The **Detailed Status** interface is useful for troubleshooting errors or other issues with Data Workbench server computers.

This includes any **Transform** profiles running on those computers, or Report computers that are clients of the Data Workbench server. You can access **Master Server** and **Query Server Detailed Status** interfaces through the **Admin** menu. To access the **Detailed Status** interface for other computers, in the **Servers Manager**, right-click the node of the server for which you want to view status and click **Detailed Status**. See [Servers Manager](#) on page 476.

For more information about the Data Workbench server, see the *Server Products Installation and Administration Guide*.

Search:

-Detailed Status for Release Machine

- +Component Status map
- +Disk Space
- +Execution Engine
- +Hardware
- +Memory Status ServerMemoryStatus
- +Network Connections
- +Performance
- +Poll Latency
- +Processing Status
- +Profiles
- +Query Memory Usage
- Location
- Product
- Report Server Status
- Server Common Name VSDEVTEST02
- Service Status running
- Status Generated At 06/09/08 11:34:24
- Status Generation Took (s) 0.091145660728216171
- Version 5.20i (08052811)



Note: To update the information in a **Detailed Status** interface, right-click the **Detailed Status** heading and click **Refresh**.

The following table lists the tasks that can be completed using the **Detailed Status** interface.

Detailed Status interface tasks

To perform this task...	Do this...
To display each computer component and its current status	Click Component Status .
To display how much memory on the computer is being used	Click Memory Status > Address Space Load . For more information about monitoring address space load, see the <i>Server Products Installation and Administration Guide</i> .
To determine whether the computer is configured to use the /3GB Switch	Click Memory Status > Process Address Space . If the Total field displays more than 3000000 KB, your computer is configured to use the /3GB Switch. For more information about the /3GB Switch, see the <i>Server Products Installation and Administration Guide</i> .
To monitor the amount of disk space and memory used to store each dimension as well as that used to store the names of its elements	Click Performance > Dimensions > Disk Usage > <profile name> or Performance > Dimensions > Memory Usage > <profile name> . The Disk Usage fields provide the name and amount of disk space (in MB) required to store each dimension. Large disk usage numbers can adversely affect query times because the Data Workbench server has to read through all the data to complete related queries. Lowering the disk usage for a dimension can reduce the time it takes to complete related queries. The Memory Usage fields provide the number of elements in each dimension and the amount of memory required to store the list of element names. Large numbers of elements can adversely affect

To perform this task...	Do this...
	<p>the amount of memory being used during a query because the Data Workbench server has to read through each element. Reducing the number of elements in a dimension can reduce the time it takes to complete related queries.</p> <p>Example:</p> <pre>+ Performance - Dimensions - Disk Usage - ProfileName - DimensionName 1.386 MB ... - Memory Usage - ProfileName - DimensionName 21 elements, 0.001 MB ... </pre>
To monitor the CPU usage for the stages within Log Processing and Transformation	<p>Click Performance > CPU Usage > Log Processing > <profile name> or Performance > CPU Usage > Transformation > <profile name>.</p> <p>Each of these sets of fields provides you with the CPU Usage (in seconds) for each of the stages within Log Processing and Transformation.</p> <p>Example:</p> <pre>+ Performance - CPU Usage - Log Processing - ProfileName 158.9 sec - Built-in 158.1 sec - StageName 13.0 sec ... - Log Processing\ProfileName 0.8 sec - StageName 0.8 sec </pre> <p>The time that it takes to complete a query is usually proportional to the total size of all of your dimensions. After reviewing the size of each dimension, you can evaluate whether a particular dimension is useful enough and used often enough to justify the performance cost of the dimension. If it is not, you can delete the dimension in the Profile Manager. See Profile Manager on page 456.</p> <p>A dimension whose list of element names is excessively large (that is, more than 128 MB) may cause “Out of memory” errors even if the total address space usage is not near the limit.</p> <p>Also, if you are using a Data Workbench server cluster but not using centralized normalization, a dimension whose list of element names is large has a significant impact on send memory budgets. For more information about centralized normalization, see the <i>Dataset Configuration Guide</i>. If the amount of memory required to store all of the lists of element names combined is more than 100 MB across all of the servers in the cluster, you might receive “Send memory budget exceeded” errors even when query activity is light. For example, if you have a four-server cluster with more than 25 MB on each server being used to store the lists of element names, you might receive errors.</p>
To monitor the time spent in Log Processing and Transformation	<p>Click Performance > CPU Usage > Log Processing > <profile name> or Performance > CPU Usage > Transformation > <profile name>.</p> <p>Reviewing the fields in these sections enables you to identify filters and transformations that may be negatively affecting the amount of time needed for Log Processing and Transformation. You then can make design decisions regarding individual filters and transformations with long processing times.</p>
To monitor disk space usage and increase query speed	Click Performance > Log Processing Fields > <profile name> .

To perform this task...	Do this...
	Each line item in this section corresponds to a parameter in the <code>Log Processing.cfg</code> file. Reviewing these fields enables you to see how much memory each parameter is using. You then can make design decisions regarding individual items that are quite large.
To determine the elapsed time of previous reprocessing or transformation	<p>Click Processing Status > <profile name> > Processing Mode History.</p> <ul style="list-style-type: none"> • Real Time - time that the Data Workbench server was available for making queries. • Fast Input - this time plus the Fast Merge time is the total time needed for processing the dataset. • Fast Merge - total time needed for transforming the dataset.
To diagnose “As-of time” issues	<p>Click Processing Status > <profile name> > As Of Time > Sources as-of.</p> <p>Reviewing the as-of times for each source can help you determine which source(s) may be negatively affecting the Overall As-of. You then can address the problems with those particular sources.</p>
To estimate how long a running query takes to complete	<p>Click Execution Engine.</p> <p>Reviewing the Data Sweep Time field provides you with an estimate of how long it takes for a query to complete.</p>
To list all profiles available on this computer and details about their status	Click Profiles .
To view Replication status	<p>Click Component Status.</p> <p>Check the status of the Replicate component. If Replication is running, OK displays. If the Replicate component has failed, an error message displays.</p>
To view Report Server status for a Report computer that connects to the Data Workbench server	<p>Click Report Server Status.</p> <p>This section of the Detailed Status interface includes a copy of the <code>Report Server.cfg</code> file, information about the number of reports that are running (Current Slice), and information about the most recent error (Last Error).</p> <p>For steps to edit the <code>Report Server.cfg</code> file, see the <i>Data Workbench Report Guide</i>.</p> <p> Note: If the Report Server Status section does not appear in the Detailed Status interface, you may need to configure the Data Workbench server to display Report Server Status. For steps, see the <i>Data Workbench Report Guide</i>.</p>
To view memory usage information for Transform	<p>Click Processing Status > Transform.</p> <p>For more information about Transform, see the <i>Server Products Installation and Administration Guide</i> and the <i>Dataset Configuration Guide</i>.</p>
To save the Detailed Status interface as a <code>*.cfg</code> file that can be opened in a text editor such as Notepad or distributed to others	<p>Right-click the Detailed Status heading and click Save Copy As.</p> <p> Note: Right-clicking the Detailed Status heading and clicking Save to server name>Status/ does not work in the Detailed Status interface. The following error message appears: Unable to save /Status/. 403 Forbidden</p>
To view Rows per Log Source metric	If Rows per Log Source metric needs to be reported in Detailed Status, then the Data Workbench Administrator should define the "Log Source ID" and provide a unique name in Custom Profile's Log Processing.cfg.

Profile Manager

Navigation title:

The **Profile Manager** is the main management tool in Data Workbench.

It enables you to do the following:

- Edit configuration files (except for `report.cfg` files) and other types of files.
- Publish local changes to a profile on the Insight Server.
- Create new folders.
- Edit your menu structure. See [Customize a menu](#) on page 484.
- Delete workspaces, templates, reports, and so on from Insight Server.
- Manage metrics and dimensions.

For an overview of profiles, see [Profiles](#) on page 202.

Profile Manager

Navigation title:

The folder and file names included in your implementation are displayed on the left side of the **Profile Manager**.

The profiles that make up your application are displayed as the individual columns in the **Profile Manager**. These profiles include multiple inherited profiles and a single working profile.



Note: Your working profile (either a dataset profile or a role-specific profile) is the profile that you load when you open Data Workbench.

The check marks (and their colors) indicate the profile folder(s) on the Data Workbench server and Data Workbench computers in which each file resides, whether multiple copies of a file exist, and whether those multiple copies have the same Modified date and time. These files are synchronized between the Data Workbench server and Data Workbench computers during profile download.

Following is a sample **Profile Manager** for a HBX implementation:

File	Base	Traffic	Value	Marketing	Geography	HBX	Commerce	Dataset	User
profile.cfg									
Color Legends\		✓	✓	✓			✓		
Context\	✓	✓	✓	✓	✓	✓	✓		
Dataset\	✓	✓		✓	✓	✓	✓		✓
Dimensions\	✓	✓	✓	✓	✓	✓	✓		
Export\	✓								✓
Filters\		✓				✓			
Images\	✓								
Maps\					✓	✓			
Menu\	✓	✓	✓	✓	✓	✓	✓		
Metrics\	✓	✓	✓	✓	✓	✓	✓		
Reports\	✓					✓	✓		
Workspaces\	✓					✓	✓	✓	✓

From the **Profile Manager** menu, you can open any of the other managers (for example, the **Dimensions Manager** or **Reports Manager**), which display only particular portions of **Profile Manager**. You also can create new profile managers. See [Create a Profile Manager](#) on page 488.

A check mark next to a file name in a particular column indicates that a file by that name resides in the folder named in that column (profile). As you move to the right in the **Profile Manager**, the files take precedence over those to the left, that is, each inherited profile builds on the profiles to its left in the **Profile Manager**. For example, if you have a file of the same name and in the same location in the **Base** profile (column) and in the **User** profile (column), the file in the **User** profile is used instead of the file in the **Base** profile.

Search for profiles

With Data Workbench 5.5, a search field has been added to find required profiles in the **Profile Manager**.



The following types of columns appear in the **Profile Manager**:

- The *inherited profile name* columns contain check marks for files that reside in each profile folder. Inherited profiles include internal profiles provided by Adobe as well as any company-specific or role-specific profiles that you create and maintain. In the example above, the internal profiles include Base, Traffic, Value, Marketing, and so on. The internal **Base** profile, which contains the basic building blocks and configuration information needed to run your Adobe application, is provided with every implementation. The other internal profiles contain elements (workspaces, metrics, derived dimensions, and so on) related to particular types of information, such as web traffic or marketing. Adobe provides only those profiles that are appropriate for the type of data you are analyzing and for your industry.



Note: By default, internal profiles (those provided by Adobe) cannot be changed. All customization must occur in your dataset or role-specific profiles or other profiles that you create. If you are building a new application and need to change an internal profile, you must change the `Modify Internal Profiles` parameter in the `Insight.cfg` file. See [Configuration parameters](#) on page 510 for more information. Before doing so, contact Adobe Consulting Services.

- The *working profile name* column, which is always the next-to-last column, contains check marks for files that reside in the current working profile's folder. In the example above, the working profile is Dataset. Your working profile is either a dataset profile or a role-specific profile. The files in this folder take precedence over any files with the same names in any inherited profile folder.
- The **User** column, which is always the last column, contains check marks for files and folders that reside as local files in the `User\profile name` folder. The directory structure of the User folder mimics that of the working profile, and each `User\profile name` folder contains local copies of the workspaces, metrics, dimensions, and configuration files for that particular profile. These local copies take precedence over any files with the same names in any inherited or working profile folder. The files in the **User** column were either created and saved to only the `User\profile name` folder, or they reside in an internal or working profile as well as in the `User\profile name` folder. The files in each folder may or may not be identical and may or may not have the same Modified date and time.



Note:

- To avoid changing your dataset only locally, the Data Workbench server ignores the local copies of the `profile.cfg` file and any files in the Dataset or Export folders in the `User\profile name` folder. Ignored files are identified by a red background in the **User** column and an “Ignored in User directory” warning in the context menu. To implement the changes you make in your local copies of these files, you must save them to your working profile so that they can be synchronized with the Data Workbench server. For steps to save files to your working profile, see [Publish files to your working profile](#) on page 459.
- A hyphen (-) instead of a check mark in a column identifies an empty (zero-byte) file. Data Workbench treats zero-byte files as non-existent, which enables you to use them to hide files included in a profile to the left. See [Hide a file by emptying it \(zero-byte\)](#) on page 460.

Determine file versions

As mentioned in the previous section, the check marks in the **Profile Manager** are color-coded so that you can easily identify where a file resides and whether the multiple copies of a file were modified at different times.

If a file or a collapsed directory is exactly the same as the file or directory to its left, it has the same color check mark as the file or directory in that column (profile). If it is different from any file or directory to its left, or the file or directory exists only in the **User** column, the check mark is white.

File	Base	Movies	User
profile.cfg		✓	
Context\	✓		✓
Dataset\	✓	✓	
Dimensions\	✓		
Export\	✓		
Filters\		✓	
Images\	✓		
Menu\	✓	✓	✓
Metrics\			
A New Metric.metric			✓
Average Score.metric		✓	✓
Average Score Error.metric		✓	✓

The **Profile Manager** shown in the example above indicates the following:

- A white check mark for the `A New Metric.metric` file appears only in the **User** column, which indicates that you have only a local copy of that file—it has not been published (or uploaded) to the Data Workbench server for other Data Workbench users to access.
- Check marks for the `Average Score.metric` file name appear in the **Movies** and **User** columns. The check mark in the **User** column is the same color as the check mark in the **Movies** column, which indicates that the local copy of the file has the same Modified date and time as the file in the **Movies** folder.
- Check marks for the `Average Score Error.metric` file name appear in the **Movies** and **User** columns. The check mark in the **User** column is white, which indicates that the local copy of the file has a different Modified date or time than the file in the **Movies** folder.

Create folders in the user profile

Navigation title:

Steps to create a new folder in the User Profile.



Note: Creating folders might affect what appears in your menus. For more information about altering your menu structures, see [Customize a menu](#) on page 484.

1. Right-click in the **User** column for the directory in which you want to create a new folder and click **Create > Folder**.
2. (Optional) To add an `order.txt` file to this folder, right-click in the **User** column for the new folder and click **Create > order.txt**.

For more information about using the `order.txt` file to control the appearance of your menus, see [Customize a menu](#) on page 484.

You now can add local files to this folder. When you publish the local files that you create in this folder to your working profile, this folder is automatically saved there as well. For more information about publishing files, see [Publish files to your working profile](#) on page 459.

Modify local files in the user profile

Navigation title:

You can use the **Profile Manager** to download files that you want to modify.

You might want to download a file to overwrite your existing, local file with the original version of the file or open, view, and modify files that cannot be accessed from the workspace menus.

To download a file

1. In the **Profile Manager**, open the necessary folders and subfolders to locate the file that you want to download.
2. Right-click the check mark next to the name of the file and click **Make Local**.



Note: Configuration (.cfg), dimension (.dim), and metric (.metric) files can be edited directly in a profile folder and saved to the server without making them local and separately saving them to the server. Simply right-click the check mark next to the name of the file and click **Open > in workstation**.

After the file has been downloaded to the local computer, a check mark appears in the **User** column, which indicates that a local copy of the file resides in the `User\profile name` folder on your computer. Note that the check mark is the same color as the check mark in the *profile name* column. This indicates that the local file has the same Modified date and time as the file in the *profile name* folder.

To modify the file

1. Right-click the check mark next to the file name in the **User** column.
2. Click one of the following menu options, depending on how you want to edit the file:
 - **Open > in workstation** if you are editing a .vw file or a .cfg file in a workspace.
 - **Open > in vw. Editor** if you are editing a .vw file to add new parameters.
 - **Open > In Notepad** if you are opening a text file or need to add new parameters to a .cfg file.
 - **Open > folder** if you want to open the folder in which the file is located on your computer
3. Edit the file as desired, then save the file.

In the **Profile Manager**, note that the check mark in the **User** column for the file you edited has changed color. This indicates that the local file is now different from the version in the *profile name* folder. If necessary, you can publish the revised version of the file to the profile for use by other users working with this profile.

Publish files to your working profile

Navigation title:

The **Profile Manager** enables the management of visualizations, segments, metrics, dimensions, and so on.

For example, if you create a metric that is useful to other users, you can share this metric with other Data Workbench users working with this profile by publishing (or uploading) it to the Data Workbench server.

Although you can save workspaces to your working profile on the Data Workbench server using the **Profile Manager**, you can do so directly from the **Worktop** by right-clicking a workspace thumbnail and selecting **Save to server**, which is the suggested method.



Note: Only users with the appropriate permissions can save workspaces to the Data Workbench server. For more information, contact your system administrator.

To publish a saved file

1. In the **Profile Manager**, open the necessary folders and subfolders in the **Profile Manager** to locate the file that you want to publish. The **Profile Manager** displays a white check mark in the **User** column next to the file indicating that a local copy of this file now resides in the `User\profile name` folder on your computer.

-
2. Right-click the white check mark for the file in the **User** column and click **Save to > <profile name>**.

The **Profile Manager** displays a check mark next to the file name in the selected profile's column and removes the white check mark in the **User** column. Your file is now available for use by all users and is downloaded automatically to them (and you) when the profile is next loaded.

Delete files from your working profile

Navigation title:

Steps to delete a file from your working profile.

1. In the **Profile Manager**, open the necessary folders and subfolders in the **Profile Manager** to locate the file that you want to delete.



Note: All workspaces are located in the Workspaces folder.

2. Right-click the check mark in the *working profile name* column (for example, Dataset) for the workspace or file that you want to delete and click **Remove > Yes**.

The workspace or file is deleted from the server and is no longer available for use.

Hide a file by emptying it (zero-byte)

Navigation title:

If you do not have permission to delete files from a profile or you do not want to delete a file permanently, you can use empty (zero-byte) files to hide files.

In the **Profile Manager**, a hyphen (-), instead of a check mark, in a column identifies a zero-byte file.

File	Base	Traffic	Value	Marketin	Geograp	IP C
Menu\						
Add Table.1d	✓					
order.txt	✓					
Add Annotation\	✓					
Add Custom\	✓					
Add Legend\						
Color.vw	✓					
Confidence.vw		✓				
Metric.vw	✓	-				
Value.vw			✓			
order.txt		✓				
Metric\						
Metric Legend.vw	✓		✓			
Metrics by Page View.vw	✓					
Metrics by Session.vw	✓		✓			
Order.txt	✓					

Unlike the other methods of hiding files (such as `order.txt`, the `Show` parameter, and the `Hidden` parameter), Data Workbench treats zero-byted files as non-existent. For example, if you zero-byte a dimension that has been used in a visualization or a metric definition, Data Workbench produces an error for that visualization or metric, respectively.

This functionality is useful for a number of reasons, including when you want to do the following:

- **Make a file unusable** in Data Workbench without needing the profile permissions required to delete the file.
- **Move a metric, dimension, or filter** to another location without needing the profile permissions required to delete the file from the original location.

-
- **Hide menu items.** For example, the **Base** profile has a **Metric Legend** defined in the **Metric.vw** file. Say your company has created three metric legends that you want to appear on a Add Legend > Metric submenu. You can zero-byte the **Base** profile **Metric.vw** file so that only your new submenu and three new metric legends appear.

To hide a file

1. In the **Profile Manager**, open the necessary folders and subfolders to locate the file that you want to zero-byte.
2. Right-click the check mark next to the name of the file and click **Make Local**.
3. Open the local file and delete its contents.
4. Save and close the file.

Copy and paste a file as an attachment

Navigation title:

Steps to copy and paste a file as an email attachment in Microsoft Outlook.

- Right-click the check mark for the file and click **Copy**. In the body of your email, press **Ctrl+v** to attach the file.

Work with derived metrics

Navigation title:

You define new metrics (referred to as derived metrics) and edit existing metric definitions using the **Metric Editor**.

For more information about metrics than is provided in this section and in *Query language syntax* on page 515, see the *Metric, Dimensions, and Filters Guide*.

Create a derived metric

You use a **Metric Editor** to define a new metric by name, formula, and format, which is saved to the `User\profile_name\Metrics` folder for later use.

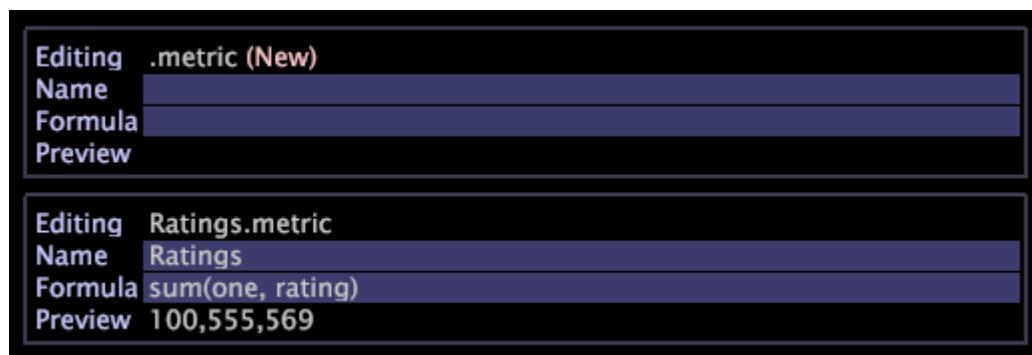
1. Open a new **Metric Editor** using the **Admin > Profile** menu option or by right-clicking the **User** column for the folder in which you want to create the metric and clicking **Create > New Metric**.

A **Metric Editor** displays.

2. In the **Name** parameter, type a name for the new metric.

Note that spaces () are allowed while underscores (_) are not. In addition, you cannot use the following symbols:

+ - * /



3. In the **Formula** parameter, type an expression for the new metric. Note that filters must be defined within brackets [] in the expression.

For additional metric expression syntax rules, see [Syntax for metric expressions](#) on page 516.

The following table provides sample expressions for extended metrics.

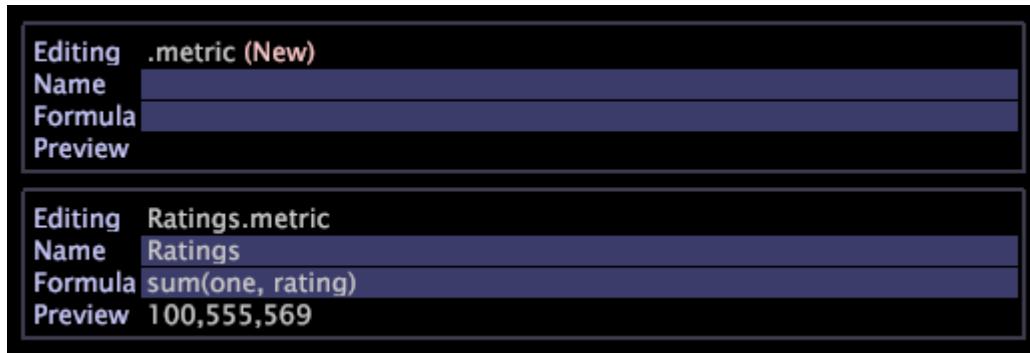
Extended Metric Name	Expression
Percent First Sessions	Sessions [Session_Number="1"] / Sessions
Conversion First Sessions	Conversion [Session_Number="1"]
Average Value Per Visitor	Value/Visitors



Note: When an appropriate expression is entered, the preview line displays the value of the new metric. If there is an error in the expression, the preview line displays an error message.

4. Right-click (**New**) and click **Save**.

When you save the metric, a file representing the new metric is created on your computer in the Data Workbench Installation directory \User\profile name\Metrics folder.



You now can use the new metric throughout the current profile by selecting it as you would any built-in metric. To change the order in which your metrics appear on the metrics menu, see [Customize a menu using order.txt files](#) on page 486.

If you would like all users of the profile to use the metric that you created, you must publish it to the working profile using the **Profile Manager**. See [Publish files to your working profile](#) on page 459.

Edit a derived metrics

1. In the **Profile Manager** or **Metrics Manager**, in the *profile name* column, right-click the check mark for the metric file that you want to edit, then click **Make Local**.
2. Right-click the check mark for the metric file in the **User** column and click **Open > from the workbench**.



Note: You also can open a **Metric Editor** by right-clicking any metric-related area within a visualization to display the metric menu. For more information, see [Metric and dimension menus](#) on page 286.

3. In the **Metric Editor**, edit and save the metric definition as necessary using Steps 2-4 in [Create a derived metric](#) on page 461.

If you would like all users of the profile to use the metric that you edited, you must publish it to the working profile using the **Profile Manager**. See [Publish files to your working profile](#) on page 459.

Work with derived dimensions

Navigation title:

The new dimensions that you create using Data Workbench (referred to as derived dimensions) are client-side dimensions.

Instead of defining these dimensions during the dataset construction and update process (in the `Transformation.cfg` file) on your Data Workbench server computers, derived dimensions are created and stored individually as `.dim` files in a profile. As a result, you can change existing and create new derived dimensions without reprocessing your dataset.



Note: For more information about dimensions than is provided in this section, see the appropriate Data Workbench application guide.

For more information about the dataset configuration and update process, see the *Dataset Configuration Guide*.

Create a derived dimensions

To create a derived dimension, you can either copy and modify an existing dimension or save a dimension from a visualization.

Create a derived dimensions from an existing dimension

Users most often want to create new time dimensions from existing ones. For example, you can create a new “Last 5 Days” dimension from the existing “Last 7 Days” dimension.

1. In the **Profile Manager**, in the *profile name* column, right-click the check mark for a dimension that is similar to the dimension that you want to create and click **Copy**.

For example, to copy the `Last 7 Days.dim` from the Reporting folder of the **Traffic** profile, you right-click the check mark for the file name in the **Traffic** column and click **Copy**.

Last 14 Days.dim		✓	Path Dimensions\Time\Reporting\ File Last 7 Days.dim From Traffic Date 04/29/05 15:15:00 Size 0.3 KB
Last 24 Hours.dim		✓	
Last 3 Months.dim		✓	
Last 4 Weeks.dim		✓	
Last 48 Hours.dim		✓	
Last 7 Days.dim		✓	
Last 7 Days and Today.dim		✓	
Last 8 Weeks.dim		✓	
Last Month.dim		✓	

2. Right-click in the **User** column for the folder in which you want to store the copied dimension and click **Paste**.

The dimension appears in the selected Dimensions folder with a check mark in the **User** column.

3. To rename the new dimension, right-click its check mark in the **User** column and type the new name in the **File** field.
4. From the right-click menu, click **Open > from the workbench**. The defining parameters for the dimension appear.
5. Modify the parameters as needed to define the new dimension.

For time dimensions, you most likely need to modify only the Count and Range parameters.

6. To save the file, right-click (**modified**) at the top of the window and click **Save**.

If you would like all users of a profile to use the dimension that you created, you must upload it to the profile using the **Profile Manager**. For more information, see [Publish files to your working profile](#) on page 459.

You now can use the new dimension throughout the current profile by selecting it as you would any built-in dimension.

Save a dimension from a visualization

You can save extended dimensions from process maps and segments. For steps to save a dimension from a process map, see [Save a dimension from a process map](#) on page 400. For steps to save a segment dimension, see [Create a segment dimensions](#) on page 413 on page 82.

Saving a segment as a dimension

You also can save defined segments as a dimension. For steps, see [Reusing a segment visualization](#) on page 413.

Edit an existing derived dimension

1. In the **Profile Manager**, in the *profile name* column, right-click the check mark for the dimension file that you want to edit and click **Make Local**.
2. Right-click the check mark for the dimension file in the **User** column and click **Open > from the workbench**.
3. Complete the parameters as needed. For more information, contact Adobe Consulting Services.
4. To save the file, right-click (**modified**) at the top of the window and click **Save**.

If you would like all users of a profile to use the modified dimension, you must upload it to the profile using the **Profile Manager**. For more information, see [Publish files to your working profile](#) on page 459.

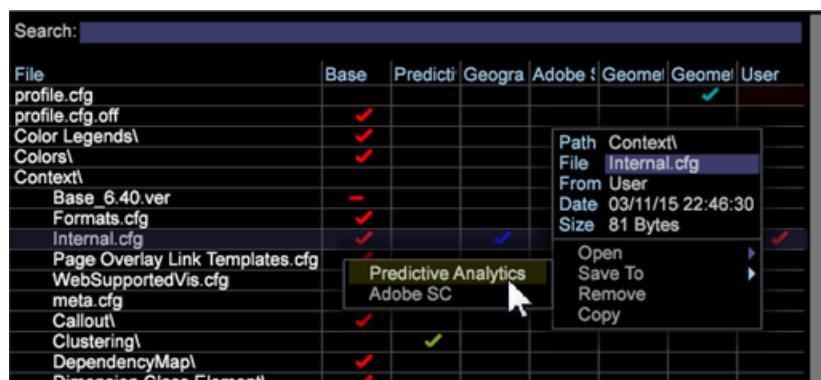
Locking Profiles in the Workstation

The **Internal.cfg** file applied in the Profile Manager prevents changes by users to your custom profiles by the Profile, Dimensions, Reports, Workspaces, Metrics, and Filters managers.

You can prevent profile files from being modified and overwritten when using the managers by saving the **Internal.cfg** file to your custom profile in the Profile Manager. This configuration file prevents users from overwriting multiple files when working in the managers (accessed from the **Admin > Profile** menu).

Locking Profiles in the Profile Manager

1. In the workspace, right-click **Admin > Profile Manager**.
2. In the **Profile Manager**, right-click **Context > Internal.cfg** and **Make Local**.
3. Right-click checkmark in **User** column and save to a <custom profile>.



Note: Only changes to profile files by the managers are prevented when saving the **Internal.cfg** to a custom profile in the Profile Manager. You can still save workspaces to the server from the worktop using the **Save to server** command.

Dataset managers

Navigation title:

Information about dependency maps and field viewers.

Dependency Maps

Navigation title:

Dependency maps enable you to visualize and manage the configuration of the components of your profile.

- **Dataset components:** Log sources, filters, fields, transformations, and extended dimensions defined in your dataset's `Log Processing.cfg`, `Transformation.cfg`, and **dataset include** files.
- **Query model components:** Metrics, dimensions, and filters defined in the Dimensions, Metrics, and Filters folders.
- **Workspaces and visualizations:** Workspaces, reports, menu options, and globe layers.

For more information about working with query model components, workspaces, and visualizations in dependency maps, see the *Data Workbench User Guide*.

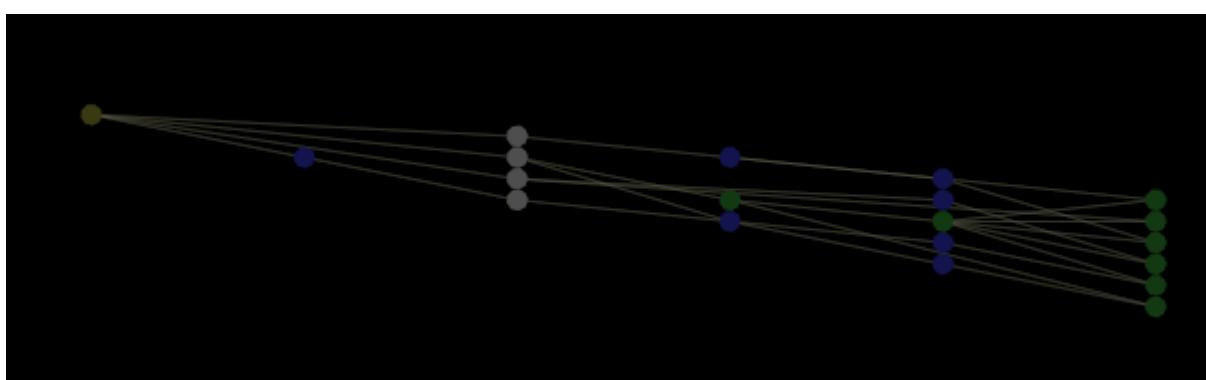
Profile components are represented by colored dots (nodes) in the map. The lines connecting the nodes depict dependencies, that is, how the components relate to one another. A line between two nodes means that an output of the node on the left is an input of the node on the right, that is, the right node depends on the left node.

Displaying Dataset Components

1. Right-click within the dependency map and click **Display**.
2. Choose **Dataset**. An X appears to the left of **Dataset**.

For more information about the other display options see the *Data Workbench User Guide*.

The following figure shows a dependency map whose nodes represent a dataset's log sources, fields, transformations, and extended dimensions.



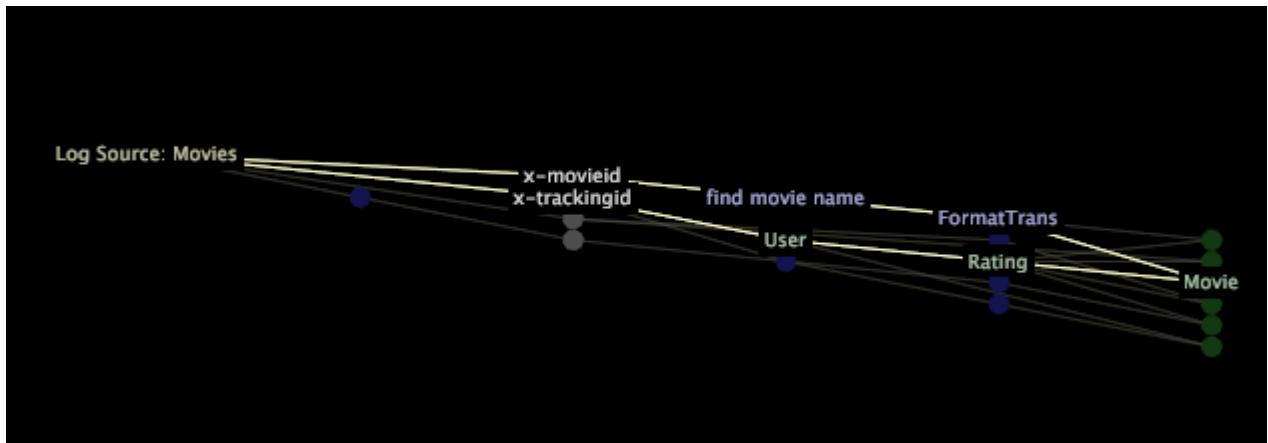
- A yellow-green node represents one or more log sources or a filter defined in the dataset. A node for a log source always appears furthest to the left in the map.
- A gray node represents a field that is listed in the Fields parameter in a `Log Processing.cfg` or **Log Processing Include** file.
- A blue node represents a transformation.
- A green node represents an extended dimension.



Note: If your dataset has a single log source, the map displays Log Source: *log source name*. If your dataset has multiple log sources, the map displays *number* Log Sources, where number is the count of log sources. For example, if you have three log sources in your dataset, your map displays 3 Log Sources.

If you cannot see all of the nodes on the map, you can move the map or zoom in or zoom out to display the entire map or to focus on a particular section. For more information about zooming, see the Working with Visualizations chapter of the *Data Workbench User Guide*.

When you click a node, all of the nodes that depend on that node and all of the nodes on which that node depends are highlighted and their names display.



Note: A highlighted path in a dependency map does not constitute a selection.

When you right-click a node, you can see identifying information about each component shown on the map and choose menu options that enable you to view more detail about the component or to edit the component. In addition, you can perform text searches and display performance information for transformations and extended dimensions.

For information about these functions for dependency maps, see the Administrative Interfaces chapter of the *Data Workbench User Guide*.

Display profile components

Navigation title:

You can choose to display your profile's dataset components, query model components, or workspaces, reports, menu options, and globe layers in the dependency map.

To select the components to display

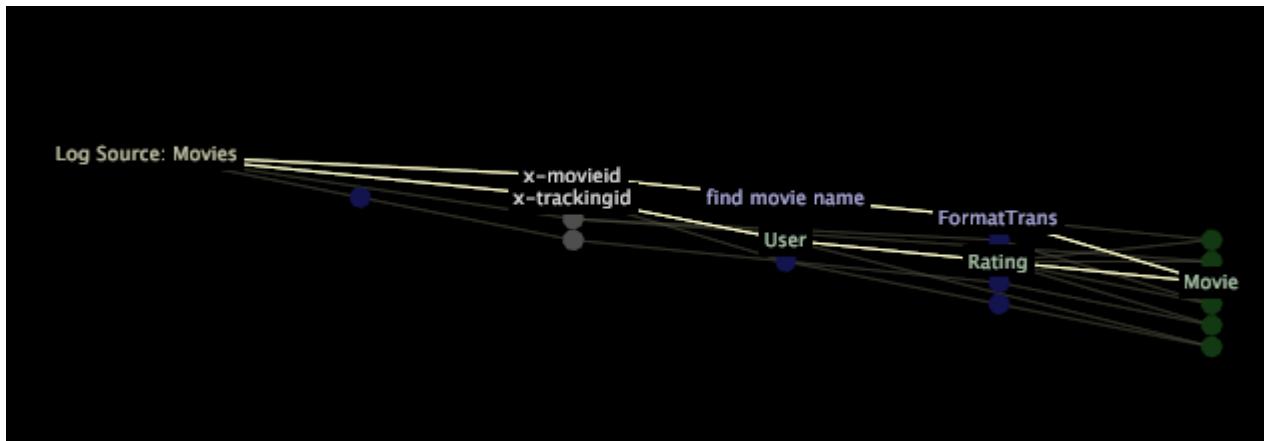
1. Right-click within the dependency map and click **Display**.
2. Choose one or more of the following options to display on the map. An X appears to the left of each display option that you enable.
 - **Dataset** to display dataset components. See [Dataset components](#) on page 467. If you choose to display the dataset components, you have the option to **Include File Blocks** on the map. See [File blocks](#) on page 473.
 - **Query Model** to display query model components. See [Query model components](#) on page 468.
 - **Workspaces and Visualizations** to display workspaces, reports, menu options, and globe layers. See [Workspaces and visualizations](#) on page 469. This option works only if the **Query Model** display option is enabled.



Note: If the **Query Model** display option is not enabled when you choose the **Workspaces and Visualizations** display option, an error message appears.

If you cannot see all of the nodes on the map, you can move the map or zoom in or zoom out to display the entire map or to focus on a particular section. For more information about zooming, see [Zoom in or out in a visualization](#) on page 292.

When you click a node, all of the nodes that depend on that node and all of the nodes on which that node depends are highlighted and their names display.



Note: A highlighted path in a dependency map does not constitute a selection.

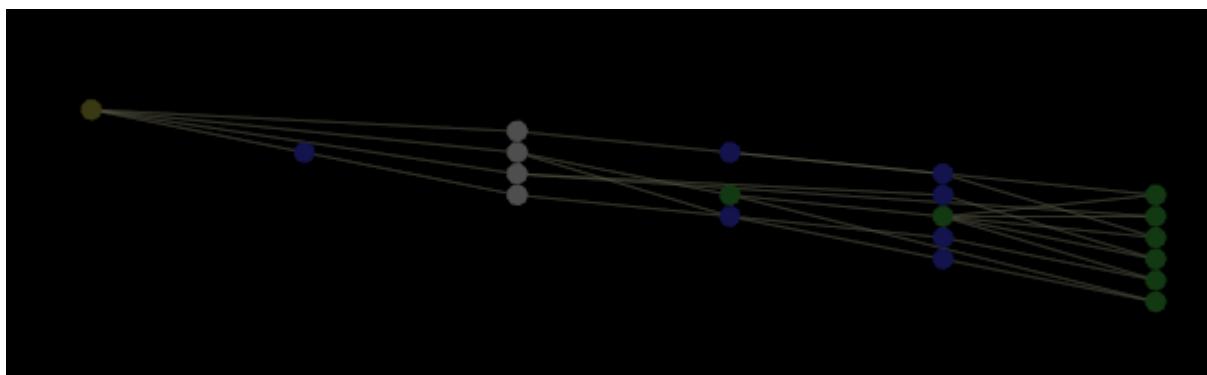
When you right-click a node, you can see identifying information about each component shown on the map and choose menu options that enable you to view more detail about the component or to edit the component. In addition, you can perform text searches and display performance information for transformations and extended dimensions.

Dataset components

Navigation title:

Conceptual information about dataset components.

The following figure shows a dependency map whose nodes represent a dataset's log sources, fields, transformations, and extended dimensions.



- A yellow-green node represents one or more log sources or a filter (such as a Log Entry Condition) defined in the dataset.

-
- A node for a log source always appears furthest to the left in the map. If your dataset has a single log source, the map displays Log Source: *log source name*. If your dataset has multiple log sources, the map displays *number* Log Sources, where number is the count of log sources. For example, if you have three log sources in your dataset, your map displays 3 Log Sources.
 - The map displays one Log Entry Condition node for each **log processing dataset include** file but only one Log Entry Condition node for transformation (if defined in the Transformation.cfg file). If the Log Entry Condition is empty, it does not display on the map.
 - A gray node represents a field that is listed in the Fields parameter in a Log Processing.cfg or **Log Processing include** file.
 - A blue node represents a transformation.
 - A green node represents an extended dimension.



Note: If your profile's Dataset folder contains the file Insight Transform.cfg, the dependency map shows the log sources, transformations, and exporters defined for use with Transform. For information about Transform, see the *Dataset Configuration Guide*.

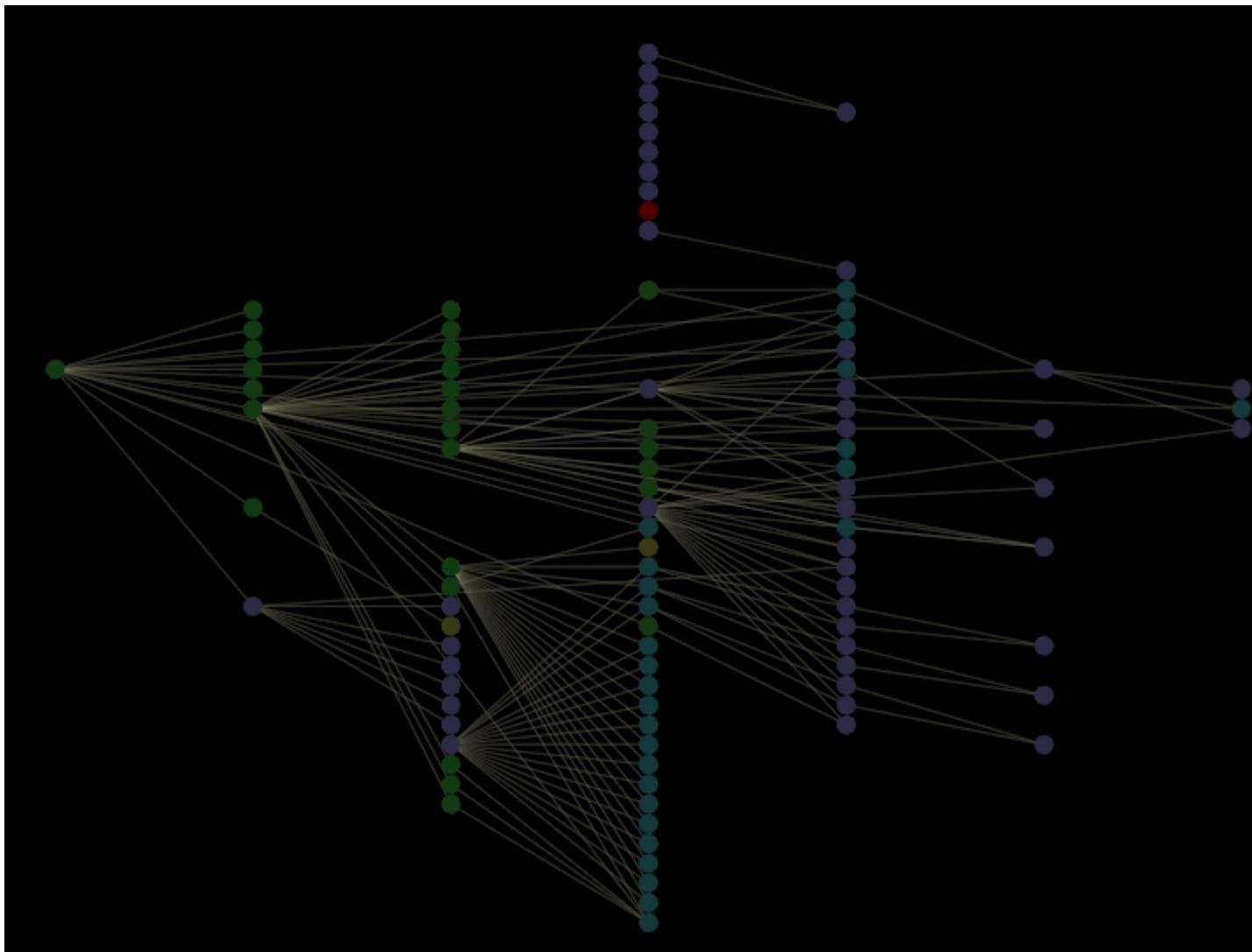
When you enable the Include **File Blocks** display option, the map displays a single blue node for all of the transformations defined in one dataset configuration file and a single green node for all of the extended dimensions defined in one dataset configuration file. For more information about this display option, see [File blocks](#) on page 473.

Query model components

Navigation title:

Conceptual information about query model components.

The following figure shows a dependency map whose nodes represent a query model's metrics, derived dimensions, and filters defined in the Dimensions, Metrics, and Filters folders within the profile as well as the extended dimensions defined in the dataset that relate to them in some way.



- A yellow-green node represents a filter.
- A purple node represents a metric.
- A blue-green node represents a derived dimension.
- A green node represents an extended dimension (defined in the dataset).
- A red node represents a metric, derived dimension, or filter with a broken or circular dependency or other error.



Note: Because the dependency map is designed to accommodate acyclic dependencies, nodes involved in circular dependencies may not display properly on the map. You can search for circular dependencies by typing “circular dependency” in the **Search** text box. For more information about the **Search** feature, see [Search within a map](#) on page 472.

Workspaces and visualizations

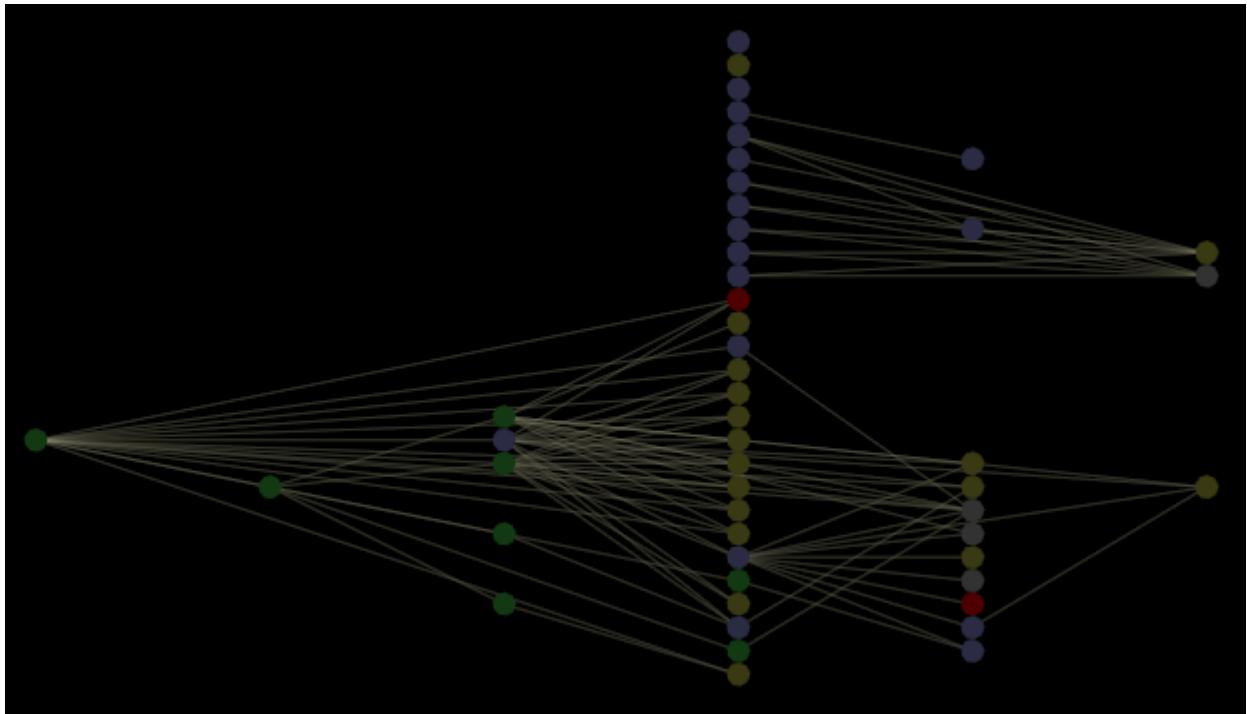
Navigation title:

Conceptual information about workspaces and visualizations.

The following figure shows a dependency map whose nodes represent the workspaces, reports, menu options, and globe layers defined in the profile. This option works only if the **Query Model** display option is enabled.



Note: If the **Query Model** display option is not enabled when you choose the **Workspaces and Visualizations** display option, an error message appears.



- A gray node represents a workspace or a report.
- A yellow-green node represents a menu option.
- A red node represents a workspace, report, menu option, or globe layer with a broken or circular dependency or other error.



Note: Because the dependency map is designed to accommodate acyclic dependencies, nodes involved in circular dependencies may not display properly on the map. You can search for circular dependencies by typing “circular dependency” in the **Search** text box. For more information about the **Search** feature, see [Search within a map](#) on page 472.

For descriptions of other nodes on the map, see [Query model components](#) on page 468.

View node details

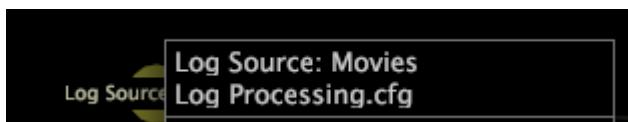
Navigation title:

Information about viewing identifying information about a profile component.

- Right-click the desired node and view the identifying information at the top of the window. The identifying information that you see depends on the type of component.

Log Source

The name of the log source and the name of the file in which the log source is defined.



Field

The name of the field.



Transformation

The transformation's type and name, the name of the file in which the transformation is defined, and the transformation's number in the file.



 **Note:** Performance information for the transformation is displayed only if the **Show Performance Data** option is enabled for the dependency map. For more information, see *Display performance data* on page 473.

Extended Dimension

The extended dimension's name and type, the name of the file in which the dimension is defined, and the dimension's number in the file.



 **Note:** Performance information for the extended dimension is displayed only if the **Show Performance Data** option is enabled for the dependency map. For more information, see *Display performance data* on page 473.

Metric

The metric's name.



Derived Dimension

The derived dimension's name.



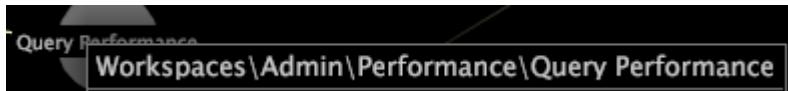
Filter

If defined in the dataset, the path of the configuration file in which the filter is defined and the filter's name.



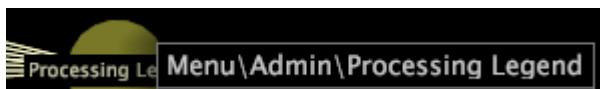
Workspaces and Reports

The path of the workspace or report in the profile's Workspace or Report directory.



Menu options

The path of the menu item in the profile's Menu directory.



To view inputs or outputs for a profile component

- Right-click the desired node and click **Inputs** or **Outputs**. A list of inputs or outputs displays.

Display a highlighted node

Navigation title:

Information about displaying highlighted nodes in another dependency map.

If a dependency path is highlighted in the dependency map, you can choose to show only those nodes in another map.

- Right-click one of the nodes in the dependency path and click **Show Only These Nodes**.

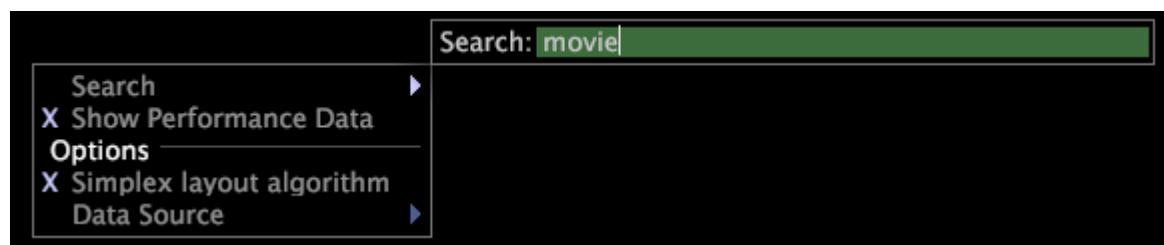
The new map opens in a callout, and all of the other menu options are available when you right-click a component in the callout.

Search within a map

Navigation title:

Steps to search for components containing specific text.

1. Right-click within the dependency map and click **Search**.
2. In the text box, type the text for which to search.



The map highlights the nodes for the components whose names or parameter settings contain the search text.

Edit profile components

Navigation title:

Steps to edit the configuration of a profile component.

From a dependency map, you can open the configuration or editor window for a log source, transformation, extended dimension, derived dimension, or metric.

1. Right-click the desired node and click one of the following options:

- **Edit Configuration**. Available for log sources, transformations, extended dimensions, and derived dimensions. The callout that appears contains only the configuration information for the component.
- **Edit Whole Configuration File**. Available for log sources, transformations, and extended dimensions. The callout that appears contains the entire configuration file in which the component is defined.

- **Edit Formula.** Available for metrics. The callout that appears contains a metric editor window. For information about the **metric editor**, see [Work with derived metrics](#) on page 461.
2. Edit the parameters as desired.
For more information about the parameters for dataset components, see the *Dataset Configuration Guide*.
 3. To save the file locally, right-click **(modified)** or **(New)** at the top of the window and click **Save**.

To make the changes take effect, save the file to the server using the **Profile Manager**. See [Publish files to your working profile](#) on page 459.

File blocks

Navigation title:

When you display dataset components on a dependency map, you have the option to enable the **Include File Blocks** display option.

When this option is enabled, the map displays a single blue node for all of the transformations defined in one dataset configuration file and a single green node for all of the extended dimensions defined in one dataset configuration file. For example, if a **log processing dataset include** file includes the definitions of three transformations, the map displays one blue node representing the three transformations. Similarly, if a **transformation dataset include** file includes the definitions of two extended dimensions, the map displays one green node representing the two extended dimensions.

Transformation blocks

Each blue node is a transformation block and has the following options:

- To view the input fields of the transformation block, right-click the node for the block and click **Inputs**.
- To view the output fields of the transformation block, right-click the node for the block and click **Outputs**.
- To edit any of the transformations in the block, right-click the node for the block and click **Edit Configuration**.
The callout that displays contains the entire configuration file in which all of the transformations are defined. You then can edit the parameters as desired. For more information about these parameters, see the *Dataset Configuration Guide*.
- To see all of the transformations in the block, right-click the node for the transformation block and click **Show Details**. The callout that displays contains another dependency map showing nodes for all of the transformations in the block.

Dimension blocks

Each green node is a dimension block and has the following options:

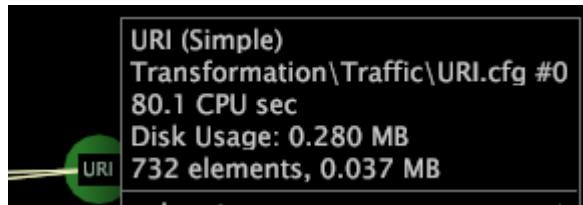
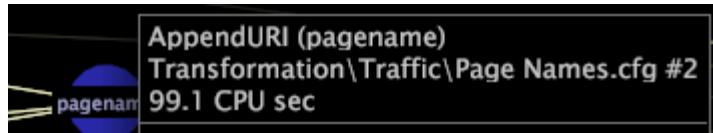
- To view the input fields or the parent dimension of the dimension block, right-click the node for the block and click **Inputs**.
- To view the output dimensions of the dimension block, right-click the node for the block and click **Outputs**.

Display performance data

Navigation title:

Information about displaying performance data for a transformation or extended dimension.

- Right-click within the dependency map and click **Show Performance Data**. When enabled, an X appears to the left of **Show Performance Data**. When you right-click a transformation or extended dimension node, the following performance information displays:
 - Number of CPU seconds required to perform the transformation or create the extended dimension
 - Amount of memory required to store the extended dimension
 - Number of elements in the extended dimension
 - Amount of memory required to store the list of element names in the extended dimension



Open a field viewer

Navigation title:

Information about opening a field viewer for a log source, transformation, or extended dimension.

- Right-click the appropriate node and click **View Fields**. For extended dimensions, click **View Input Fields**.

The callout that appears contains a field viewer, which is a table showing all of the fields that are either inputs or outputs of that component. (Inputs and outputs can be fields or dimensions.) For a log source or transformation, the field viewer shows its inputs and outputs. For a dimension, the field viewer shows its inputs. See [Field viewer](#) on page 474.



Note: To open a field viewer, your Data Workbench computer must have network access to the Data Workbench server on which the dataset resides. If your dataset resides on a Data Workbench server cluster, your Data Workbench computer must have network access to the master Data Workbench server in the cluster.

Field viewer

Navigation title:

A field viewer is a table containing the values of one or more data fields.

The fields whose values display are inputs or outputs of a dataset's log sources, transformations, or extended dimensions. The field's name is shown in the column heading, and each row contains the field's value for a single row of source data. Because field viewers enable you to see a field's values, they are helpful in writing and testing [Regular Expressions](#) on page 642.

You can open a field viewer as a callout from a **Transformation Dependency** map or as a standalone visualization from the **Admin** menu:

- Creating a field viewer from a **Transformation Dependency** map. When you open a field viewer from a **Transformation Dependency** map, the viewer is populated automatically based on the log source, transformation, or dimension that you right-click. For a log source or a transformation, the fields in the viewer are inputs or outputs of the log source or transformation. For a dimension, the fields are inputs of the dimension. You can add and remove fields as desired.
- Creating a field viewer as standalone visualization. When you open a field viewer as a standalone visualization, you can create a **Log Processing Field Viewer** or a **Transformation Field Viewer**. The viewer is blank, and you must add the desired fields to the viewer. For a **Log Processing Field Viewer**, you can add fields from the **Log Processing.cfg** file or any **Log Processing Dataset Include** file. For a **Transformation Field Viewer**, you can add fields from the **Transformation.cfg** file or any **Transformation Dataset Include** file.

x-trackingid (Pre-Log Processing)
1488844
822109
885013
30878
823519
893988
124105
1248029
1842128
2238063
1503895
2207774
2590061
2442
543865
1209119
804919
1086807



Note: Field viewers are not table visualizations; therefore, they do not have the properties associated with tables.

For information about adding and removing fields and filtering within a field viewer, see [Administrative interfaces](#) on page 446.

Create a field viewer

Navigation title:

You can open a field viewer as a callout from a dependency map or as a standalone visualization from the **Admin** menu.

Create a field viewer from a dependency map

When you open a field viewer from a dependency map (as shown in [Open a field viewer](#) on page 474), the viewer is populated automatically based on the log source, transformation, or dimension that you right-click. For a log source or a transformation, the fields in the viewer are inputs or outputs of the log source or transformation. For a dimension, the fields are inputs of the dimension. You can add and remove fields as desired.

Create a field viewers as a standalone visualization

When you open a field viewer as a standalone visualization, you can create a **Log Processing Field Viewer** or a **Transformation Field Viewer**. The viewer is blank, and you must add the desired fields to the viewer. For a **Log Processing Field Viewer**, you can add fields from the `Log_Processing.cfg` file or any **Log Processing dataset include** file. For a **Transformation Field Viewer**, you can add fields from the `Transformation.cfg` file or any **Transformation dataset include** file.

For more information about configuration and include files, see the *Dataset Configuration Guide*.

Add and remove a field

To add a field to a field viewer

- Right-click within the field viewer and click **Fields** > `<field name>` > `<instance>`.

-or-

-
- Right-click within an existing column in the field viewer and click **Fields** > <**field name**> > <**instance**>.

Field name refers to the field's name, and instance refers to where in the dataset configuration the field is used, such as a dataset processing stage or a transformation. Some fields are used in multiple places within the dataset configuration (for example, a field can be modified by multiple transformations), so you must select which instance of the field to add to your field viewer.

In the list of available fields, an X appears to the left of any field already displayed in the viewer.

To remove a field from a field viewer

- Right-click within the column for the field that you want to remove and click **Remove Field**.

Filter within a field viewer

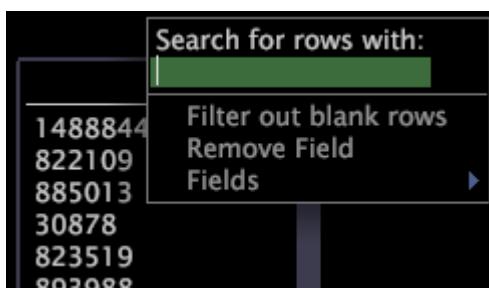
Navigation title:

You can filter the values displayed in a field viewer by searching for specific text or by filtering out blank rows.

To filter within a field viewer

1. Right-click within the column for the field whose values you want to search.
2. I

In the text box under **Search for rows with**, type the text for which to search. You also can enter regular expressions for pattern matching.



Field values that contain the text or satisfy the regular expression display in the field viewer. When the search results display, the column name is highlighted in red.

For information about regular expressions, see the *Dataset Configuration Guide*.

To filter out blank rows

- Right-click within the column for the field for which you want to filter out blank rows and click **Filter out blank rows**.

When enabled, an X appears to the left of the option, and the column name is highlighted in red.

Servers Manager

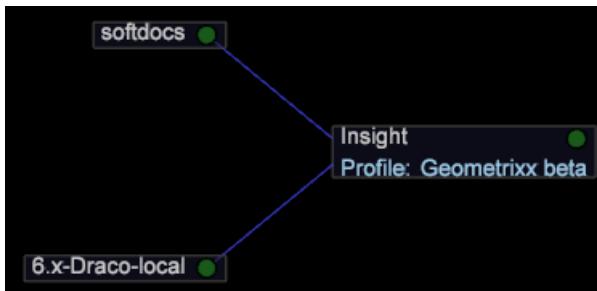
Navigation title:

The primary tool used by system administrators is the Servers Manager.

It is the main interface for determining overall system status and for performing system configuration, file management, and error monitoring functions.

The Servers Manager displays a colored dot (node) for each Data Workbench server and **Sensor** installation in your system and provides at-a-glance system status for each installation. It also displays a node for your Data Workbench installation.

Green nodes represent active connections, red nodes represent connections that are disabled or otherwise inaccessible, and gray nodes represent connections whose states are undetermined.



Right-clicking a node displays information about the connecting component and provides access to any related menus.

The following tables describe the information provided when you right-click a node for Data Workbench, Data Workbench server (including a master Data Workbench server in a cluster), or **Sensor**.

Data Workbench Servers Manager

Item	Description
Product	Product name, version, and build number. Example: Data Workbench 5.3 (00000001)
Address	IP address of the Data Workbench computer. Example: 100.0.0.1
Configure	A link to your Data Workbench's configuration file. Click Configure > Insight.cfg to display the Data Workbench configuration window. Any changes that you make and save in this window are reflected in the <code>Insight.cfg</code> file in your Data Workbench installation directory.
Product	Product name, version, and build number. Example: Data Workbench server 5.3 (00000001)
CN	The common name of the Data Workbench server computer. Example: <code>myserver1.mycompany.com</code>
Address	IP address or fully qualified domain name of the server as configured in the Addresses file on the computer and the Network Location parameter in the <code>Insight.cfg</code> file. Example: 100.0.0.1 For information about the Addresses file, see the <i>Server Products Installation and Administration Guide</i> .
Status	Current status of the Data Workbench server. This field displays OK when the Data Workbench server is running normally. If an error has occurred and the Data Workbench server node is red, the field displays the error (for example, "403 Forbidden").
Detailed Status	A link to the Data Workbench server Detailed Status interface, which is useful for troubleshooting errors or other issues with the Data Workbench server. For more information, see <i>Detailed Status interface</i> on page 452.
Remote Desktop	Opens a Remote Desktop session to the Data Workbench server computer. For more information, see <i>Remote desktop option</i> on page 478.
Server Files	A link to the Server Files Manager , which displays the directories and files in the Data Workbench server installation directory.

Item	Description
	For more information, see Server Files Manager on page 479.
Server Monitor	A link to the Server Monitor interface, which is useful for troubleshooting or tracking performance parameters. For more information, see Server Monitor interface on page 481.
Related Servers	For Data Workbench server clusters only. A menu that lists the common names of the computers listed in the master Data Workbench server's *.address file. This list usually includes all of the processing Data Workbench servers in the cluster. This menu appears only if Data Workbench has a copy of the master Data Workbench server's *.address file. When you click Related Servers , you can click either: <ul style="list-style-type: none"> • Server Monitor List, which displays the Server Monitor interface listing details for all of the related servers • The common name of any Data Workbench server, which displays a context menu that enables you to open any of the following for that particular server:<ul style="list-style-type: none"> • Detailed Status. See Detailed Status interface on page 452. • Remote Desktop. See Remote desktop option on page 478. • Server Files Manager. See Server Files Manager on page 479. • Server Monitor. See Server Monitor interface on page 481.

Sensor Servers Manager

Item	Description
Product	Product name, version, and build number. Example: Sensor 3.76; J3.67
ID	The Sensor ID specified in the Sensor configuration file for this installation.
IP	IP address of the web or application server on which Sensor is installed. Example: 100.0.0.1
PID	Process ID assigned by the operating system.
SSL	Whether Sensor and the Data Workbench server communicate using SSL.
Time	Time (HH:MM:SS) that the Sensor last established a connection with the Data Workbench server.

Remote desktop option

Navigation title:

The **Remote Desktop** option enables administrators to access the Data Workbench server or Report computers directly from Data Workbench using the computer's IP address.

Just as with other remote desktop functionality outside of Adobe products, your Data Workbench must have network access to the computer that you want to access remotely. This requires that the remote desktop port on the Data Workbench server (port 3389 by default) be open and listening for connections.

For more information about Data Workbench server, see the *Server Products Installation and Administration Guide*. For more information about Report, see the *Data Workbench Report Guide*.

To open a remote desktop

1. In the **Servers Manager**, right-click the node of the **Data Workbench server** or **Report** computer that you want to access remotely and click **Remote Desktop**.
2. When the log on window for the computer appears, type your user name and password, then click **OK**. The desktop displays and you now can navigate the computer.

Server Files Manager

Navigation title:

The Server Files Manager enables you to remotely administer and manage Data Workbench server computers from any authorized Data Workbench by providing access to all of the directories and files in the product's installation directory, including configuration and look-up files.

You can access the **Server Files Manager** using the **Admin** menu as well as by right-clicking the node of the Data Workbench server computer in the **Servers Manager** and clicking **Server Files**.

File	Visual Server	Temp
Access Control\	✓	
Addresses\	✓	
Audit\	✓	
Bin\	✓	
Components for Processing Servers\	✓	
Components\	✓	
Events\	✓	
Logs\	✓	
Lookups\	✓	
Profiles\	✓	
Software\	✓	
Trace\	✓	
Users\	✓	



Note: You can create new server files managers that display selected directories. See [Create a Server Files Manager](#) on page 489.

The left column of **Server Files Manager** lists file and folder names. The check marks in the center and right columns indicate where in the file structure these directories and files reside.

If a file resides in the product's installation directory, the *server name* (for example, Data Workbench server) column contains a check mark. If a file resides on the Data Workbench user's computer in the *Data Workbench installation directory\Temp* directory, the **Temp** column contains a check mark. The color of the check marks indicates whether the files that reside in different locations were modified at the same time.

- A red check mark in the server name column indicates that the folder or file resides on the Data Workbench server computer.
- A red check mark in the **Temp** column indicates that the local copy of the file or folder has the same Modified date and time as the file or folder on the Data Workbench server computer.
- A white check mark in the **Temp** column indicates that the file or folder in the *Data Workbench installation directory\Temp* directory has a different Modified date and time than the file or folder on the Data Workbench server computer.

The following graphic shows the **Server Files Manager** with both red and white check marks:

File	Visual Server	Temp
Access Control\		
Access Control.cfg	✓	✓
Addresses\	✓	
Audit\	✓	
Bin\		
EventMessages.dll	✓	✓
VisualServer.err	✓	✓
VisualServer.exe	✓	✓
Components for Processing Servers\	✓	
Components\	✓	
Events\	✓	
Logs\	✓	
Lookups\	✓	
Profiles\	✓	
Software\	✓	
Trace\	✓	
Users\	✓	

To manage directories and files using the Server Files Manager

You can use the **Server Files Manager** to manipulate directories and files on a Data Workbench server computer.

The following table lists the tasks that can be completed using the **Server Files Manager**:

To perform this task...	Do this...
To see the files within any directory	Click the directory name to view its contents.
To hide the content of a directory	Click the directory name.
To see details about a directory	Right-click the cell next to the directory in either the server name or Temp column. You see the following information: <ul style="list-style-type: none"> • Path. The path of the directory. • Dir. The name of the directory. • From. The location of the directory, Remote or Temp. • Date (Temp column only). Creation date or the date of the last revision to the local copy.
To see details about a file	Right-click the check mark next to the file in either the server name or Temp column. You see the following information: <ul style="list-style-type: none"> • Path. The path of the file. • File. The name of the file. • From. The location of the directory, Remote or Temp. • Date. Date of the last revision to the file. • Size. The size of the file.
To download a directory to your local computer	Right-click the check mark in the <i>server name</i> column for this directory and click Make Directory Local . A check mark for the directory appears in the Temp column.
To download a file to your local computer	Right-click the check mark in the <i>server name</i> column for this file and click Make Local . A check mark for the file appears in the Temp column.
To download the last portion of a log file to your local computer	To avoid having to download an entire log file (especially when you know that the error message is close to the end of the file), right-click the check mark in the server name column for the file,

To perform this task...	Do this...
	click Tail , and select the size of the portion you want to download. A check mark for the file appears in the Temp column. The local file contains only the amount of data that you specified, starting from the end of the file.
To open a directory	Right-click the check mark for the directory in the Temp column and click Open > folder .
To open a file	Right-click the check mark for the file in the Temp column, click Open , then click in Data Workbench , in Notepad , or folder .
Save a local copy of a directory to the Data Workbench server	Right-click the check mark for the directory in the Temp column and click Save Directory to > <profile name> .
Save a local copy of a file to the Data Workbench server	Right-click the check mark for the file in the Temp column and click Save to > <profile name> .
Remove a local copy of a directory or file	Right-click the check mark for the directory or file in the Temp column and click Remove .
Copy and paste a file as an email attachment in Microsoft Outlook	Right-click the check mark for the file in the Temp column and click Copy . In the body of your email, press Ctrl+v to attach the file.

Server Monitor interface

Navigation title:

The **Server Monitor** interface is useful for troubleshooting or simply tracking the performance parameters of Data Workbench server computers and Report computers that are clients of Data Workbench server computers.

The Server Monitor interface displays either a green dot or a red dot at its top, to the left of the computer name. A green dot indicates that the computer is functioning without issue. A red dot indicates that one or more errors have occurred on the computer.

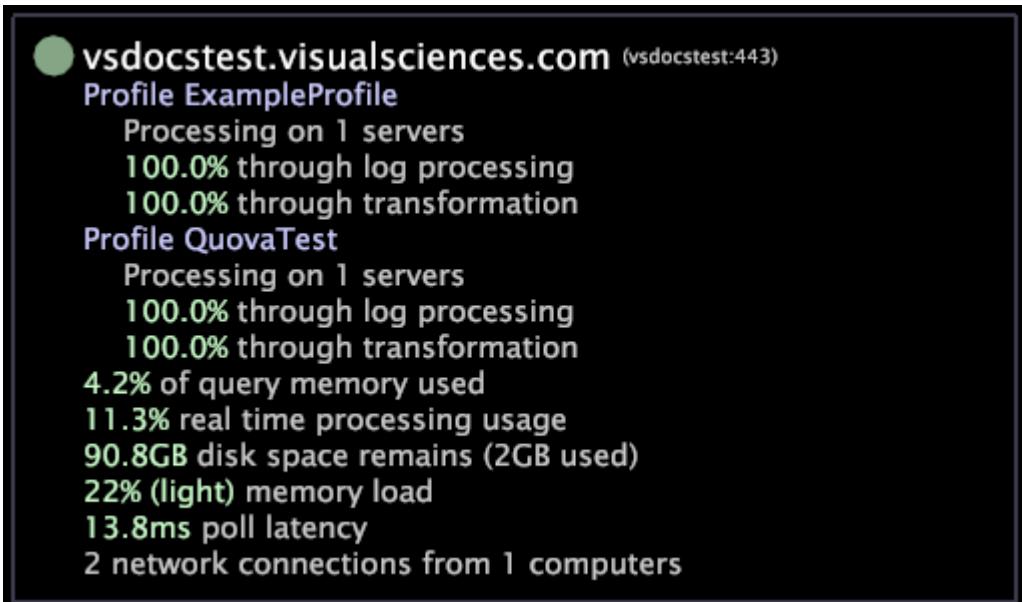
The lower portion of the Server Monitor interface lists the processing status of each of your available profiles as well as performance details about the computer.

For more information about Data Workbench servers, see the *Server Products Installation and Administration Guide*. For more information about Report, see the *Data Workbench Report Guide*.

To open the Server Monitor interface

- In the Servers Manager, right-click the node of the Data Workbench server or Report computer. t

Click **Server Monitor** to view details about one server, or click **Related Servers > Server Monitor List** to view details about a cluster of related servers.



The **Server Monitor** interface updates automatically every 10 seconds.

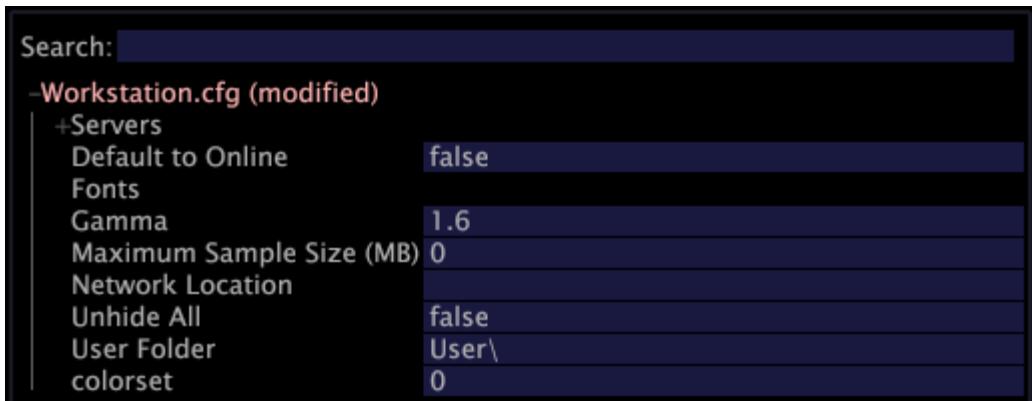
The following table lists the tasks that can be completed using the **Server Monitor** interface.

To perform this task...	Do this...
To check the log processing status of a profile	View the Profile <i>Profile Name</i> vector. In the example above, you would view the Profile ExampleProfile vector to see that the ExampleProfile profile processes on one server and its log processing is 100% complete.
To determine how long the computer takes to respond to requests	View the poll latency field. If this value is greater than 1000ms, contact Adobe Support Services.
To view an estimate of how long it might take to complete transformation or querying	View the sweep time (hh:mm:ss) field, which is present only during transformation or querying.
To determine the current number of network connections to the computer	View the last line of the computer's Server Monitor information. In the example above, you see that 2 network connections are currently coming from one computer.

Configuration option

Navigation title:

The Configuration option opens your `Insight.cfg` file, which controls your connections to various servers.

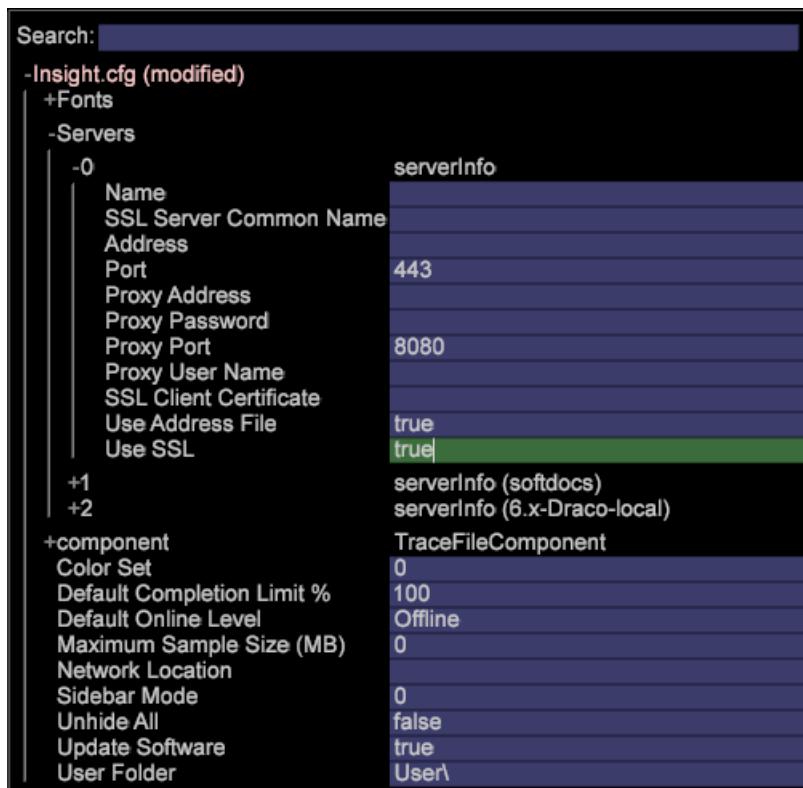


To edit the Insight.cfg file

1. In the Insight.cfg window, modify the parameters as desired. For detailed descriptions of the parameters in the Insight.cfg file, see [Configuration parameters](#) on page 510.
2. To save your configuration settings, right-click **Insight.cfg (modified)** at the top of the window and click **Save as Insight.cfg**.

To add new servers

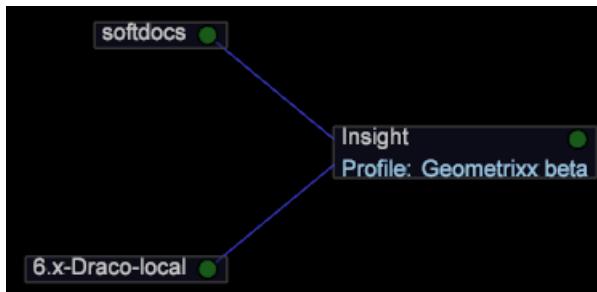
1. In the Insight.cfg window, right-click **Servers** and click **Add new child > Server**.



2. Complete or modify the server parameters to provide Data Workbench with access to the desired server. For detailed descriptions of the parameters in the Insight.cfg file, see [Configuration parameters](#) on page 510.
3. Repeat Step 1 and Step 2 for each server to which you want to configure a connection.

-
4. To save your configuration settings, right-click **Insight.cfg (modified)** at the top of the window and click **Save as Insight.cfg**.

Data Workbench attempts to connect to the server(s) using the settings that you have specified. If a connection is established, a green node appears in the **Servers Manager** as shown below. If Data Workbench cannot connect to the server, a red node appears.



Interface and analysis features

Navigation title:

You can customize the Data Workbench interface to meet your needs and configure certain features so that you can use them in your analyses. These sections describe how to perform customizations and configurations.

Customize Data Workbench locally

Navigation title:

Data Workbench is extremely flexible in the number of ways it can be customized.

All customizations made by a user to their instance of Data Workbench reside in the User folder within the Data Workbench installation directory. This directory is synchronized with the Data Workbench server to provide a backup of your local information.

To remove any customizations to Data Workbench, delete all files and folders within the Data Workbench installation directory except the five elements installed originally:

- Base folder
- Certificates folder
- Configuration folder
- `Insight.cfg` file
- `Insight.exe` file

Customize a menu

Navigation title:

You can customize the appearance of menus including the workspace window menu (accessed by right-clicking in any workspace) and menus listing metrics, dimensions, and map layers.

The hierarchy of any menu mirrors the structure of its directory in the **Profile Manager**. For example, the workspace window menu mirrors the structure of the `Menu` directory and the metrics menu mirrors the structure of the `Metrics` directory. For any menu, the corresponding directory may contain the following items:

- **Files:** These files represent the visualizations, legends, annotations, administrative interfaces, metrics, dimensions, or map layers that you can open by clicking the corresponding menu item.
- **An `order.txt` file:** This file specifies in what order the menu displays its items.
- **Subdirectories:** Each subdirectory represents a submenu. Each subdirectory may contain its own files, subdirectories, and `order.txt` file.

For example, the **Add Legend** menu contains the menu items Metric, Color, Value, and Confidence, in that order. In comparison, the **Menu\Add Legend** directory in the **Profile Manager** contains the files **Color.vw**, **Confidence.vw**, **Metric.vw**, and **Value.vw** files in alphabetical order, as well as an **order.txt** file to control the order of the other files.

Create a workspace menu and menu item

Navigation title:

These steps apply only to creating submenus and menu items for the **Workspace Window** menu.

You can customize the metric and dimension menus by creating new folders and new derived metrics and dimensions. For more information, see [Work with derived metrics](#) on page 461 and [Work with derived dimensions](#) on page 463.

To create a new workspace window menu

1. In the **Profile Manager**, click the **Menu** directory to view its contents.
2. In the **User** column for the **Menu** directory, click **Create > Folder**. A folder named **New Folder** appears in the **File** column for **Menu**.



Note: You also can create a new folder for any subdirectory within the **Menu** directory.

3. Rename the new folder by right-clicking in the **User** column for the folder and typing the new name in the **Dir** parameter.
4. Add the desired files to the new folder.
5. (Optional) In the **User** column for the new folder, click **Create > order.txt**.

An **order.txt** file appears in the **File** column for the new folder, and a check mark for the new file appears in the **User** column.

6. (Optional) Edit the **order.txt** file as needed. See [Customize a menu using order.txt files](#) on page 486.
7. (Optional) To make the changes available to all users of the working profile, right-click the white check mark for the folder in the **User** column and click **Save Directory to > <working profile name>**.

To add a new menu item to an existing menu

1. Complete one of the following steps:
 - Create a new item (visualization, annotation, and so on) to add to a menu:
 1. Open a workspace in Data Workbench and create the desired item.
 2. Save the item to the following directory:
Data workbench installation directory\User\working profile name\Work
 - 3. In the **Profile Manager**, click the **Work** directory to view its contents.
 - 4. In the **User** column, right-click the check mark for the desired file and click **Copy**.
 - 5. In the **User** column for the desired folder, click **Paste**.A copy of the file appears in the **File** column for the new folder, and a check mark for the new file appears in the **User** column.
- Copy and paste an existing item from one menu (folder) to another:
 1. In the **Profile Manager**, click the **Menu** directory to view its contents.
 2. In the *working profile name* column, right-click the check mark for the desired file and click **Make Local**.
 3. Right-click the check mark for the file in the **User** column and click **Copy**.

-
4. In the **User** column for the desired folder, click **Paste**. A copy of the file appears in the **File** column for the new folder, and a check mark for the new file appears in the **User** column.
 2. (Optional) Edit the `order.txt` file as needed. See *Customize a menu using order.txt files* on page 486.
 3. (Optional) To make the changes available to all users of the working profile, right-click the white check mark for each file in the **User** column and click **Save to > <working profile name>**.
 4. (Optional) To avoid having a menu item appear in multiple menus, you should delete the corresponding file from its original folder by right-clicking the check mark for the file in the *working profile name* column and clicking **Remove > Yes**.

Repeat this step for the file's check mark in the **User** column.

Customize a menu using order.txt files

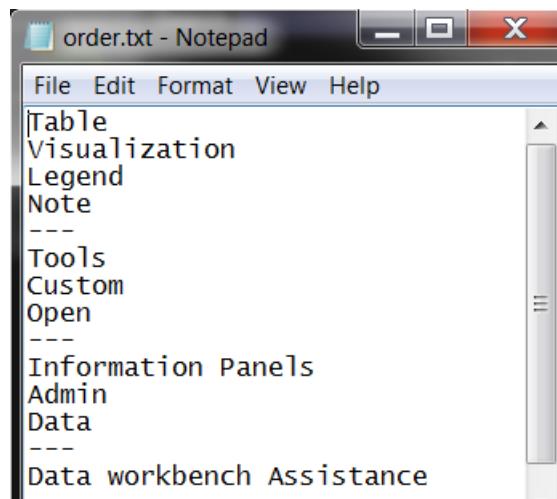
Navigation title:

You can customize the appearance of any menu by editing the `order.txt` file associated with that menu.

The steps in this section apply to all types of menus.

To edit the order.txt file to customize a menu

1. In the **Profile Manager**, in the *profile name* column, right-click the check mark for the `order.txt` file and click **Make Local**.
2. Right-click the check mark for the `order.txt` file in the **User** column and click **Open > in Notepad**. The `order.txt` file displays.



3. (Optional) Add or change the [Inclusive] or [Exclusive] setting at the top of the file if desired. This setting controls whether items not listed in the `order.txt` file but present in the **Profile Manager** is listed on the menu. The options include:
 - **[Inclusive]:** This is the default setting. This setting results in menu items that are not specified in the `order.txt` file being listed at the bottom of the menu in alphabetical order. For example, if the **Profile Manager** contained a Profile item in addition to those listed in the `order.txt` above, Profile would display below Data.
 - **[Exclusive]:** This setting results in menu items that are not specified in the `order.txt` file being excluded from the menu. For example, if the **Profile Manager** contained a Profile item in addition to those listed in the `order.txt` above, Profile would not be displayed anywhere on the menu.
 - **blank:** If neither [Inclusive] or [Exclusive] appears at the top of the file, Data Workbench displays the menu items as if the setting were [Inclusive].

4. Complete one or more of the following steps:

To perform this task...	Do the following...
Reorder menu items	<p>Type the item names in the order that you want them to appear in Data Workbench.</p> <p>For example, as long as each menu item name matches its corresponding file or folder name, the following would result in Add Table appearing first, then Add Visualization, Add Legend, and Add Note appearing last.</p> <p>Add Table Add Visualization Add Legend Add Note</p>
Rename a menu item	<p>Rename the corresponding file or folder in the Profile Manager, then change the name of the item in the <code>order.txt</code> file.</p> <p>For example, to rename Add Annotation to New Annotation, rename the Add Annotation folder in the Profile Manager to New Annotation, then change the name of the Add Annotation item in the <code>order.txt</code> file to New Annotation.</p>
Hide a menu item	<p>To hide the menu item but not delete the item itself, type a minus sign (-) at the beginning of its name.</p> <p>For example, the following results in Add Annotation not appearing in the menu.</p> <p>Add Legend -Add Annotation</p> <p>To again show the hidden menu item, simply remove the minus sign (-) or use the Unhide All parameter in the <code>Insight.cfg</code> file, see Configuration parameters on page 510.</p> <p>You can also hide menu items using the following methods:</p> <ul style="list-style-type: none"> The Show parameter in a <code>.filter</code>, <code>.metric</code>, or <code>.dim</code> file hides filters, derived metrics and dimensions, and extended dimensions from their respective menus. When using this option, the item is not listed in the menu, but it is still in the profile and available to be used. To use this parameter to hide filters and derived metrics and dimensions, add the following line to the end of the <code>.metric</code>, <code>.dim</code>, or <code>.filter</code> file: <pre>show = bool: false</pre> To use this parameter to hide extended dimensions, see Chapter 10 of the <i>Dataset Configuration Guide</i> for instructions. You can temporarily unhide items hidden using this method by setting the Unhide All parameter in the <code>Insight.cfg</code> file. For more information about this parameter, see Configuration parameters on page 510. The Hidden parameter in the <code>Transformation.cfg</code> file or any dataset include file hides extended dimensions from the dimension menu. When using this option, the item is not listed in the menu, but it is still in the profile and available to be used. <p> Note: When hiding extended dimensions using this method, you must retransform your dataset for the dimensions to be hidden.</p> <p>You can temporarily unhide items hidden using this method by setting the Unhide All parameter in the <code>Insight.cfg</code> file. For more information about this parameter, see Configuration parameters on page 510.</p> <ul style="list-style-type: none"> Zero-byte files hide any type of item on any menu. When using this option, an empty (zero-byte) file hides the presence of a file with the same name that contains data. Data Workbench treats zero-byte files as if they do not exist. For more information, see Hide a file by emptying it (zero-byte) on page 460.

To perform this task...	Do the following...
Delete a menu item	<p>If this file is set to use the [Exclusive] option, you can simply delete the menu item from this file. The item itself is still in the profile, but it is not listed in the menu.</p> <p>If this file is set to use the [Inclusive] option, you must remove the menu item name from this file and either delete or zero-byte the corresponding file to remove the item from the menu.</p> <p>For information about deleting files, see Delete files from your working profile on page 460. For information about zero-byte files, see Hide a file by emptying it (zero-byte) on page 460.</p>
Add a group header	<p>Type three hyphens before and after the heading text that you want to appear.</p> <p>For example, the following would result in a Manage group header for a set of related menu items.</p> <p>---Manage---</p> <p>Profile</p> <p>Dataset</p> 
Add a line to separate sections of a menu	<p>Type three hyphens where you want a line to appear.</p> <p>For example, the following results in a line separating Add Annotation and Add Custom.</p> <p>Add Annotation</p> <p>---</p> <p>Add Custom</p>

5. Save and close the file.
6. (Optional) To make the changes available to all users of the working profile, right-click the white check mark for the `order.txt` file in the **User** column and click **Save to > <working profile name>**.

Customize the profile manager and server file manager

Navigation title:

The Profile Manager and Server Files Manager display a large number of directories and files.

In some cases, you may want to access a subdirectory of one of these managers without having to navigate its entire directory structure. To do so, you can create a new manager that displays only those directories that you select.

Create a Profile Manager

Navigation title:

The **Profile Manager** displays all of the directories associated with your working profile.

You may want to access a subdirectory of the **Profile Manager** without having to navigate its entire directory structure. For example, the **Metrics** and **Workspaces** menu options available on the **Manage** menu of the workspace window menu enable you to open the Profile Manager Metrics and Workspaces folders, respectively.

For more information about the **Profile Manager**, see [Profile Manager](#) on page 456.

By default you have access to the following managers:

- **Metrics Manager:** Displays the contents of the Profile Manager's Metrics folder. You can open, edit, remove, or copy the metrics defined within each profile.
- **Reports Manager:** Displays the contents of the Profile Manager's Reports folder. You can open, edit, remove, or copy report workspaces or `report.cfg` files.

-
- **Workspaces Manager:** Displays the contents of the Profile Manager’s Workspaces folder. All of the files for configuring the **Worktop**’s tabs are located here. See [Customize a worktop tab](#) on page 490.

Data Workbench enables you to create additional profile managers that display one subdirectory from the **Profile Manager**. Each manager that you create must have a .vw file that specifies the **Profile Manager** directory whose contents it shows and the properties of that window. You can use the .vw file for any of the provided managers as a template.

To create a Profile Manager

1. In the **Profile Manager**, click the **Menu** directory to view its contents.
2. Within the **Menu** directory, click the **Admin** directory and then the **Profile** directory. The .vw files for the existing managers are located here.
3. In the *profile name* column, right-click the check mark for the one of the .vw files (for example, **Workspaces.vw**), then click **Make Local**.
A check mark for the file appears in the **User** column.
4. Right-click the check mark for the .vw file in the **User** column and click **Open > in Notepad**.
5. In the **Profile Path** field, type the **Profile Manager** directory for which you want to create a new manager. Be sure to include the slash (/) after the directory name.

```
window = simpleBorderWindow:  
client = scrollWindow:  
client = fileManager:  
    Profile Path = string: directory name/  
    size = v3d: (820, 5649, 0)  
    scroll_offset = v3d: (0, 0, 0)  
    size = v3d: (830, 881, 0)  
    pos = v3d: (525, 162, 0)  
    size = v3d: (830, 900, 0)
```

6. In Notepad, click **File > Save As** to save the edited file to the *Data Workbench installation folder\User\working profile name\Menu\Admin\Profile Management*.

Be sure to change the name of the .vw file to reflect the directory in the **Profile Manager** to which it corresponds.

7. (Optional) To make the changes available to all users of the working profile, right-click the check mark for the .vw file in the **User** column and click **Save to > <working profile name>**.

Create a Server Files Manager

The **Server Files Manager** displays all of the directories in the Data Workbench server installation directory, including configuration and look-up files.

You may want to access a portion of the **Server Files Manager** without having to navigate its entire directory structure or display only a few of the subdirectories to address a particular need. For example, you might want to create a separate **Server Files Manager** that displays only the Access Control and Users subdirectories, enabling you to manage your users and their access.

Each manager that you create must have a .vw file. You can use the **Server Files.vw** file as a template.

For more information about the **Server Files Manager**, see [Server Files Manager](#) on page 479.

To create a Server Files Manager

1. In the **Profile Manager**, click the **Menu** directory, then the **Admin** directory. The **Server Files.vw** file is located here.
2. In the *profile name* column, right-click the check mark for the **Server Files.vw** file and click **Make Local**. A check mark for the file appears in the **User** column.

-
3. Right-click the check mark for the **Server Files .vw** file in the **User** column and click **Open > from the workbench**. The **Server Files .vw** file opens.
 4. To remove a directory, right-click the top border of the **Server Files Manager** and click **Server Directories > Remove > <directory name>**.
 5. To add a directory, right-click the top border of the **Server Files Manager**, click **Server Directories > Add > Directory**.

Type the directory's URI in the **URI** field, and click **OK**.



Note: You can designate multiple directories in the **URI** field. For example, if you type **Profiles\Marketing**, the server files manager contain the Marketing subdirectory, but no other subdirectory within the Profiles directory.

6. Right-click the top the top border of the **Server Files Manager** and click **Save**.
7. To create a new manager, change the file name in the **File Name** field, then click **Save**. To overwrite the existing manager, click **Save**.
8. (Optional) To make the changes available to all users of the working profile, right-click the check mark for the **.vw** file in the **User** column.

Then, click **Save to > <working profile name>**.

Customize a worktop tab

Navigation title:

Each tab or subtab in the **Worktop** corresponds to a particular type of information, such as Dashboards, Activity, Acquisition, and so on.

For example, the **Acquisition** tab might contain workspaces that provide data about referring domains, search engines, and campaigns.

Each tab that appears in the **Worktop** corresponds to a folder in the *working profile name\Workspaces* folder within the Data Workbench installation directory. The order of the tabs in the **Worktop** is controlled by the **order.txt** file in that same folder. For example, if you have an **Acquisition** subfolder in the **Workspaces** folder and then you add **Acquisition** as the first entry in the **order.txt** file, **Acquisition** is the first tab in the **Worktop** and everything in that subfolder displays in the **Acquisition** tab.



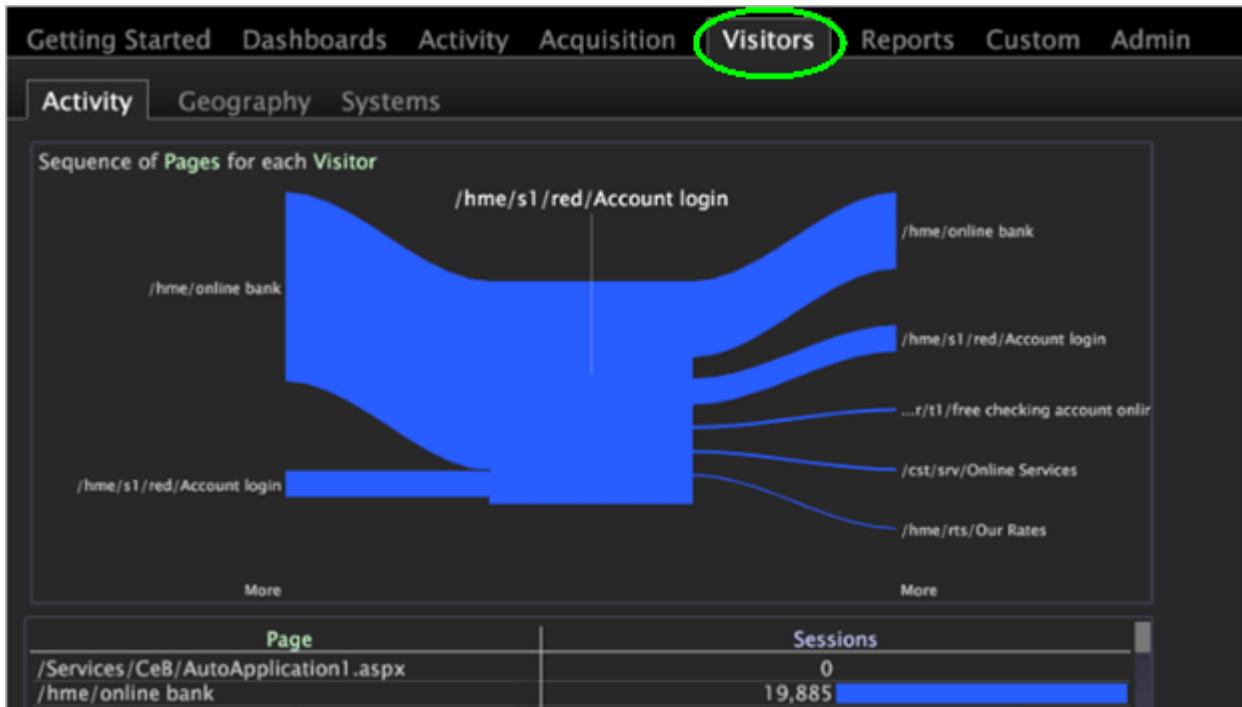
Note: For information about using the **order.txt** file to customize the workspace window menu, see [Customize a menu](#) on page 484.

Add a tab to the worktop

Navigation title:

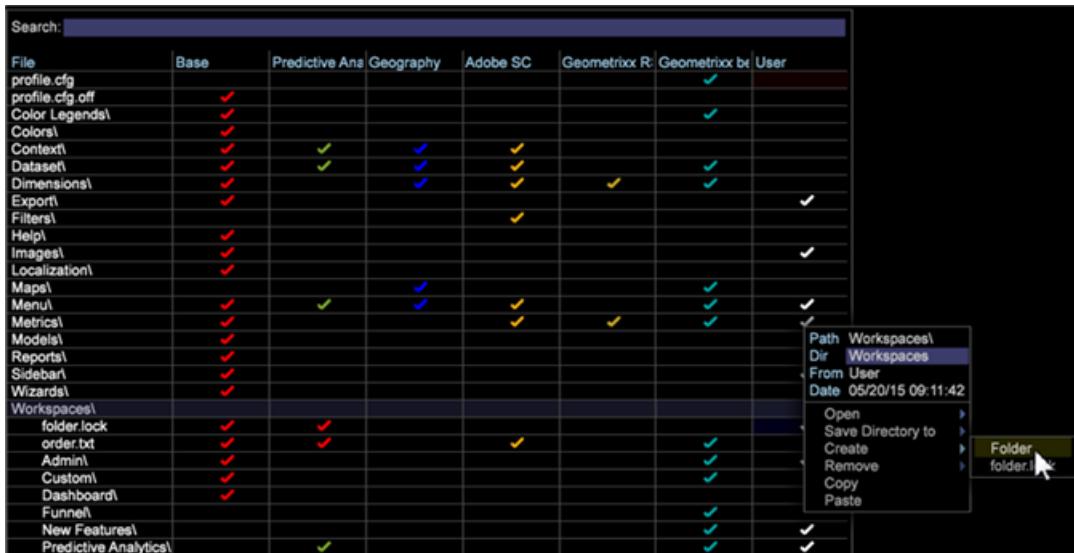
Add a tab to the Worktop to quickly access a workspace.

You can add a workspace as a tab on the Worktop. For example, you can add a **Visitors** tab for quick access to an overview of your traffic.



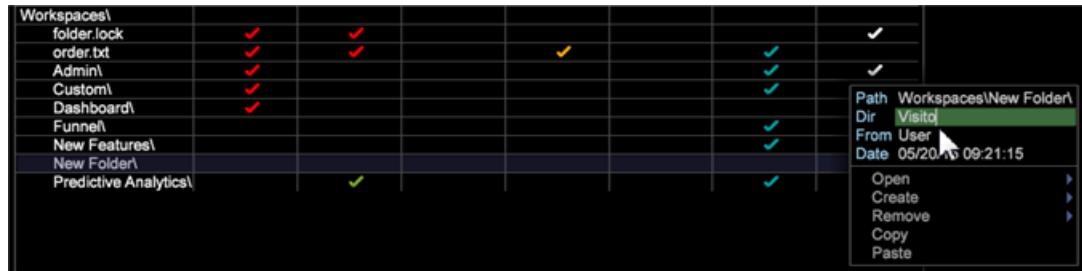
Add a folder in the Profile Manager

1. Open the workspace you want to associate with the new tab, then right-click **Admin > Profile Manager**.
2. In the **Profile Manager**, click **Workspaces** to view its contents.
3. Right-click in the **User** column for **Workspaces** and select **Create > Folder**.

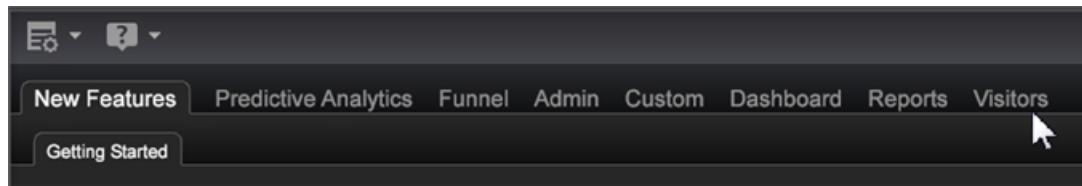


This creates a new subfolder to contain the files that you want to be available through your new tab. A new folder appears named New Folder.

4. To rename, right-click in the **User** column for the New Folder and type a name in the **Dir** parameter.



The folder name will appear in the worktop as a tab.



5. (Optional) To make the changes available to all users of the working profile, right-click the white check mark for the .vw file in the **User** column and click **Save to > <working profile name>**.

To display only the full-sized version of the description workspace and not the thumbnail of this workspace, you must save the workspace to the working profile. To do so, in the **Profile Manager**, right-click the white check mark for the description.vw file in the **User** column, then click **Save to > <working profile name>**.

Change the tab order

Navigation title:

Steps to change the order of the tabs that appear in the **Worktop**.

1. In the **Profile Manager**, click **Workspaces** to view its contents.
 2. In the *working profile name* column, right-click the check mark for the **Workspaces\order.txt** file and click **Make Local > Yes**.
 3. Right-click the check mark for the **Workspaces\order.txt** file in the **User** column and click **Open > in Notepad**.
 4. Rename, delete, or add tab names exactly as you want them to appear in the **Worktop**. For example, the following would result in the **Getting Started** tab appearing first and the **Acquisition** tab appearing last.
- Getting Started
Dashboards
Activity
Acquisition
5. Save and close the file.
 6. (Optional) To make the change available to all users of the working profile, right-click the white check mark for the **Workspaces\order.txt** file in the **User** column and click **Save to > <working profile name>**.

Display subfolders as subtabs

Navigation title:

By default, newly created tabs display the subfolders within the associated directory as hierarchical, drop-down subdirectories instead of as subtabs.

You can display subfolders as subtabs (as shown in the following example) by placing an `empty.folder.useTabs` file in the `working profile name\Workspaces\tab name folder` within the Data Workbench installation directory.

The following example shows the **Custom** tab with drop-down subdirectories.

Getting Started Dashboards Activity Acquisition Visitors Reports **Custom** Admin

Activity Collapse

Visitors by Hour of Day

Visitors by Hour of Day provides information about visitors, one-time, and repeat visitors by day, day of week, and hour of day.

Weekly Traffic Summary

The Weekly Traffic Summary provides a day-by-day view of many of your site's critical metrics and key performance indicators, as well as a summary of the cumulative page views of each page on your web site.

Dashboards 2 files

If you place an `empty.folder.useTabs` file in the `Workspaces\Custom` folder, all of the subfolders within the Custom folder display in the **Worktop** as subtabs, as shown in the following example:

Getting Started Dashboards Activity Acquisition Visitors Reports **Custom** Admin

Activity Dashboards

Visitors by Hour of Day

Visitors by Hour of Day provides information about visitors, one-time, and repeat visitors by day, day of week, and hour of day.

Weekly Traffic Summary

The Weekly Traffic Summary provides a day-by-day view of many of your site's critical metrics and key performance indicators, as well as a summary of the cumulative page views of each page on your web site.

To display subfolders as subtabs in the Worktop



Note: Each directory level must have a `Tab Name.useTabs` file for the contents of the subfolder to appear as subtabs instead of hierarchical, drop-down subdirectories.

1. In the **Profile Manager**, click **Workspaces** to view its contents.
2. In the *working profile name* column, right-click the check mark for one of the `folder.useTabs` files and click **Copy**.
3. Right-click in the **User** column for the `Workspaces\tab name` folder and click **Paste**. The subfolders within that tab now display as subtabs.
4. (Optional) To make this change available to all users of the working profile, right-click the white check mark for the new `folder.useTabs` file in the **User** column and click **Save to > <working profile name>**.

Hide or remove a tab

Navigation title:

Steps to hide, but not delete, a tab from the **Worktop**.

1. In the **Profile Manager**, click **Workspaces**, then click the desired *tab name* folder.
2. If you have an `order.txt` file in the *tab name* folder, right-click the check mark for the `order.txt` file in the *working profile name* column and click **Make Local > Yes**.

-or-

If you do not have an `order.txt` file in the *tab name* folder, right-click the check mark for an `order.txt` file in another folder and click **Copy**. Right-click in the **User** column for the `Workspaces\tab name` folder and click **Paste**.

3. Right-click the check mark for the `tab name\order.txt` file in the **User** column and click **Open > in Notepad**.
4. In the text file, type a minus sign at the beginning of the tab name. For example: “-Dashboard.”
5. Save and close the file.
6. (Optional) To make this change available to all users of the working profile, right-click the white check mark for the `order.txt` file in the **User** column and click **Save to > <working profile name>**.

Remove a tab from the **Worktop**

1. In the **Profile Manager**, click **Workspaces**, then click the desired *tab name* folder.
2. In the appropriate *profile name* column, right-click the check mark for the tab that you want to delete and click **Remove > Yes**.

Installing the Input Method Editor

Data workbench now supports the Input Method Editor (IME) as a secondary text entry process for international languages.

IMEs allow you to enter international characters using a variety of methods suited for your local language. Data workbench provides an input dialog box that allows you to open and use your desired IME for text fields.



Note: For the data workbench 6.1 release, only the virtual Simplified Chinese keyboard will be supported. Inputting other languages through the IME could result in unexpected behavior.

Using an IME

To use the floating IME text input feature:

-
1. Click **Alt + Space** for any text input area.
 2. Enter values using your system's IME.
 3. Close the input dialog by selecting the **Enter** key or clicking the **OK** button.

The dialog will disappear and the characters will then appear in the selected field.

Updating the Insight.cfg file

To employ the IME, you must update the `Insight.cfg` file with this setting:

```
Localized IME = bool: true
```

If this setting does not exist in the configuration file, then pressing **Alt + Space** will not engage the IME feature.

Starting Insight in another language: To better support localized assets like a splash screen and to support multiple languages in the future, data workbench requires command-line arguments identifying the language to load. The default language is English.

Starting data workbench in Chinese requires you to invoke `Insight.exe` with the "`-zh-cn`" argument:

```
Insight.exe -zh-cn
```

(These command line arguments are not case sensitive.)

Setting up Localized Languages

Set up `insight.zbin` file to set the language of the client application.

Update the data workbench server components

The administrator must first complete these tasks to update these server components:

1. Update to data workbench server 6.x.

You need to update the data workbench server for localization by updating the `base\localization*.zbin` file. This `insight.zbin` file will then be copied to the client.

An `insight.zbin` file is included in the installation folder alongside the `insight.exe` file. If you connect to a server that doesn't provide you with language-specific `.zbin` files, then data workbench will proceed to use this file.

The backup `insight.zbin` file can be provided in any language. As a result, if you use data workbench in Chinese and connect to a server that doesn't support this language, then your data workbench client will still be in Chinese, even if the server changes your base profile and removes your `.zbin` files from the `Base/Localization` folder.

2. Update the data workbench report server.

The `insight.zbin` at the root folder of data workbench report server will be in English by default. As the administrator, you will be required to select and copy the `.zbin` file from the updated report server package and place it in the root directory of the data workbench report server. Like the client, the report server also requires the proper arguments for the selected language, such as `Insight.exe -zh-cn`

- a. Stop the report server services.
- b. Copy the `Localization` folder from the new report server package.
- c. From the `Localization` folder, copy the `Insight.zbin` file and place it in the root directory of the report server where the `Insight.exe` is located.
- d. Add any required arguments, such as `insight.exe -zh-cn`
- e. Restart the report server.

Update the data workbench client

After updating the server, follow these steps to update each client.

1. To make sure the client does not get updated from the server during this update, set your `Insight.cfg` argument to False.

```
Update Software = bool: false
```

2. Restart the client.
3. Navigate to the Software and Docs profile (SoftDocs profile) and download the required `insight.zbin` file from the client package: `Software\Insight Client\Insight_6.1.zip`
4. Move the `insight.zbin` file to the folder where `insight.exe` is located.
5. To make sure that the client files now gets updated from the server, change the `Insight.cfg` file argument to True:

```
Update Software = bool: true
```

I



Note: Your client will synchronize with the server and you will see a message stating that it is updating. At the conclusion of the download, you will get a message asking if you want to restart your client.

6. Click **OK** to restart the client.

If you get the following message, then it means the `zbin` file was not placed in the same location as the `Insight.exe`.

```
Insight Terminated: The backup dictionary file insight.zbin  
is missing.
```

Localized Splash Screens

Data workbench looks for the following splash screen files:

- English (default): `Base/Images/<version_product>_Splash.png`
- Chinese (when started with -zh-cn): `Base/Images/<version_product>_Splash_zh-cn.png`.

If a splash screen is requested but missing, data workbench will access the English splash screen by default.

Configure a locked workspace

Navigation title:

Data Workbench can be configured to allow only certain users to change certain workspaces.

While a workspace is locked, users can make selections in most visualizations and sort the data in tables but cannot otherwise change the workspace. This functionality prevents users from making unintentional changes to the workspaces.

The following three elements work together to control the locking of workspaces:

- **A folder.lock or workspace name.lock file:** A `folder.lock` file specifies whether the workspaces in a particular folder are locked, while a `workspace name.lock` file specifies whether a particular workspace is locked.
- **The Unlock parameter in a user's Insight.cfg file:** This parameter specifies whether that user can temporarily unlock locked workspaces.
- **The Temporarily Unlock menu option:** When the `Unlock` parameter in the user's `Insight.cfg` is set to true, this option appears automatically in the title bar menu of each locked workspace to provide a way for the user to unlock it.

For information about `folder.lock` and `workspace name.lock` files, see the following section. For information about setting the **Unlock** parameter, see [Set the unlock parameter](#) on page 498. For information about the **Temporarily Unlock** menu option, see [Unlocking a workspace](#) on page 222.

Folder.lock and workspace.lock files

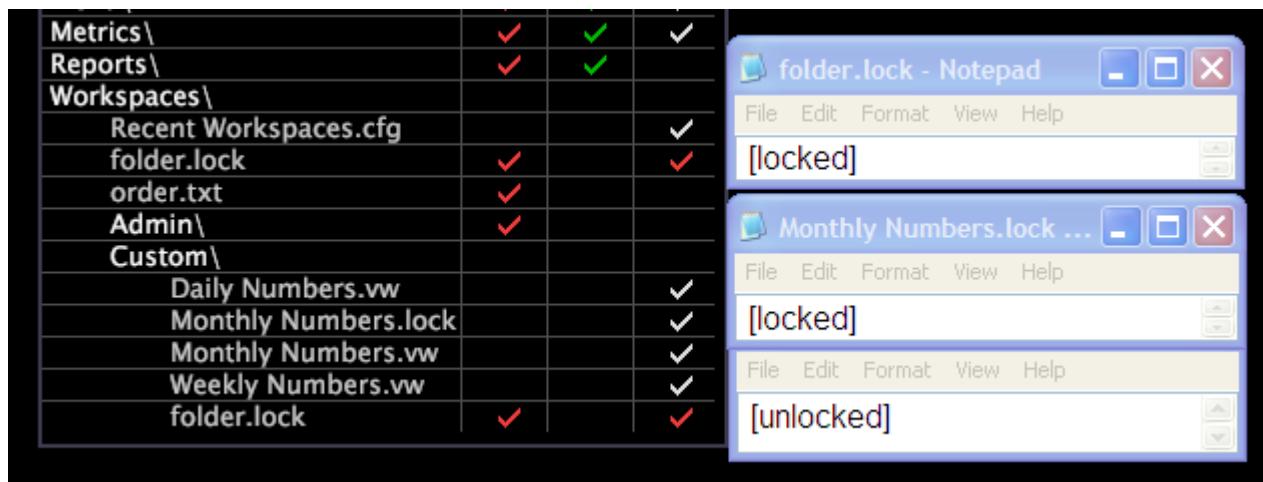
Navigation title: Using Folder.lock and Workspace.lock Files

Within the `Workspaces` folder of your Data Workbench installation directory, a `folder.lock` file specifies whether the workspaces in that particular folder are locked, while a `workspace name.lock` file specifies whether a particular workspace is locked.

When locking entire folders, you can lock them at the `Workspaces` folder level or at the sub-folder (tab) level. You also can lock or unlock all of your folders (at the `Workspaces` folder level), then specify the exceptions for particular sub-folders (using `folder.lock` files) or particular workspaces (using `workspace name.lock` files).

The following example of the **Profile Manager** highlights three elements:

- A `folder.lock` file for the main `Workspaces` folder
- A `Monthly Numbers.lock` file for the `Monthly Numbers.vw` workspace file
- A `folder.lock` file for the `Workspaces\Custom` sub-folder



In this example, the `Workspaces` `folder.lock` file is set to locked, which locks all of the workspaces in this instance of Data Workbench. The `Workspaces\Custom` sub-folder `folder.lock` file is set to unlocked, which unlocks all of the workspaces on the **Custom** tab. Finally, the `Monthly Numbers.lock` file is set to locked, which locks the `Monthly Numbers` workspace.

Creating .lock Files

You can create a new `folder.lock` file using the **Create** menu option in the **Profile Manager** or the **Workspaces Manager**. You also can create a `folder.lock` or `workspace name.lock` file by copying and pasting an existing `.lock` file to the appropriate folder, changing the name of the file (for `workspace name.lock` files only), and changing the setting in the file if necessary.

To create a new `folder.lock` file

1. In Data Workbench, open the **Workspaces Manager** by right-clicking within a workspace and clicking **Manage > Profile > Workspaces Manager**.
2. Click the folder for which you want to create a `folder.lock` file.
3. In the **User** column for that folder, right-click in the cell and click **Create > folder.lock**. A new `folder.lock` file appears. New `folder.lock` files are set to [unlocked] by default.
4. (Optional) If you need to change the setting in the file:

-
- a. Right-click the check mark for the file.
 - b. Click **Open > in Notepad**. The `folder.lock` file opens.
 - c. Change the setting to [locked].
 - d. Save and close the file.
5. To make this the setting for all users working with the same working profile, right-click the check mark for the file and click **Save to > <working profile name>**.

The workspaces in this folder are now locked or unlocked according to the setting in the new file.

To create a .lock file from an existing file

- 1. In the **Profile Manager** or **Workspaces Manager**, right-click the check mark for an existing `.lock` file and click **Copy**.
- 2. In the **User** column for the folder into which you want to paste the `.lock` file, right-click in the cell and click **Paste**.
- 3. If this file is used to lock an individual workspace, right-click the check mark for the `.lock` file in the **User** column and change its name in the **File** field to match the name of the workspace that you want to lock.

For example, to lock the `Monthly Numbers.vw` workspace, you would name the file “`Monthly Numbers.lock`.” If this file is used to lock an individual workspace, right-click the check mark for the `.lock` file in the **User** column and change its name in the **File** field to match the name of the workspace that you want to lock. For example, to lock the `Monthly Numbers.vw` workspace, you would name the file “`Monthly Numbers.lock`.”

- 4. To change the setting in the file:
 - a. Right-click the check mark for the file.
 - b. Click **Open > in Notepad**. The `.lock` file opens.
 - c. Change the setting to [locked] or [unlocked].
 - d. Save and close the file.
- 5. To make this the setting for all users working with the same working profile, right-click the check mark for the file and click **Save to > <working profile name>**.

The selected workspaces is now locked or unlocked according to the setting in the new file.

Set the unlock parameter

Navigation title:

The `Unlock` parameter in a user’s `Insight.cfg` file specifies whether that user has permission to temporarily unlock locked workspaces for editing.

If the `Unlock` parameter is already present in the user’s `Insight.cfg` file, you can edit the parameter using Data Workbench. If it is not already present in the user’s `Insight.cfg` file, you must add it to the file using a text editor, such as Notepad.

To set an existing Unlock parameter in the `Insight.cfg` file

- 1. In a workspace, right-click **Admin > Client Configuration**. The `Insight.cfg` file opens.
- 2. For the `Unlock` parameter, type either true or false. True enables the user to temporarily unlock any locked workspace, while false does not.
- 3. To save your configuration changes, right-click **Insight.cfg (modified)** at the top of the window and click **Save as Insight.cfg**.

To add the `Unlock` parameter to the `Insight.cfg` file

-
1. Navigate to your Data Workbench installation directory and open the `Insight.cfg` file using a text editor, such as Notepad.
 2. Add the `Unlock` parameter to the end of the file, as in the following example:

```
Unlock = bool: false
```
 3. Set the value to either true or false. True enables the user to temporarily unlock any locked workspace, while false does not.
 4. Save and close the file.

Configure open functionality

Navigation title:

Open functionality enables you to open such items as documents or URIs in such external applications as a text editor or a web browser.

Open functionality is currently configured only in the **Site** application and only for opening URIs. This section provides information about configuring Open URI functionality for **Site**. For information about configuring Open functionality for another application or to open other items, contact Adobe Consulting Services.

In **Site**, you can right-click a URI from a page overlay or table to display the URI in the application for which it was formatted. For example, from a URI table visualization, you can click a URI to display a web page in a web browser.

To open a URI from a visualization, you first must edit the `Open_URI.cfg` file for the URI dimension to identify the location and naming conventions that Data Workbench uses to open the URI. To view a URI in its native format (such as HTML), Data Workbench must have access to the referenced location and the application needed to open that item. For example, to view a web page, Data Workbench would need access to the Internet as well as have a web browser installed.

To edit `Open_URI.vw`

1. In the **Profile Manager**, click **Context > Dimension Element > URI**.
2. In the URI folder, right-click the check mark next to the `Open_URI.vw` file and click **Make Local**. A check mark for this file appears in the **User** column.
3. Right-click the newly created check mark and click **Open > in Insight** if the file is a `.cfg` file or **Open > in Notepad** if it is a `.vw` file.
4. Click **Command**, then **Parameters** to view the contents of the file.
5. Modify the **Site** parameter and the **Template** parameter as necessary:

For this parameter...	Provide this information...
Site	The location of the URIs that you want to open. Example: <code>mysite.com</code>
Template	The parameters that Data Workbench should use to locate and open the URIs. Example: <code>http://%Site%%URI%</code> The default template shown in the example tells Data Workbench to open a web browser, look for the location defined in the Site parameter, then locate the URI dimension element you are attempting to open.

6. Save the file.
7. To make this change available to all users of the working profile, right-click the check mark for the `.vw` file in the **User** column and click **Save to > <working profile name>**.

Configure a callout

Navigation title:

Callouts bring attention to a particular dimension element by creating a new visualization with a virtual selection of a particular dimension element in a visualization.

You can add or edit callouts by configuring the callout files stored in a Profiles\profile name\Context\Callout folder of the Server installation folder. Callouts that bring attention to a particular metric in a worksheet visualization are called metric callouts. Metric callout files are stored in the Profiles\profile name\Context\Metric Callout folder.

For instructions to add a callout or metric callout to a visualization, see [Adding callouts to a workspace](#) on page 293.

To create a new type of callout

1. In any workspace, create a visualization containing the data that you want to appear in the new callout type. For example, if you want your callout to be a table, you create a table visualization showing the desired metric and dimension.
2. Right-click the top border of the callout window and click **Save**.
3. In the **Save** window, click  , double-click **Context**, then double-click **Callout**. In the **File Name** field, type a name for the file and click **Save**. The callout file is saved to the User\working profile name\Context\Callout folder.



Note: When naming your callout, choose a descriptive name that reflects the visualization type and the metric and dimension that it shows. For example, if you want to create a callout from a table visualization showing the Sessions metric over the Day dimension, you could name the callout “Sessions by Day Table.”

4. (Optional) To make this change available to all users of the working profile:
 - a. In the **Profile Manager**, click **Context**, then click **Callout**. This folder contains all of the visualization files (.vw) that define the existing callout types.
 - b. Right-click the check mark next to the file name of the new callout in the **User** column and click **Save to > <working profile name>**.

To change a callout to a metric callout

1. In the **Profile Manager**, click **Context**, then click **Callout**. This folder contains all of the visualization files (.vw) that define the existing callout types.
2. Right-click the check mark next to the file name of the type of callout that you want to change and click **Make Local**. After the file has been downloaded to the local computer, a check mark appears in the **User** column.
3. Right-click the check mark next to the file name in the **User** column and click **Open > In Notepad**.
4. Locate the metric_y = ref : entry in the callout file and replace the existing value with the word Metric. The highlighted text in following file fragment shows where you insert this word.

```
window = simpleBorderWindow:  
    client = graph:  
        bars = bool: true  
        dim_x = ref: wdata/model/dim/dimension name  
        lines = bool: false  
        metrics = vector: 1 items  
            0 = gr_metric:  
                metric_y = ref: Metric  
                yaxis = axisLegend:  
                    max_value = double: maximum y-value
```

```
min_value = double: minimum y-value  
zoom_max = double: maximum y-zoom  
zoom_min = double: minimum y-zoom  
.
```

5. Click **File > Save As**. In the **Save As** window, click once, then double-click **Metric Callout**. In the **File Name** field, type a name for the file and click **Save**. The metric callout file is saved to the `User\working profile name\Context\Metric Callout` folder.
6. (Optional) To make this metric callout available to all users of the working profile, in the **Profile Manager**, right-click the check mark next to the file name in the **User** column and click **Save to > <working profile name>**.

Configure a latency table

Navigation title:

A latency table is a table that includes a latency dimension, which is a type of derived dimension that measures the time that has elapsed since a particular event occurred.

You define the event for measuring latency by making selections within one or more visualizations and setting those selections as the event using the **Set Event menu** option within the latency table. For instructions to create and reuse latency tables, see [Latency tables](#) on page 375.

By default, latency tables provide information about the visitors' sessions that occurred as many as seven days before or after the event. Using the instructions in the following sections, you can configure latency tables to provide information about different countable and time dimensions.

Create a latency dimension

Navigation title:

Latency dimensions are constructed from a parent countable dimension, such as Sessions, and a time dimension, such as Day.

When you create a latency table in Data Workbench, you automatically add a latency dimension to the visualization file (`.vw`). You can edit a table's latency dimension by following the steps below.

To edit a latency dimension

1. Open the latency table that you created in a text editor such as Notepad. It is located in the `User\working profile name\Work` folder within your Data Workbench installation directory.

The defined latency dimension includes the parameters shown in the following example. (The definition of your latency dimension may include additional parameters.) The line `entity = LatencyDim:` indicates the start of the latency dimension's definition.

```
entity = LatencyDim:  
Name = string: dimension name  
Level = ref: wdata/model/dim/level  
Clip = ref: wdata/model/dim/clip  
Time = ref: wdata/model/dim/time dimension  
Format = printf_format:  
format = string: %+0.0f time string  
offset = double: offset  
Time Before = int: time before  
Time After = int: time after
```

2. Edit the values for the Name, Level, Clip, Time, Format, Time Before, or Time After parameters using the following table as a guide:

For this parameter...	Provide this information...
Name	Optional. The name of the latency dimension that appears in the context menu when you right-click the dimension label or elements.
Level	A countable dimension that is the parent of the latency dimension. Examples include Session, Visitor, and Page View.
Clip	A countable dimension that has a one-to-many relationship with the level of the latency dimension. Latency is not calculated across the boundaries of this dimension. For example, if you specify a level of Page View and a clip of Session, latencies are calculated for those page views that occurred during the same session as the event. For information about one-to-many (simple) dimensions, see the <i>Dataset Configuration Guide</i> .
Time	The dimension used to measure elapsed time for the latency dimension. This dimension can be a time dimension, such as Day or Hour, or a countable dimension, such as Visitor, Session, or Page View.
Format	Optional. Specifies the appearance of the latency visualization in Data Workbench. Within the Format parameter, you can edit the following values: <ul style="list-style-type: none">• Time string. The unit of time shown in the latency visualization, such as day or week. Be sure to change the time string when you change the time dimension.• Offset. A whole number that is equal to the negative of the value for Time Before. For example, if Time Before is 7, the offset should be -7.
Time Before	The maximum amount of time (expressed in the units of the Time dimension) before the event for which latency is calculated. If this value is set to 0 or not set at all, latency is calculated only for the forward direction. If you change this value, be sure to change the offset value in the Format parameter: The offset is the negative of the value for Time Before.
Time After	The maximum amount of time (expressed in the units of the Time dimension) after the event for which latency is calculated.

- Save the .vw file to the User\working profile name\Work folder.

Following are the settings for the default latency dimension:

```
entity = LatencyDim:
Name = string:
Level = ref: wdata/model/dim/Session
Clip = ref: wdata/model/dim/Visitor
Time = ref: wdata/model/dim/Day
Time Before = int: 7
Time After = int: 7
```

In the following latency dimension, the latency of each session event is calculated in hours and Time Before is set to zero. Therefore, latency is calculated for only those sessions that occurred within 24 hours after the defined event.

```
entity = LatencyDim:
Name = string:
Level = ref: wdata/model/dim/Session
Clip = ref: wdata/model/dim/Visitor
Time = ref: wdata/model/dim/Hour
Time Before = int: 0
Time After = int: 24
```

Configure a page overlay

Navigation title:

Page overlays are configured only in the **Site** application, but they can be configured for other applications.

For information about configuring page overlay for another application, contact Adobe Consulting Services.

The page overlay visualization is a tool for HTML link analysis. When you request an overlay for a particular page, Data Workbench takes a snapshot of the actual page as it would appear in a web browser and parses the HTML code that represents links according to a list of regular expressions that you define. For each link on the selected page, the software attempts to find a regular expression pattern match by working down the list until the first match is found. If there is a match, the link appears highlighted in the page overlay.

Page overlay shows data only when you add a color legend to the workspace containing the page overlay.



Note: Configuring page overlay requires careful configuration work, and it is possible to create misleading results if links are inappropriately mapped to the data. The work involved in configuring page overlay for a specific site depends on how links are presented within the HTML code on the site's pages.

Page overlay, by its nature, suggests to the user the mental model that it displays “where people click.” If the data backing the visualization does not match this model, the potential for confusion is high.

In **Site**, a link typically represents an element from the Next URI or Next Link dimension, but you can map a link to any dimension that makes sense for your analysis. For information about configuring page overlay for other dimensions, contact Adobe Consulting Services.



Note: Using the Page dimension for page overlay is not recommended. Users can rename the elements of the Page dimensions, thereby altering the link syntax on which the page overlay functionality relies.

To configure page overlay for **Site**, you must edit two files:

- **Page Overlay.vw:** This file is a template file for creating page overlay visualizations. At least one template file must be present in the profile for which you are configuring page overlay.
- **Page Overlay Link Templates.cfg:** When the page overlay visualization loads a page, it automatically identifies the links in the page and their destinations. To relate these links to elements in the data, you must define a set of regular expressions in this file.

You can define multiple regular expressions to match against the elements of the dimension. The order in which you define the expressions is important. When you request an overlay for a particular page, Data Workbench takes a snapshot of the actual page as it would appear in a web browser and parses the HTML code that represents links according to a list of regular expressions that you define. For each link on the selected page, the software attempts to find a regular expression pattern match by working down the list until the first match is found. The first expression to match a dimension element is the one used. Therefore, it is best to list the regular expression with the most specific matching pattern first, followed by less specific expressions. If there is a match, the link appears highlighted in the page overlay visualization.

To configure page overlay for Site

1. In the **Profile Manager**, navigate to **Context > Dimension Element > URI**.



Note: The Dimension Element directory contains the context menu items that appear when you right-click a dimension element. For example, open a URI table, then select a URI element. Right-click the URI and Page Overlay appears.

-
2. In the URI folder, right-click the check mark next to the `Page_Overlay.vw` file and click **Make Local**. A check mark for this file appears in the **User** column.
 3. Right-click the newly created check mark and click **Open > in Notepad**.
 4. Specify the Domain (and Browser Height, if required).

```
window = simpleBorderWindow:
    client = scrollWindow:
        client = PageOverlay:
            URI Template = string: http://%Domain%&Element%
            URI Parameters = map:
                Domain = string: domain name
                Element = ref: Element/Name
                Dim = ref: wdata/model/dim/URI
                Dim Element = ref: Element/Name
                Level = ref: wdata/model/dim/Page View
                Group = ref: wdata/model/dim/Session
                Browser Height = int: browser height
            pos = v3d: (518, 202, 0)
            size = v3d: (810, 610, 0)
            titleBar = editor:
                size = v3d: (61, 19, 0)
            text = string:
```

5. Save the file.
6. To make this change available to all users of the working profile, in the **Profile Manager**, right-click the check mark for the `.vw` file in the **User** column and click **Save to > <working profile name>**.



Note: You can create additional template files for other sites or subdomains. Each template that you create appears in the **Page Overlay menu**.

7. In the Context folder of the **Profile Manager**, right-click the check mark next to the `Page_Overlay Link Templates.cfg` file and click **Make Local**.
A check mark for this file appears in the **User** column.
8. Right-click the newly created check mark and click **Open > from the workbench**.
9. Right-click **Link Templates** and click **Add new > Regular Expression**.
10. Edit the parameters for LinkRegex vector as necessary:

For this parameter...	Provide this information...
Dimension	The dimension (typically the Next URI dimension) that is represented by the link.
Expression	The regular expression used to select the relevant part of the HTML link to find the next element from the Dimension. The regular expression must be an exact match, and the desired output pattern is grouped with parentheses. For details about regular expressions, see the <i>Dataset Configuration Guide</i> .
Output Pattern	The output pattern of the regular expression used to extract the resulting element of the Dimension parameter.

The following sample file shows three regular expressions:

-Page Overlay Link Templates.cfg	
+Link Parameters	map
-Link Templates	
-0	LinkRegex
Dimension	Next URI
Expression	https://www.site.com(/[^?]*\?product=[^&]*.*
Output Pattern	%1%
-1	LinkRegex
Dimension	Next URI
Expression	(/[?]*\?product=[^&]*).*
Output Pattern	%1%
-2	LinkRegex
Dimension	Next URI
Expression	https://www.site.com(/[^?]*\?spot=[^&]*.*
Output Pattern	%1%

11. To save the file, right-click (**modified**) at the top of the window and click **Save**.
12. To make this change available to all users of the working profile, right-click the check mark for **Page Overlay Link Templates.cfg** in the **User** column and click **Save to > <working profile name>**.

Configure a path browser

Navigation title:

Path browsers can be configured to work with any combination of base dimension, group dimension, level dimension, and metric that makes sense for your application and dataset.

After you configure a path browser, it is listed with other path browsers in the **Add Visualization** menu.

1. In the **Profile Manager**, click **Menu**, then click **Add Visualization** and **Path Browser**.

At least one *.vw file resides in the Path Browser directory.

2. Right-click the check mark for the desired file and click **Make Local**.
3. Right-click the check mark for the file in the **User** column and click **Open > in Notepad**.
4. Edit the parameters of the file using the following sample file and table as guides:

```
window = simpleBorderWindow:
  client = pathBrowser:
    Left Path = vector: 0 items
    Map = ref: wdata/model/dim/base dimension name
    Map Group = ref: wdata/model/dim/group dimension name
    Map Level = ref: wdata/model/dim/level dimension name
    Metric = ref: wdata/model/metric/metric name
    Right Path = vector: 0 items
    size = v3d: (673, 279, 0)
  pos = v3d: (714, 143, 0)
  size = v3d: (673, 298, 0)
```

For this parameter...	Provide this information...
<i>Base dimension name</i>	The name of the dimension whose elements appear on the path browser.

For this parameter...	Provide this information...
<i>Group dimension name</i>	The name of the dimension that determines how the elements of the level dimension are grouped to form the paths in a path browser. Specifically, the level dimension elements associated with a single path in a path browser cannot span more than one element of a group dimension.
<i>Level dimension name</i>	The name of the level (parent) of the base dimension whose elements you drag to the path browser. As you follow a path from one base dimension element to the next, you move from one level dimension element to the next. When you select a path of base dimension elements, you are selecting data for the corresponding elements of the level dimension. The selection always includes the elements of the level dimension that relate to the root, and it is refined by adding more elements to the path.
<i>Metric name</i>	The name of the metric whose value for a given element is proportional to the thickness of the path leading to that element.



Note: For more information about the base dimension, group dimension, level dimension, and metric for a path browser, see [Path browsers](#) on page 383.

5. In Notepad, click **File > Save As** to save the file with a new name based on the group dimension, that is, *Group dimension name.vw*.

Make sure that you save the file to the Path Browser directory.



Note: To make sure that your path browser is saved as a * . vw file, in the **Save As** window, set Save as type to All Files.

6. (Optional) To make the changes available to all users of the working profile, in the **Profile Manager**, right-click the check mark for the file in the **User** column and click **Save to > <working profile name>**.

Configure a process map

Navigation title:

Process maps can be configured to work with any combination of base dimension, group dimension, level dimension, and metric that makes sense for your application and dataset.

After you configure a process map, it is listed with other process maps in the **Add Visualization menu**.

1. In the **Profile Manager**, click **Menu**, click **Add Visualization**, then click the type of process map type that you want to configure (2D Metric Map, 2D Process Map, or 3D Process Map).

At least one * . vw file resides in the directory.

2. Right-click the check mark for the desired file and click **Make Local**.
3. Right-click the check mark for the file in the **User** column and click **Open > in Notepad**.
4. Edit the parameters of the file using the following sample file and table as guides:

```
window = simpleBorderWindow:
client = processMap:
Traffic Metric = ref: wdata/model/metric/metric name
center = v3d: (-0.741014, 0, 0.0639476)
color_links = bool: true
Map = PrefixDim:
Name = string: Map
Base Dimension = ref: wdata/model/dim/base dimension name
Element Prefixes = vector: 0 items
Element Names = vector: 0 items
Map Level = ref: wdata/model/dim/level dimension name
```

```

Map Group = ref: wdata/model/dim/group dimension name
node_label = vector<bool>:
node_positions = vector: 0 items
quantify_links = int: 2
range = double: 2.37386
size = v3d: (430, 290, 0)
xAxisMetric = ref: wdata/model/metric/metric name for metric map
pos = v3d: (880, 595, 0)
size = v3d: (430, 309, 0)

```

For this parameter...	Provide this information...
<i>Metric name</i>	The name of the metric whose value for a given node is proportional to the size of the node.
<i>Base dimension name</i>	The name of the dimension whose elements appear as nodes on the process map.
<i>Level dimension name</i>	The name of the level (parent) of the base dimension whose elements you drag to the process map.
<i>Group dimension name</i>	The name of the dimension that determines how the elements of the level dimension are grouped to form the connections between nodes. A connection between two nodes cannot span more than one element of a group dimension. When you make a selection based on a node within a process map, you are selecting all of the elements of the group dimension that involved that node.
<i>Metric name for metric map</i>	This parameter applies to 2D metric maps only. The name of the metric whose value determines the horizontal position of the nodes on the map.



Note: For more information about the base dimension, group dimension, level dimension, and metric for a process map, see [Process map](#) on page 390.

5. In Notepad, click **File > Save As** to save the file with a new name based on the base dimension, that is, *Base dimension name.vw*.

(If you are configuring a 2D metric map, you should save the file with a name based on the metric name for the metric map, that is, *Metric name for metric map.vw*.) Make sure that you save the file to the appropriate process map directory.



Note: To make sure that your process map is saved as a *.vw file, in the **Save As** window, set Save as type to All Files.

6. (Optional) To make the changes available to all users of the working profile, in the **Profile Manager**, right-click the check mark for the file in the **User** column and click **Save to > <working profile name>**.

Configure the Dataset Schema Interface

Navigation title:

Steps to change the default visualization.

You can control which type of visualization displays when you click a dimension name in a **Dataset Schema Interface** by adding files to the `Profiles\profile name\Context\Dimension Legend` folder of the Data Workbench server installation folder. The `Default.1d` file in this folder controls the default visualization type for all of the dimensions. By adding a `dimension name.1d` file (such as `Hour.1d`) to this folder, you can control the default visualization for that particular dimension.

For more information about **Dataset Schema Interfaces**, see [Dataset Schema interface](#) on page 446.

To change the default visualization

1. In any workspace, create a visualization containing the data that you want to appear in the new default visualization.

For example, if you want the dimension to display in a bar graph, create a bar graph visualization showing the desired metric and dimension.

2. Right-click the top border of the callout window and click **Save**.
3. In the **Save** window, click , double-click **Context**, then double-click **Dimension Legend**.
4. In the **File Name** field, type the dimension name.

The name of the .1d file must match the name of the dimension exactly. For example, Hour.1d.
5. Change the file extension to “1d” and click **Save**.

The file is saved to the User\working profile name\Context\Dimension Legend folder.
6. (Optional) To make this change available to all users of the working profile:
 1. In the **Profile Manager**, click **Context**, then click **Dimension Legend**.
 2. Right-click the check mark next to the file name of the new callout in the **User** column and click **Save to > <working profile name>**.

Localizing Time Dimensions

Configure the time dimensions to display correctly for the locale.

You can configure the displayed format of time dimensions based on locale in the **Standard Time Dimensions.cfg** file (located by default at **Server/Profiles/<my profile>/Dataset/Transformation/Time/Standard Time Dimensions.cfg**).

For example, in North America you can express the date May 3rd, 2015 as 5/3/15, or %m/%d/%y. However, in other parts of the world this could be interpreted as %d/%m/%y, or March 5th, 2015 due to an ambiguity in the values. To avoid this situation, an administrator might want to change the displayed format to match the expectations of the users in a locale.

1. Override Default Time Dimensions in Standard Time Dimensions.cfg

To enable this feature, the administrator must override the defaults by either editing the existing time dimensions or by creating new time dimensions with additional parameters.

An example of a modified time dimension follows.

The **Format** values for Week, Hour, Day, Month, and Hour of Day are set to the defaults in the example.



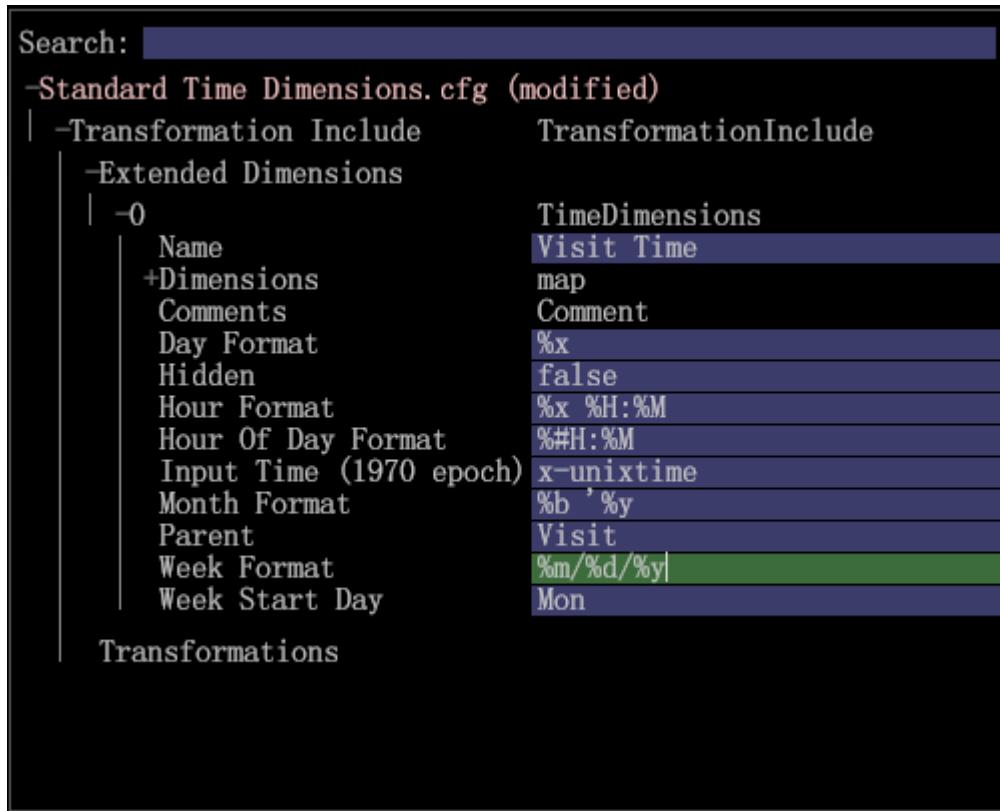
Note: If these lines are omitted, Data Workbench's behavior won't change and the dimension will be compiled using the defaults.

```
Transformation Include = TransformationInclude:  
Extended Dimensions = vector: 1 items  
0 = TimeDimensions:  
Comments = Comment: 0 items  
Dimensions = map:  
Day = string: Day
```

```

Day of Week = string: Day of Week
Hour = string: Hour
Hour of Day = string: Hour of Day
Month = string: Month
Week = string: Week
Hidden = bool: false
Input Time (1970 epoch) = string: x-unixtime
Week Format = string: %m/%d/%y
Hour Format = string: %x %H:%M
Day Format = string: %x
Month Format = string: %b '%y
Hour Of Day Format = string: %#H:%M
Name = string: Visit Time
Parent = string: Visit
Week Start Day = string: Mon
Transformations = vector: 0 items

```



2. Configure the `meta.cfg` file

Additionally, it's necessary for the package administrator to add these parameters and their defaults to profile's `meta.cfg` file. This allows editing from the workstation.

Here's an excerpt from a configured `meta.cfg` file.

```

dimensions = vector: 6 items
0 = Template:
...
...
5 = Template:

```

```

name = string: Time Dimensions
value = TimeDimensions:
  Name = string:
  Comments = Comment: 0 items
  Hidden = bool: false
  Week Format = string: %d/%m/%y
  Hour Format = string: %x %H:%M
  Day Format = string: %x
  Month Format = string: %b '%y
  Hour Of Day Format = string: %#H:%M
  Input Time (1970 epoch) = string:
  Parent = string:
  Week Start Day = string: Mon
  Dimensions = map:
    Hour of Day = string: Hour of Day
    Day of Week = string: Day of Week
    Hour = string: Hour
    Day = string: Day
    Week = string: Week
    Month = string: Month

```

Here is an example of a **meta.cfg** file in the workstation:

The screenshot shows the Logi Workbench interface. On the left, there is a 'File Manager' tree view with various configuration files listed. On the right, there is a search results panel with a tree view of the 'meta.cfg' file structure and its properties.

File	Base	Geography	Predictive	Adobe SC	Attribution	Geometrix	qa-cmp-ge	User
legacy-segment-export.conf							✓	
profile.cfg							✓	
profile.cfg.off	✓							
profile.cfg.hide.6.11								
profile.cfg.hide.6.2								
profile.cfg.hide.6.3							✓	
test.conf								
Color Legend\	✓							
Colors\	✓							
Context\								
Base_6.40.ver	-							
Formats.cfg	✓							
Internal.cfg	✓							
Page Overlay Link Templates.cfg	✓							
WebSupportedVis.cfg	✓							
meta.cfg	✓							
Callout\	✓							
Clustering\								
DependencyMap\	✓							
Dimension Class Element\	✓							
Dimension Class\	✓							
Dimension Element\								
Dimension Legend\	✓							
FileEditor\	✓							
FileNew\	✓							
FileOpen\	✓							
Filter Class\	✓							
Funnel\	✓							
Metric Callout\	✓							
Metric Class\	✓							
MultiServerDetails\	✓							
Page Size\	✓							
Path Browser\	✓							
Pearsons\								

Search results for 'Time Dimensions':

- meta.cfg
- +DecoderGroups
- +XMLDecoderTables
- +conditions
- dimensions
 - +0
 - +1
 - +2
 - +3
 - +4
 - +5
- value
 - Name
 - +Dimensions
 - Comments
 - Day Format
 - Hidden
 - Hour Format
 - Hour Of Day Format
 - Input Time (1970 epoch)
 - Month Format
 - Parent
 - Week Format
 - Week Start Day
 - name
- +logSources

The administrator is then able to go into the **File Manager**, open the file(s) where the time dimensions are configured (e.g., **Standard Time Dimensions.cfg**), and edit them using in the workstation.

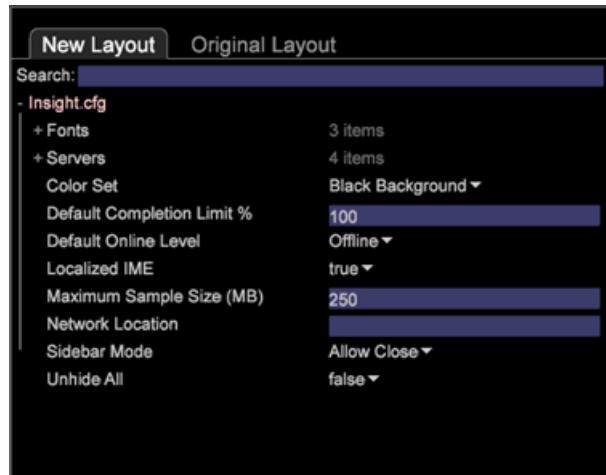
Configuration parameters

Navigation title:

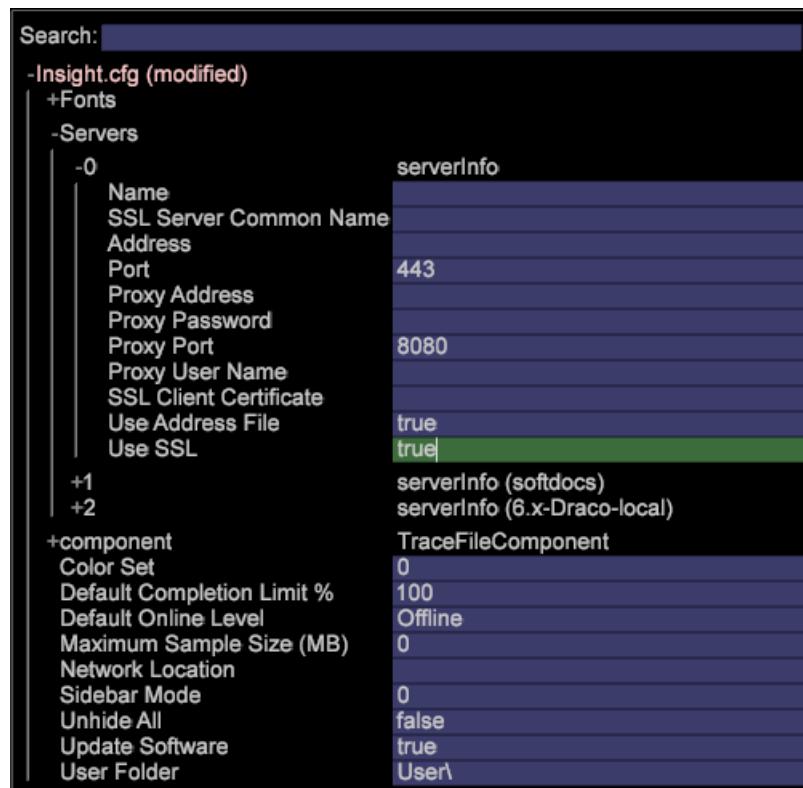
Set the parameters in the `Insight.cfg` file to specify your network connection and Data Workbench configuration settings.

The following example contains only the parameters included in the `Insight.cfg` file by default.

New Layout view



Original Layout view



Some of the parameters available in the new `Insight.cfg` file may not available in your version of the `Insight.cfg` file. If one of these parameters is needed, you must add it to the `Insight.cfg` file using **Add Custom Key**, by right-clicking on a parameter, then specifying the name and type. You can also add parameters by opening the `.cfg` file (located in the Data Workbench installation directory) using a text editor.

Following is an example of all of the parameters available for the `Insight.cfg` file that you can use as a model for adding parameter entries:

```
Licensing = serverInfo:  
    Proxy Address = string:  
    Proxy Port = int: 8080  
    Proxy User Name = string:  
    Proxy Password = string:  
Servers = vector: 1 items  
    0 = serverInfo:  
        Address = string: VS02  
        Name = string: Insight Server  
        Port = int: 443  
        Proxy Address = string:  
        Proxy Password = string:  
        Proxy Port = int: 8080  
        Proxy User Name = string:  
        SSL Client Certificate = string: Named User.pem  
        SSL Server Common Name = string: VS02  
        Use Address File = bool: false  
        Use SSL = bool: true  
Color Set = int: 0  
Default to Online = bool: false  
Fonts = vector: 0 items  
Gamma = float: 1.6  
Maximum Sample Size (MB) = double: 0  
Network Location = string:  
Unhide All = bool: false  
Unlock = bool: false  
Update Software = bool: true  
User Folder = string: User\\  
Toolbar Icons = bool: false
```

The following table provides descriptions of the available `Insight.cfg` file parameters in alphabetical order.

Parameter	Description
Address	<p>The host name or numeric IP address of your Data Workbench server computer. Example: <code>vsServer.mycompany.com</code> or <code>192.168.1.90</code></p> <p>If Address is not specified, the client uses the common name specified in the SSL Server Common Name parameter when Use Address File is set to false.</p> <p> Note: If the Use Address File parameter is set to true, the text entered in the Address parameter can be removed after the first profile is opened on the client. Then the setting for Network Location Parameter determines which addresses from the Address file for the cluster are used for connecting to the master server.</p>
Color Set	<p>Specifies the background color of your client application. The options are as follows:</p> <p>0 = black 1 = white 2 = monochrome</p> <p>The default value is 0, black.</p>

Parameter	Description
Default Online Level	<p>Optional. Enables you to make your instance of Data Workbench work by default as streaming, offline or online each time it opens. The options are Streaming, Online, or Offline. The default configuration for Data Workbench is to work Offline.</p> <p> Note: Before deciding to work online by default, make sure to weigh the benefits and consequences outlined in Working offline and online on page 201.</p>
Fonts	<p>Optional. Vector listing the fonts that the client should use to render UTF8-based unicode special characters. The number of fonts in the list is unlimited.</p> <p>The first font should always be Lucida Sans Console. If this parameter is not included in the <code>Insight.cfg</code> file, Data Workbench uses only Lucida Sans console to display text.</p> <p>Data Workbench uses the first font in the list to render all characters until it encounters a character that it cannot render. It uses the second font in the list to render that character. If that font does not render the character, Data Workbench uses the third font in the list to render that character, and so on, until it reaches the end of the font list. If the correct font is not listed in the vector, Data Workbench displays the hexadecimal value for the character.</p> <p> Note: Do not change this parameter while Data Workbench is running.</p>
Gamma	Gamma setting for the Data Workbench display. The default value is 1.6. Adobe does not recommend changing this value.
Licensing	<p>Optional. You need to modify the parameters in the Licensing vector only if you are contacting Adobe's license server through a proxy server.</p> <p>Section identifier for parameters that enable your client to contact the Adobe License Server through a proxy server. Expand the Licensing node and complete the following parameters.</p>
Proxy Address	Optional. The address of the proxy server that client must use to access the Adobe License Server.
Proxy Port	Optional. The port of the proxy server.
Proxy User Name	Optional. The user name for the proxy server.
Proxy Password	Optional. The password associated with the Proxy User Name.
Maximum Sample Size (MB)	<p>Specifies the maximum size allowed in the data cache used by the client running on a single computer.</p> <p>While the Sample Bytes parameter in the <code>Server.cfg</code> file specifies the data cache size for all clients connecting to a Data Workbench server (250e6 bytes by default), the Maximum Sample Size parameter enables you to further limit the data cache size for a particular computer. This is useful when the client is installed on a computer with limited storage or computing power.</p> <p> Note: This parameter limits the size of a local data sample queried locally by the client program. In contrast, the <code>cache.db</code> file contains data for each profile the client connects to, plus the element names and their dimensions. These element names and dimensions in the <code>cache.db</code> files are required to run local queries. Consequently, there is no way to limit its size other than reducing the number of elements in the data set.</p>
Name	Optional. The name that the client uses to identify the server.
Network Location	<p>Optional. The network location that client uses to resolve the Data Workbench server common name to an IP address. The available network locations are defined in the address file on the server. For more information, see the <i>Server Products Installation and Administration Guide</i>.</p> <p>If you do not specify a network location, client resolves common names using the default network location, "Insight."</p>

Parameter	Description
Port	The port to which the client sends requests. This port number must match the port on which the Data Workbench server is listening for requests. This is usually 443 for secure connections.
Proxy Address	If a proxy server is required to connect to your Data Workbench server computers, you must complete at least this parameter and the Proxy Port parameter. You can optionally complete the Proxy User Name and Proxy User Password parameters. The address of the proxy server that the client must use to access the Data Workbench server.
Proxy Password	Optional. The password to the proxy server.
Proxy Port	The port of the proxy server. The default value is 8080.
Proxy User Name	Optional. The user name for the proxy server.
Search	In the <code>Insight.cfg</code> (or any <code>.cfg</code> file), you can search by key name, key type, or value to quickly locate an entry, to remove the need to scroll through expanded, large files for nested information. You can locate dimension names, server names, and so on. Type a search phrase into this field to locate the data. Depending on the success of a match, the color of the field changes. Matches are shown highlighted and non-matches are dimmed. If there are no matches, the background of the search field turns red. When you press Enter, the config tree expands every place where there is a match and collapses where there is not a match. You can also use regular expressions in the Search field. For example, you can use re: <code>*zip.*</code> for any entry containing the word “zip.” To clear a search, press Escape .
Servers	Vector heading for client to server connections.
SSL Client Certificate	Optional unless you have more than one certificate. The name of the file that contains the digital certificate for this copy of Data Workbench. This is the file that you downloaded in Downloading and Installing the Digital Certificate. Example: <code>Samantha Smith.pem</code> If you leave this parameter blank, the client uses whatever certificate is present.
SSL Server Common Name	The common name of the Data Workbench server (as assigned on its digital certificate).
Toolbar Icons	False turns off the icons in the workstation user interface and displays text in the toolbar.
Use Address File	Specifies whether any settings in the address file override the Address parameter setting. The options are true or false. If set to true, the settings in the address file, if present, override the Address parameter setting. The default value is true.
Use SSL	Specifies whether SSL is used for secure communication between the Data Workbench server and the client. The options are true or false. The default value is true.
Unhide All	Optional. Enables you to temporarily unhide all metrics, dimensions, and filters that have been hidden using any of the following functionality: <ul style="list-style-type: none">• [Exclusive] setting in <code>order.txt</code> files• Hide option in <code>order.txt</code> files• Show parameter in <code>.metric</code>, <code>.dim</code>, and <code>.filter</code> files.• Hidden parameter in the <code>Transformation.cfg</code> file. For more information about these methods, see Customize a menu using order.txt files on page 486.

Parameter	Description
	The options are true or false. The default setting is false. Change this parameter to true to temporarily unhide hidden metrics, dimensions, and filters.
Unlock	Optional. Specifies whether you are allowed to unlock locked workspaces. The options are true and false. True enables you to unlock any locked workspace, while false does not. The default value is false. For more information about locked workspaces, see Unlocking a workspace on page 222.
Update Software	Optional. Specifies whether to allow this client software to be updated by the Data Workbench server. The options are true or false. The default value is true.
User Folder	Optional. Specifies the name and location of the folder that contains the local copies of any workspaces, metrics, dimensions, or configuration files for each profile. The default setting is User\, which specifies the User folder within the Data Workbench installation directory. Changing this setting is useful when two users are sharing the same computer. To accommodate both users, you can specify a shared folder to which both users have access.

Query language syntax

Navigation title:

Query language syntax for creating and editing metric, dimension, and filter expressions.

Syntax for any expression

Navigation title:

Syntax rules for formulas.

- When in a formula, keywords are case sensitive and must be typed exactly as they appear.
- When in a formula, metric, dimension, and filter names that include spaces must use underscores between words. For example: Page_VIEWS
- For strings within expressions, you must escape certain single quotation marks, double quotation marks, and backslashes. For example:
 - "can't" is acceptable and does not need to be escaped, but 'can't' is unacceptable and must be escaped as 'can\'t'.
 - c:\windows must be escaped as c:\\windows.

Syntax for identifiers

Navigation title:

Metric, dimension, and filter expressions can use identifiers to refer to named metrics, dimensions, and filters.

These identifiers are case sensitive and must be typed exactly as they are defined.

A valid identifier can contain one or more of the following:

- Underscores (_). Underscores in an identifier represent spaces in the metric, dimension, or filter name. For example, the Session Referrer dimension would be referred to as Session_Referrer in an expression.
- Percent signs (%)
- Upper case letters (A-Z)
- Lower case letters (a-z)
- Dollar signs (\$)
- Numbers (0-9), except as the first character in an identifier.
- Non-ASCII characters

All other characters are illegal in an identifier.

These same rules apply to the names of metrics, dimensions, and filters when they are used outside of expressions, except that the names can contain spaces but not underscores. For example, you can define the Session Referrer dimension in the `Transformation.cfg` file as Session_Referrer, but not `Session_Referrer`.

Syntax for metric expressions

Navigation title:

Metrics can be edited using the **Metric Editor** and saved in the Metrics directory of a profile.

For more information, see [Work with derived metrics](#) on page 461. Metric expressions also can be used in worksheets. For more information, see [Worksheets](#) on page 427. The following syntax is used to define metric expressions.

Notes:

1. Underlined words should be entered in the expression text literally.
2. The form {TEXT}? represents optional text.
3. The form {TEXT}* represents text that may occur zero or more times.
4. The form {A | B | C |...} represents text that consists of exactly one of the given options, such as A or B or C....
5. The form [A,B) represents a range of numbers, from A up to but not including B.

Identifier	An identifier references a named metric. For the rules governing legal identifiers, see Syntax for identifiers on page 515. Example: Revenue = Total_Price
(Metric)	The result of (Metric) is the same as the result of Metric. Parentheses specify the order of operations in an expression. Example: Average = (A + B) / 2
A + B	Sum of the results of Metric A and Metric B. Example: Value = Revenue + Cost_Savings
A – B	Difference of the results of Metric A and Metric B. Example: Profit = Revenue – Cost
A * B	Product of the results of Metric A and Metric B. Example: Dollars = Cents * 0.01
A / B	Ratio of the results of Metric A and Metric B. Example: Revenue_per_Session = Revenue / Sessions
A ^ B	Result of Metric A raised to the power of the result of Metric B. Example: Area = Width^2
confidence(metric)	An estimate of the standard deviation of the metric. This is calculated using a sampling technique known as jackknifing. This metric is memory-intensive and should not be used in large tables. To use this syntax, you must have a jackknife dimension (named “jackknife”) with the appropriate properties. For more information, contact Adobe Consulting Services. Example: confidence(Average_Score)  Note: The confidence metric types, including confidence(metric) and confidence(metric,jackknife), are especially useful when using Adobe's controlled

	<p>experimentation functionality. If a metric jumped from 12% to 16% during a controlled experiment, you could use a confidence callout to calculate the odds that the jump was due to random variation. This can help you to avoid drawing the wrong conclusions from limited evidence, and, conversely, provide assurance that a questionable change is actually real.</p>
confidence(metric, jackknife)	<p>An estimate of the standard deviation of the metric. This is calculated using a sampling technique known as jackknifing. This syntax enables you to determine the confidence of a metric using a jackknife dimension named something other than “jackknife.”</p> <p>This metric is memory-intensive and should not be used in large tables.</p> <p>To use this syntax, you must have a jackknife dimension (named something other than “jackknife”) with the appropriate properties. For more information, contact Adobe Consulting Services.</p> <p>Example: <code>confidence(Average_Score,SubSamples)</code></p>
eval(CellReference)	<p>Treats the content of the cell you are referencing as a metric expression. This syntax can be used only in a worksheet visualization.</p> <p>Example: <code>eval(B1)</code></p>
log (B, X)	<p>The mathematical logarithm function: Metric X is the parameter, and Metric B is the base.</p> <p>Example: <code>dB = 20*log(Amplitude,10)</code></p>
Metric[Filter]	<p>“Metric where Filter”: A new metric filtered by the given filter.</p> <p>Example: <code>Jan_Sessions = Sessions[Month=“Jan”]</code></p>
Metric by Dimension	<p>A metric evaluated at the “level” of dimension. The result of <code>(M by X)[F]</code> (the result of metric “M by X” evaluated with Filter “F”) is the result of <code>M[F by X]</code> (the result of Metric “M” evaluated with Filter “F by X”).</p> <p>Example: <code>AB_Visitors =</code> <code>(Visitors by Session)[Page=“A” and Page=“B”] =</code> <code>Visitors[(Page=“A” and Page=“B”)] by Session =</code> The number of Visitors who visited page A and page B in the same Session.</p>
number	<p>A metric with a fixed value.</p> <p>Example: <code>Pi = 3.1415</code></p>
total(Metric)	<p>Ignores any dimension over which the metric is evaluated. The metric has the same value for every element of that dimension.</p> <p>Example: <code>Pct_of_Visitors = Visitors / total(Visitors)</code></p>
all(Metric)	<p>Ignores filters applying to the metric. The metric are unaffected by selections and other filters.</p> <p>Example: <code>Benchmark_Sessions = all(Sessions)</code></p>
total(all(Metric))	<p>Ignores all filters and dimensions. It has the same value throughout a given profile regardless of what filters or dimensions are applied.</p> <p>Example: <code>Dataset_Visitors = total(all(Visitors))</code></p>
sum(One,Countable_Dimension)	<p>A metric that gives the count of a countable dimension such as Visitor or Session.</p> <p>Example: <code>Visitors = sum(One,Visitor)</code></p>
sum(Numeric_Dimension, Countable_Dimension)	<p>A metric that gives the sum of a numeric dimension over a countable dimension. The ordinal values (as opposed to the formatted values) of the numeric dimension’s elements are used, so a scale factor often needs to be applied to the result.</p> <p>Example: <code>Value = sum(Session_Value, Session)*0.01</code></p>

min(A, B)	The lesser results of metric A and metric B.
max(A, B)	The greater of the results of metric A and metric B.
format(A, B)	A metric which is identical to metric A, except it uses the formatting function of metric B.

Syntax for dimension expressions

Navigation title:

Dimension expressions are never used alone, but can be used anywhere a dimension is called for in a metric or filter expression.

1. Underlined words should be entered in the expression text literally.
2. The form {TEXT}? represents optional text.
3. The form {TEXT}* represents text that may occur zero or more times.
4. The form {A | B | C |...} represents text that consists of exactly one of the given options, such as A or B or C....
5. The form [A,B) represents a range of numbers, from A up to but not including B.

Syntax for Dimension Expressions

Identifier	An identifier references a named dimension. For the rules governing legal identifiers, see Syntax for identifiers on page 515. Example: Sessions[Session_Number = "1"] is the number of Sessions that had a Session Number of "1." Session Number is a named dimension referenced by identifier.
(Dimension)	The result of (Dimension) is the same as the result of Dimension. Parentheses specify the order of operations in an expression. Example: Sessions[(Page) = "/home"] is the number of Sessions visiting the Page "/home".
Dim by Level	Defines a dimension having the same elements as the dimension Dim, but relating to other dimensions through the dimension level. Specifically, an element of the new dimension relates to the same elements of level as the same element of Dim, and relates to those elements of any other dimension that relate to any of those elements of level. Example: Sessions[(Page by Visitor)="/home"] is the number of Sessions of Visitors who visited the Page "/home".
shift(Dim,Level,Group,N)	Defines a dimension having the same elements as the dimension Dim. The eth element of the dimension level relates to the same element of the new dimension as the element of Dim related to the e+Nth element of Level, provided that the eth and e+Nth elements of level relate to the same element of the dimension group. Example: Page_Views[shift(Page, Page_View, Session, 1)="/home"] is the number of Page Views for which the next Page viewed in the same Session is "/home".
next(Dim,Level,Group,N)	Similar to shift(Dim,Level,Group,N), except that if there are empty values in the dimension, they are skipped.
segment(Level {,String->Filter}*)	Defines a dimension that classifies elements of Level based on a list of filters. The elements of the new dimension are the strings given as arguments. Each element of Level relates to the 1st element of the segment dimension whose filter admits the element of Level. This is similar to the segment visualization. Example: segment(Visitor, "One-Time Visitors" -> Visitor_Sessions = 1, "Very Loyal Visitors" -> Visitor_Sessions > 10, "Everyone Else" -> True) creates a dimension that classifies Visitors into three groups -- One-Time Visitors are those with only one Session, Very Loyal Visitors are those with more than ten Sessions, and all other Visitors have a value of "Everyone Else."

<code>bucket(Level, Metric, Count, Format {, Start {, Size}? }?)</code>	<p>Defines a dimension whose elements are ranges of numbers (of fixed size, e.g. [0-9], [10-19],...). Elements of Level relate to the element of the bucket dim whose range contains the value of Metric for that element of level. Format is the printf format string used to format the elements of Metric.</p> <p>Example: If Page_Duration_Minutes is a Page View-level dimension representing the number of minutes spent on each page, then <code>bucket(Session, sum(Page_Duration_Minutes, Page_View), 100, "%0.0f minutes", 0, 5)</code> is a Session-level dimension representing the number of minutes spent in each Session; its elements are 5 minute intervals {[0-5], [5-10],..., [495-500]}.</p> <p>Start is the starting value of the first interval (default: 0) and Size is the size of the interval (default: 1).</p>
<code>prefix(Level {,ElementName->(Prefix{,Prefix}*)})</code>	<p>Defines a dimension whose elements are the given ElementName strings and are associated with the corresponding sets of Prefix strings. Elements of Level relate to the element of the prefix dim, which is associated with the longest prefix matched by the name of the given element of level. Prefixes ending with the special character '\$' must be matched exactly.</p> <p>For example, <code>prefix(URI, "Products" -> ("products/"), "Services" -> ("services/", "/products/service/"), "Warranties" -> ("products/warranty.html\$", "/services/warranty.html\$"), "Everything Else" -> (""/"))</code> creates a dimension that classifies URIs into the four listed categories. The effect on various pages is as follows:</p> <ul style="list-style-type: none"> /products/warranty.html Goes into Warranty, since it matches the /products/warranty.html\$ prefix exactly. /products/cars/specialcar.html Goes into Products, since it matches the /products/ prefix and no longer prefix /products/service/something.html Goes into Services, since it matches the /products/service/ prefix which is longer than the /products/ prefix. /companyinfo/aboutus.html Goes into the "Everything Else" category, since the only prefix it matches is "/".
<code>latency(Level, Clip, Dim, Filter, MaxBefore, MaxAfter, FormatString)</code>	<p>See Create a latency dimension on page 501.</p>
<code>cartesian_product(Separator {,Dim}*)</code>	<p>Defines a dimension whose elements are all the combinations ("the cartesian product") of the elements of the dimensions given. The name of each element is made out of the concatenation of the corresponding elements in the input dimensions, separated by the given Separator string.</p> <p>For example, if the dimension D1 has elements {"a", "b"} and the dimension D2 has the elements {"x", "y"}, then cartesian product("-", D1, D2) has the elements {"a-x", "a-y", "b-x", "b-y"}.</p> <p>Note that internally, each of the input dimensions is treated as if the number of its elements is the next higher power of two. This results in the cartesian product having some dummy elements. When using the Data Workbench API, depending on the output format, these elements may be omitted, or they may be shown as "#nnn", where nnn is the ordinal of the element (and should be ignored by the client).</p> <p>For example, in the example above, if D2 had the three elements {"x", "y", "z"}, it would be treated as if it had four elements, and the cartesian product would have the elements {"a-x", "a-y", "a-z", "#3", "b-x", "b-y", "b-z", "#7"}.</p> <p>If no dimensions are given, the result is a dimension with one element, "#0", which is equivalent to the None dimension.</p>
<code>nearest_countable(Dim)</code>	<p>Refers to an already existing dimension: the nearest countable ancestor of Dim in the schema. For example, <code>nearest countable(URI)</code> is identical to <code>Page_View</code>.</p>
<code>normalized(Dim,Count)</code>	<p>Defines a normalized dimension from the denormal dimension Dim, with up to Count elements.</p>
<code>last_n(Dim, TimeMetric, FormatString, Count, Offset, TrimToData {, WeekStart}?)</code>	<p>Defines a dimension which has a subset of the elements of the dimension Dim, whose elements represent slices of time -- for example, days, weeks, or years.</p> <p>The subset is a range around a specified time, the value of the constant metric TimeMetric, which is interpreted as a time value in seconds since midnight UTC of January 1, 1970. The range has Count elements, the last of which is Offset elements after the given Dim's element whose name is the result of formatting the value of the metric with the given FormatString string. FormatString uses the same % escapes as the standard C library function strftime.</p>

	<p>If trimToData is true then any elements at the beginning of the resulting dimension, which would be before the beginning of Dim, are removed. When it is false, there will always be the exact number of elements specified by Count. Note that there may always be elements at the end of the resulting dimension that are not actually in Dim.</p> <p>The optional WeekStart, if specified, must be one of { "Sun", "Mon", "Tue", "Wed", "Thu", "Fri", "Sat" }. It modifies TimeMetric by moving it backwards to the most recent occurrence of that weekday.</p> <p>Example: If Week has the elements { "10/03/10", "10/10/10", ..., "12/12/10" } and the built-in As Of metric has the value 1292348109 (representing a time in the middle of December 14th, 2010), then last n(Week, As_Of, "%m/%d/%y", 4, 0, false, "Sun") defines the dimension with elements { "12/12/10", "12/19/10", "12/23/10", "12/30/10" }.</p> <p>Example 2: If the Week dimension only has elements { "12/19/10", "12/26/10", ..., "01/30/11" }, and the As Of metric is as above, then last n(Week, As_Of, "%m/%d/%y", 4, 0, true, "Sun") gives a dimension with elements { "12/19/10", "12/23/10", "12/30/10" }.</p>
days_of_previous_months(Dim, TimeMetric, FormatString, nMonths, includeThisMonth, TrimToDate)	<p>Defines a dimension which has a subset of the elements of Dim, whose elements represent days. The subset is a range around a specified time, the value of the constant metric TimeMetric, which is interpreted as a time value in seconds since midnight UTC of January 1, 1970. The range will include the elements corresponding to each day in the nMonths months preceding the specified time. If includeThisMonth is true, the range also includes each day of the month that contains the specified time.</p> <p>FormatString specifies the formatting of the elements of Dim, using "%" escapes as in the standard C library function strftime.</p> <p>If trimToDate is true then any elements at the beginning of the resulting dimension, which would be before the beginning of Dim, are removed. When it is false, there will always be the exact number of elements specified by Count. Note that there may always be elements at the end of the resulting dimension that are not actually in Dim.</p> <p>Example: If Day has the elements { "01/01/10", "01/02/10", ..., "12/31/10" } and the built-in As Of metric has the value 1292348109 (representing a time in the middle of December 14th, 2010), then days_of_previous_months(Day, As_Of, "%m/%d/%y", 2, false, false) will have elements { "10/01/10", "10/02/10", ..., "11/30/10" }.</p>
days_of_current_month(Dim, TimeMetric, FormatString, allMonth, trimToDate)	<p>Similar to days of previous months, except the elements correspond only to days of the same month as the time specified by the TimeMetric. If allMonth is true, there will be an element for each day of the appropriate month; otherwise, only days from the first of the appropriate month through the day containing the specified time will be part of the dimension.</p>
days_of_future_months(Dim, TimeMetric, FormatString, nMonths, includeThisMonth, TrimToDate)	<p>Similar to days of previous months, except that the elements correspond to the days of months after, rather than before, the month containing the time specified by the TimeMetric.</p>
hours_of_day(Dim, Metric, TimeFormatString, nDaysForward, TrimData)	<p>Defines a dimension which has a subset of the elements of Dim, whose elements represent hours. The subset is a range around a specified time, the value of the constant metric TimeMetric, which is interpreted as a time value in seconds since midnight UTC of January 1, 1970. The range includes the elements corresponding to each hour of the day nDaysForward after the day containing the time specified by the TimeMetric.</p> <p>FormatString specifies the formatting of the elements of Dim, using "%" escapes as in the standard C library function strftime. The format string should always output a string representing midnight at the beginning of the day of the time passed in.</p> <p>If trimToDate is true then any elements at the beginning of the resulting dimension, which would be before the beginning of Dim, are removed. When it is false, there will always be the exact number of elements specified by Count. Note that there may always be elements at the end of the resulting dimension that are not actually in Dim.</p> <p>Example: If Hour has the elements { "01/01/10 00:00", "01/01/10 01:00", ..., "12/31/10 23:00" }, and the built-in As Of metric has the value 1292348109 (representing a time in the middle of December 14th, 2010), then hours_of_day(Hour, As_Of, "%x 00:00", 0, false) has elements { "12/12/10 00:00", "12/12/10 01:00", ..., "12/12/10 23:00" }.</p>

Syntax for filter expressions

Navigation title:

A filter is an expression that defines a subset of the data in a dataset.

A filter either admits or rejects each element of each dimension according to the relationships between dimensions.

Filters can be edited using the **Filter Editor**. See [Filter editors](#) on page 348.

In the following table, each syntax description includes an example of a metric expression using that filter. For example, Sessions[True] is a metric defined using the “True” filter. The Sessions[True] metric is the same as the Sessions metric because the True filter admits every element of the Session dimension.

Syntax for filter expressions

True	Constant filter. Admits every element of every dimension Example: Sessions[True] is the same as Sessions.
False	Constant filter. Rejects every element of every dimension. Example: Sessions[False] is always zero.
not Filter	Admits elements that Filter rejects. Example: Sessions[not Page=”A”] is the number of Sessions that did not visit page A.
FilterA and FilterB	Admits elements that FilterA and FilterB admit. Example: Sessions[Page=”A” and Page=”B”] is the number of Sessions that visited both page A and page B.
FilterA or FilterB	Admits elements that FilterA or FilterB admit. Example: Sessions[Page=”A” or Page=”B”] is the number of Sessions that visited page A, page B, or both.
Filter by Dim	Admits the set of elements of the dimension Dim that are admitted by Filter. Example: Sessions[Page=”/home” by Visitor] is the number of Sessions belonging to a Visitor that saw the Page “/home”.
Identifier	Reference filters defined otherwise in the profile. Example: Sessions[Broken_Session_Filter] is the number of Sessions admitted by the Broken Session Filter.
Dim = “Value”	Admits the given element of the dimension Dim. Example: Sessions[Page=”A”] is the number of Sessions that visited Page A.
Dim <> “Value”	Admits every other element of the dimension Dim.
Dim != “Value”	Example: Sessions[Page<>”A”] is the number of Sessions that visited any page other than A.
Dim = #Ordinal	Admits the element of the dimension Dim with the given ordinal value. Example: Sessions[Month=#0] is the number of Sessions in the first month of the dataset.
Dim <> #Ordinal	Admits every other element of the dimension Dim.
Dim != #Ordinal	Example: Sessions[Session_Value <> #0] is the number of Sessions having a nonzero Session Value.
Dim matches “Expr”	Admits the elements of the dimension Dim matching the given regular expression. Dim must not be a denormal or countable dimension.

	Example: Sessions[URI matches “.*/product/*”] is the number of Sessions that visited any page in a product directory.
Dim notmatches “Expr”	Admits the elements of the dimension Dim not matching the given regular expression. Dim must not be a denormal or countable dimension. Example: Sessions[URI notmatches “.*\.jsp”] is the number of Sessions that visited any page which was not a JSP page.
Dim < “Value”	Admits the elements of the dimension Dim with ordinal values less than the ordinal value of the element “Value.” If “Value” is not an element of dimension, then it is assumed to be larger than any current element of the dimension. Example: Sessions[Month < “Jul ‘04”] is the number of Sessions that took place before July 2004.
Dim > “Value”	Admits the elements of the dimension Dim with ordinal values greater than the ordinal value of the element “Value.” If “Value” is not an element of dimension, then it is assumed to be larger than any current element of the dimension. Example: Sessions[Month > “Jul ‘04”] is the number of Sessions that took place after July 2004.
Dim <= “Value”	Admits the elements of the dimension Dim with ordinal values less than or equal to the ordinal value of the element “Value.” If “Value” is not an element of dimension, then it is assumed to be larger than any current element of the dimension. Example: Sessions[Session_Number <= “2”] is the number of Sessions that were the first or second session for a visitor.
Dim >= “Value”	Admits the elements of the dimension Dim with ordinal values greater than or equal to the ordinal value of the element “Value.” If “Value” is not an element of dimension, then it is assumed to be larger than any current element of the dimension. Example: Sessions[Session_Number >= “5”] is the number of sessions that were the fifth or greater session for a visitor.
any Dim	Admits all elements of the dimension Dim. Example: Sessions[any Page_View] is the number of sessions with at least one page view.
no Dim	Admits elements that any Dim rejects. Example: Sessions[no Page_View] is the number of sessions without a page view.
(FILTER)	The same as FILTER; used to group a part of a filter expression.

Syntax for math expressions

Data workbench allows you to employ additional mathematical calculations.

Math functions can be used evaluate and display in math transformation, metric editor visualizations, and worksheet data.

Calculation	Function	Input
absolute	abs()	Math transformation, Metric editor and Worksheet cell
arccosine	acos()	Math transformation, Metric editor and Worksheet cell
arcsine	asin()	Math transformation, Metric editor and Worksheet cell
arctangent	atan()	Math transformation, Metric editor and Worksheet cell
cosine	cos()	Math transformation, Metric editor and Worksheet cell

Calculation	Function	Input
exponential	exp()	Math transformation, Metric editor and Worksheet cell
logarithm	log()	Math transformation, Metric editor and Worksheet cell
maximum	max()	Math transformation and Statistics callout
minimum	min()	Math transformation and Statistics callout
ln (natural logarithm)	ln()	Math transformation, Metric editor and Worksheet cell
sine	sin()	Math transformation, Metric editor and Worksheet cell
tangent	tan()	Math transformation, Metric editor and Worksheet cell

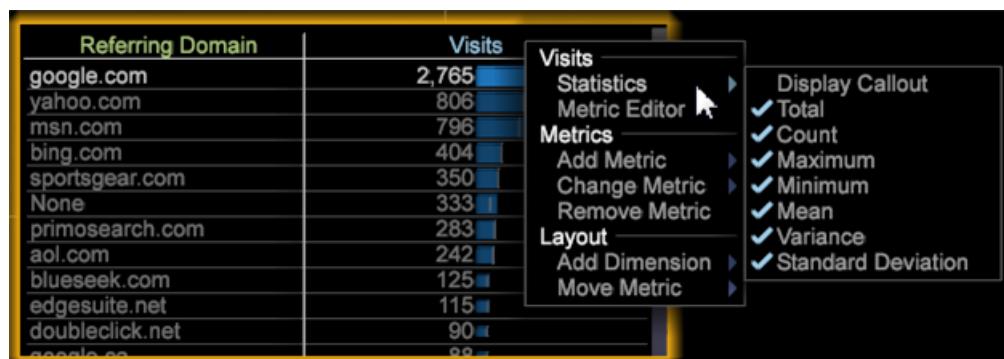
Statistical Callouts

Statistical callouts measure meaningful relationships to identify hidden opportunities and variables of interest for more advanced data mining capabilities in audience clustering and visitor response scoring.

Statistical callouts expand the algorithms so that more types of data can be correlated, such as binomial variables (yes/no, 0/1, or purchaser/non-purchaser) correlated with countable metrics (visits, orders, or downloads).

To add statistical callouts:

1. In a table, right-click the metric header.
2. Select **Statistics** and then select or clear the checkmarks for each required setting. All in the Display Callout are selected as the default setting.



The callout can return statistical values factored into the dataset columns.

Calculation	Description
Count	Returns the number of rows in a dataset.
Maximum	Identifies the maximum Metric value across all elements of the dimension.
Minimum	Identifies the minimum Metric value across all elements of the dimension.

Calculation	Description
Mean	The mean is the arithmetic average of the Metric values of elements in the Dimension, calculated by the total sum divided by the count (sum/count).
Standard Deviation	The standard deviation shows how much variation exists from the expected mean. A lower standard deviation shows the data points close to the mean. A higher standard deviation shows that the data points are spread across a large range of values.
Total	Returns the total sum of the Metric values.
Variance	A measure of the variance of the Metric values from the Metric mean for that dimension. It is equal to the square of the standard deviation.

Regular expressions

Data workbench utilizes regular expressions (regex) for search and sort operations.

Within the **Search** field you can perform a search following the "re:" statement using common expressions, for example:

`re: * .s`

metacharacter	description
. (dot)	Matches a single character, for example: <code>re: x . z</code> matches "xyz" or "xxz".
* (star)	Matches one or more characters, for example: <code>re: Z *</code> matches "ZZZ".
? (wildcard)	Matches 0 or 1 of previous expression to force minimal matching, for example: <code>re: xY? z</code> matches "xy" and "xyz".

Additional common regular expressions can also be used to create more complex search strings. Regular expressions are used across all Data Workbench search fields including the query entity panels.

See in-depth information at [regular expressions](#).

Dataset Configuration

Overview of the dataset construction process and the **dataset configuration** files that control the processing phases.

Understanding Dataset Construction

An Adobe dataset contains the data that has been loaded and processed by the data workbench server.

The steps involved in the loading and processing of the data by the data workbench server (InsightServer64.exe) make up the dataset construction process.



Note: A data workbench server that processes and serves data from an Adobe dataset is called a data processing unit or DPU. It is sometimes referred to as a processing server or a query server. Data workbench and **Report** clients interact with DPUs directly.

During dataset construction, the data workbench server reads source data from log sources, applies transformations to specific fields of data, and defines extended dimensions to be created from the transformed fields. The construction process occurs in two phases: *Log Processing* and *Transformation*. After the dataset is constructed,

you can use the dataset's extended dimensions to create derived metrics and dimensions for your specific analysis purposes.

Dataset construction is like a manufacturing process. You select the data (the raw materials) to be used to build the dataset, and you define the data transformations (the process steps) that manipulate the information available in the data to create extended dimensions (the manufactured products).

Log Processing

During the log processing phase of dataset construction, the sources of data for the dataset are specified, and the desired data fields are extracted and manipulated into more useful forms.

The logs are filtered, and the fields of data that are to be passed to the transformation phase are identified. At the end of the log processing phase, the data is grouped by tracking ID (that is, all log entries with the same tracking ID are grouped together) and ordered in time. During the log processing phase, you cannot access the processed data to use for analysis.

Specifying Log Sources

Log sources are files that contain the data to be used to build a dataset. The data available in the log sources is called event data because each data record represents a transaction record or a single instance of an event. In addition, each record, or log entry, contains a value referred to as a tracking ID.



Note: When selecting log sources, make sure that each log entry contains a tracking ID for the entity that is to represent the highest level at which your data is to be grouped. For example, if you are working with data collected from website traffic, you are likely to choose visitor to be this entity. Each visitor has a unique tracking ID, and all of the data about a particular site visitor can be grouped together. For assistance, contact Adobe.

A log sources event data is collected in real-time by **Sensors** or extracted from archived data sources by Insight Server. Event data collected by Sensors from HTTP and application servers is transmitted to Insight Servers, which convert the data into highly compressed log (.vs1) files. Event data that resides in a flat file, XML file, or an ODBC data source is read by Insight Server, which provides decoders that you define to extract a common set of log fields from these different formats.

Defining Transformations

A transformation is a set of instructions that you can define to extract or manipulate information in the event data. Each transformation that you define is applied to each event data record (log entry) to update existing log fields or produce new fields. The results of transformations are used along with log entry conditions to evaluate which log entries are filtered out of the dataset during log processing.

Not all types of transformations can be used during the log processing phase of the dataset construction process.

Filtering Logs

The dataset contains several parameters used to filter the data flowing out of the transformations. Filtering is used to specify which log entries are used in subsequent processing steps. For example, filters can be defined by, time range, the status of the server's response, or IP address and user-agent information. The **Log Entry Condition** is a customizable filtering test. The test looks for certain conditions in the fields of each log entry to determine whether that entry should proceed further in the dataset construction process. If a log entry does not meet the condition, the entry is removed from the construction process.

Identifying Fields for Transformation

If a field of data is to be passed from the log processing phase to the transformation phase for further processing, you must identify it during log processing. This requirement applies regardless of whether the field is available from the log sources or created from data transformations applied to the data during log processing.

Transformation

During the transformation phase of dataset construction, processing occurs on the grouped and ordered data that is output from log processing.

During the transformation phase of dataset construction, processing occurs on the grouped and ordered data that is output from log processing. Additional data transformations are performed and extended data dimensions are created for use in your analyses. During the transformation phase, you can access a statistical sample of the data that gets larger as the transformation phase nears completion.

Defining Transformations

You can define transformations to be used during the transformation phase of the dataset construction process to facilitate the creation of the extended dimensions. Each transformation is applied to each event data record (log entry) passed from log processing.

Filtering Logs

The **Log Entry Condition** can be applied during transformation to look for specific conditions in the fields of each log entry coming from log processing. If a log entry does not meet the condition, the entry is removed from the construction process.

Defining Extended Dimensions

Extended dimensions are the final products of the dataset construction process. They represent relationships between the log fields in the data. You use them to create visualizations, build extended metrics, or perform analysis to understand the operations and issues specific to your business.

Understanding Dataset Configuration

Dataset configuration refers to the process of editing the configuration files whose parameters provide the rules for dataset construction.

The constructed dataset physically resides in the `temp.db` file stored on the data workbench server computer, but the configuration files for the dataset reside within a directory for a profile. A profile contains a set of configuration files that construct a dataset (including its extended dimensions) for a specific analysis purpose. In addition, a profile contains the definitions of entities such as metrics, derived dimensions, workspaces, reports, and visualizations that enable analysts to interact with the dataset and obtain information from it.

The profile whose dataset configuration files you are editing is referred to as your dataset profile. A dataset profile references multiple inherited profiles, which can be any profiles that you create and maintain so that you can configure your Adobe application to best fit your analysis needs. A dataset profile also may reference internal profiles that are provided with your Adobe application to form the basis for all of the functionality available in your application.

For more information about the different types of profiles that are available with Adobe applications, see the *Data Workbench User Guide*.

Required Configuration Files

The dataset profile's required configuration files control the processing phases (log processing and transformation) through which incoming data flows to keep the dataset up to date and specify the inherited profiles that are to be included in the dataset profile.

A dataset profile for any Adobe application must contain the following configuration files on the Insight Server machine:

- **Profile.cfg:** Lists the inherited profiles and processing servers for the profile. Processing servers are the Insight Server DPUs that process the data for the profile. If you have installed an Insight Server cluster, you can specify multiple Insight Server computers to run a single profile.

For instructions to add inherited profiles to a dataset profile's `Profile.cfg` file, see the *Server Products Installation and Administration Guide*. For information about installing an Insight Server cluster or configuring a dataset profile to run on an Insight Server cluster, see the *Server Products Installation and Administration Guide*.

- **Dataset\Log Processing.cfg:** Controls the log processing phase of the dataset construction process. See [Log Processing](#) on page 525. For more information about the `Log Processing.cfg` file, see [Log Processing Configuration File](#) on page 528.
- **Dataset\Transformation.cfg:** Controls the transformation phase of the dataset construction process. See [Transformation](#) on page 526. The `Transformation.cfg` file typically configures the dataset for profile-specific analysis. For more information about the `Transformation.cfg` file, see [Transformation Configuration File](#) on page 551.
- **Dataset Include Files:** A **dataset include** file contains a subset of the parameters contained in the `Log Processing.cfg` or `Transformation.cfg` file for the dataset profile but is stored and managed within an inherited profile. **Dataset include** files supplement the main dataset configuration files. For more information, see [Dataset Include Files](#) on page 557.

The dataset profile provided to you during the implementation of your Adobe application contains a set of dataset configuration files that you can open, edit, and save using the **Profile Manager**.

For information about the **Profile Manager**, see the *Insight User Guide*.

Additional Configuration Files

Additional dataset configuration files reside in the profile's **Dataset directory** on the data workbench server computer.

Although not required for all datasets, these files enable you to control other aspects of the dataset construction process:

- **Log Processing Mode.cfg:** The `Log Processing Mode.cfg` file lets you pause processing of data into a dataset, specify offline sources, or specify the frequency at which the data workbench server saves its state files. See [Additional Configuration Files](#) on page 570.
- **Server.cfg:** The `Server.cfg` file specifies the default data cache size (in bytes) for data workbench machines that connect to the data workbench server. See [Additional Configuration Files](#) on page 570.
- **Transform.cfg and Transform Mode.cfg:** These files are available only if you have licensed the data transformation functionality to use with your Adobe application. The `Transform.cfg` file contains the parameters that define the log sources and data transformations for transformation functionality. The transformations that you define manipulate the source data and output it into a format that you specify. The `Insight Transform Mode.cfg` file enables you to pause processing of data into a dataset, specify offline sources, or specify the frequency at which the Insight Server running transformation functionality saves its state files. See [Transform Functionality](#) on page 617.

Next Steps

Information about specific dataset configuration tasks.

For information about specific dataset configuration tasks, use the table below to locate and read about the tasks of interest of you:

If you would like to...	See...
Define log sources	Log Sources on page 532
Determine which log entries enter the dataset during log processing	Data Filters on page 546 Log Entry Condition on page 546
Enable the splitting of tracking IDs with large amounts of event data	Key Splitting on page 547
Configure an Insight Server to run as a file server unit	Configuring a Data Workbench Server File Server Unit on page 540
Configure an Insight Server to run as a centralized normalization server	Configuring a Data Workbench Server File Server Unit on page 540
Set the time zone to be used for creating time dimensions and making time conversions	Time Zones on page 554
Make minor changes to the dataset configuration files included with the internal profiles provided by Adobe	Editing Existing Dataset Include Files on page 557
Specify new fields of data to be passed from log processing to transformation	Creating New Dataset Include Files on page 558 Log Processing Dataset Include Files on page 559
Define transformations	Data Transformations on page 573 Creating New Dataset Include Files on page 558 Transformation Dataset Include Files on page 567
Create extended dimensions	Extended Dimensions on page 606 Creating New Dataset Include Files on page 558 Transformation Dataset Include Files on page 567
Define parameters to use throughout log processing or transformation	Defining Parameters in Dataset Include Files on page 568
Learn about the Insight interfaces that enable you to monitor or manage your dataset	Working With Dataset Configuration Interfaces on page 628
Hide certain extended dimensions so they do not show on the dimension menu in Insight	Hiding Dataset Components on page 632
Override certain dataset configuration files in a profile that you cannot or do not want to modify	Hiding Dataset Components on page 632
Reprocess your dataset	Reprocessing and Retransformation on page 627

Log Processing Configuration File

Information about the `Log_Processing.cfg` file and how its parameters affect the log processing phase of the dataset construction process.

About the Log Processing Configuration File

The `Log Processing.cfg` file controls the log processing phase of dataset construction, during which unordered data is read from the data sources (referred to as log sources), and filters and transformations are applied to the data.

You must edit the `Log Processing.cfg` file to perform any of the following dataset configuration tasks:

- Specifying log sources. See [Log Sources](#) on page 532.
- Determining which log entries enter the dataset during log processing. See [Data Filters](#) on page 546 and [Log Entry Condition](#) on page 546.
- Enabling the splitting of tracking IDs with large amounts of event data. See [Key Splitting](#) on page 547.
- Configuring a data workbench server to run as a file server unit. See [Configuring a Data Workbench Server File Server Unit](#) on page 540.
- Configuring a data workbench server to run as a centralized normalization server. See [Configuring a Data Workbench Server File Server Unit](#) on page 540.



Note: **Log Processing Dataset Include** files can contain additional instructions for the log processing phase of dataset construction. These files exist within the `Dataset\Log Processing` directory for any inherited profile. They typically define application-specific parameters (such as web-specific configuration parameters for the Site application). For information about **Log Processing Dataset Include** files, see [Dataset Include Files](#) on page 557. For information about web-specific configuration parameters for Site, see [Configuration Settings for Web Data](#) on page 649.

Editing the Log Processing Configuration File

Steps to edit the `Log Processing.cfg` file for a dataset profile.

1. While working in your dataset profile, open the **Profile Manager** and click **Dataset** to show its contents. For information about opening and working with the **Profile Manager**, see the *Data Workbench User Guide*.



Note: A Log Processing subdirectory may exist within the `Dataset` directory. This subdirectory contains the **Log Processing Dataset Include** files that have been created for one or more inherited profiles. See [Dataset Include Files](#) on page 557.

2. Right-click the check mark next to `Log Processing.cfg` and click **Make Local**. A check mark for this file appears in the **User** column.
3. Right-click the newly created check mark and click **Open > in Workstation**. The `Log Processing.cfg` window appears.

You also can open the `Log Processing.cfg` file from a **Transformation Dependency Map**. For information about transformation dependency maps, see [Dataset Configuration Tools](#) on page 628.

4. Edit the parameters in the configuration file using the following table as a guide.

When editing the `Log Processing.cfg` file within a data workbench window, you can use shortcut keys for basic editing features, including cut (`Ctrl+x`), copy (`Ctrl+c`), paste (`Ctrl+v`), undo (`Ctrl+z`), redo (`Ctrl+Shift+z`), select section (click+drag), and select all (`Ctrl+a`). You can also use the shortcuts to copy and paste text from one configuration file (`.cfg`) to another.



Note: A **Log Processing Dataset Include** file for an inherited profile contains a subset of the parameters described in the following table as well as some additional parameters. See [Dataset Include Files](#) on page 557.

`Log Processing.cfg`

Parameter	Description
Log Sources	The sources of data. See Log Sources on page 532.
End Time	<p>Optional. Filter data to include log entries with timestamps up to but not including this time. Adobe recommends using one of the following formats for the time:</p> <ul style="list-style-type: none"> • January 1 2013 HH:MM:SS EDT • Jan 1 2013 HH:MM:SS GMT <p>For example, specifying July 29 2013 00:00:00 EDT as the End Time includes data through July 28, 2013, at 11:59:59 PM EDT. See Data Filters on page 546.</p> <p>You must specify a time zone. The time zone does not default to GMT if not specified. For a list of time zone abbreviations supported by the data workbench server, see Time Zone Codes on page 658.</p> <p> Note: The Use Start/End Times parameter for Sensor, log file, and XML sources is related to this parameter. See the sections of Log Sources on page 532 that discuss these source types.</p>
Fields	Optional. Adobe recommends defining Fields in one or more Log Processing Dataset Include files. See Log Processing Dataset Include Files on page 559.
Group Maximum Key Bytes	<p>Maximum amount of event data that the Server can process for a single tracking ID. Data exceeding this limit is filtered from the dataset construction process. This value must be set to 2e6 when key splitting is active and 1e6 when key splitting is not active. See Key Splitting on page 547.</p> <p> Note: Do not change this value without consulting Adobe.</p>
Hash Threshold	<p>Optional. A sampling factor for random sub-sampling of rows. If set to a number n, then only one out of each n tracking IDs enters the dataset, reducing the total number of rows in the dataset by a factor of n.</p> <p>To create a dataset that requires 100 percent accuracy (that is, to include all rows), you would set Hash Threshold to 1.</p> <p>Values:</p> <p>Hash Threshold = 1 (100 percent of data including all rows.)</p> <p>Hash Threshold = 2 (1/2 of data and includes half the rows.)</p> <p>Hash Threshold = 3 (1/3 of data and includes one of three rows, but rounds to 34% in Query Completion)</p> <p>Hash Threshold = 4 (1/4th of data and includes one out of four rows.)</p> <p> Note: Using a Hash Threshold = 8 provides 1/8th of the data, which is 12.5%. However the Query Completion value in the rounds to 13% for this value. Additional examples include a Hash Threshold = 6 that results in 17% query resolution. A Hash Threshold = 13 results in 8% query resolution.</p> <p>If Hash Threshold is specified in both the Log Processing.cfg and Transformation.cfg files, it is not applied in sequence; the maximum value set in either configuration file applies. See Data Filters on page 546.</p>
Log Entry Condition	Optional. Defines the rules by which log entries are considered for inclusion in the dataset. See Log Entry Condition on page 546.
Reprocess	Optional. Any character or combination of characters can be entered here. Changing this parameter and saving the file to the data workbench Server machine initiates data reprocessing. See Reprocessing and Retransformation on page 627.
Split Key Bucket Space	Parameter involved in key splitting. Its value should be 6e6 when key splitting is active. See .

Parameter	Description
	 Note: Do not change this value without consulting Adobe.
Split Key Bytes	Parameter involved in key splitting. Its value should be 1e6 when key splitting is active and 0 when key splitting is not active. See Key Splitting on page 547.  Note: Do not change this value without consulting Adobe.
Split Key Space Ratio	Parameter involved in key splitting. Its value should be 10 when key splitting is active. See Key Splitting on page 547.  Note: Do not change this value without consulting Adobe.
Stages	Optional. The names of the processing stages that can be used in Log Processing Dataset Include files. Processing stages provide a way to order the transformations that are defined in Log Processing Dataset Include files. This parameter is very helpful if you have defined one or more transformations within multiple Log Processing Dataset Include files and you want specific transformations to be performed at specific points during log processing. The order in which you list the stages here determines the order in which the transformations in the Log Processing Dataset Include files are executed during log processing. Preprocessing and Postprocessing are built-in stages. Preprocessing is always the first stage, and Postprocessing is always the last stage. By default, there is one named stage called Default. To add a new processing stage <ul style="list-style-type: none"> • In the <code>Log_Processing.cfg</code> window, right-click Stages and click Add New > Stage. • Enter a name for the new stage. To delete an existing processing stage <ul style="list-style-type: none"> • Right-click the number corresponding to the stage that you want to delete and click Remove<#stage_number>.  Note: When you specify a Stage in a Log Processing Dataset Include files, the name of the stage must match exactly the name that you enter here. See Dataset Include Files on page 557.
Start Time	Optional. Filter data to include log entries with timestamps at or after this time. Adobe recommends using one of the following formats for the time: <ul style="list-style-type: none"> • January 1 2013 HH:MM:SS EDT • Jan 1 2013 HH:MM:SS GMT For example, specifying "July 29 2013 00:00:00 EDT" as the Start Time includes data starting from July 29, 2013, at 12:00:00 AM EDT. See Data Filters on page 546. You must specify a time zone. The time zone does not default to GMT if not specified. For a list of time zone abbreviations supported by the data workbench server, see Time Zone Codes on page 658.  Note: The Use Start/End Times parameter for Sensor , log file, and XML sources is related to this parameter. See the sections of Log Sources on page 532 that discuss these source types.
Time Zone	Optional. Time zone of the the data workbench server that is used for time conversions (such as the conversion represented by the x-local-timestring field) during log processing.  Note: You must specify the Time Zone if you want to access the converted time field during the log processing phase of dataset construction. Otherwise, the data workbench server records an error in the event logs. See Time Zones on page 554.

Parameter	Description
Transformations	Optional. Adobe recommends defining transformations for log processing in one or more Log Processing Dataset Include files. See Log Processing Dataset Include Files on page 559.

5. Right-click (**modified**) at the top of the window and click **Save**.
6. In the **Profile Manager**, right-click the check mark for `Log Processing.cfg` in the **User** column, then click **Save to > <dataset profile name>** to make the locally made changes take effect. Reprocessing of the data begins after synchronization of the dataset profile.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

For more information about reprocessing your data, see [Reprocessing and Retransformation](#) on page 627.

Log Sources

Log sources are files that contain the data to be used to build a dataset.

The data available in the log sources is called event data because each data record represents a transaction record or a single instance of an event. The data workbench server can process log sources that are derived from data collected by **Sensors** or extracted from other data sources.

- **Data Collected by Sensors:** Data collected by **Sensors** from HTTP and application servers is transmitted to data workbench servers, which convert the data into highly compressed log (.vsl) files. See [Sensor Files](#) on page 533.
- **Data Extracted by Insight Server:** The data workbench server reads event data contained in flat files, XML files, or ODBC-compliant databases, and uses its decoders to extract the desired elements of the data. Such event data does not have to be memory-resident, but the records that contain the data must include a tracking ID. See [Log Files](#) on page 534, [XML Log Sources](#) on page 536, and [ODBC Data Sources](#) on page 547.

To add a log source

1. Open `Log Processing.cfg` in data workbench.
2. Right-click **Log Sources**, then click **Add New**.
3. Select one of the following:
 - **Sensor**
 - **Log File**
 - **XML Log Source**
 - **ODBC Data Source**
4. The specific parameters used to define a dataset vary based on the type of log source to be used in the dataset's configuration process. Specify the parameters as indicated in the section corresponding to the appropriate log source:
 - [Sensor Files](#) on page 533
 - [Log Files](#) on page 534
 - [XML Log Sources](#) on page 536
 - [ODBC Data Sources](#) on page 547
5. After you have defined your log source (and made changes to other parameters) in the `Log Processing.cfg` file, save the file locally and save it to your dataset profile on the data workbench server.



Note: A data workbench server **File Server Unit** can receive and store **Sensor** files, log files, and XML files and serve them to the data workbench server's **Data Processing Units** that construct the dataset. See [Configuring a Data Workbench Server File Server Unit](#) on page 540.

You can open the configuration of any log source from a **Transformation Dependency Map**. For information about **Transformation Dependency Map**, see [Dataset Configuration Tools](#) on page 628.

Sensor Files

Information about the requirements and parameters of **Sensor** files.

Requirements

Event data collected by **Sensors** from HTTP and application servers is transmitted to data workbench servers, which convert the data into highly compressed log (.vsl) files. The .vsl file format is managed by the data workbench server, and each file has a name of the format:

YYYYMMDD-SENSORID.VSL

where YYYYMMDD is the date of the file, and **SENSORID** is the name (assigned by your organization) that indicates which **Sensor** collected and transmitted the data to the data workbench server.

Parameters

For **Sensor** files, the following parameters are available:

Log Processing.cfg: Sensor Files

Parameter	Description
Log Paths	<p>The directories where the .vsl files are stored. The default location is the Logs directory. A relative path refers to the installation directory of the data workbench server.</p> <p>You can use wildcard characters to specify which .vsl files to process:</p> <ul style="list-style-type: none">• * matches any number of characters• ? matches a single character <p>For example, the log path <code>Logs*.*.vsl</code> matches any file in the Logs directory ending in .vsl. The log path <code>Logs*-SENSOR?.vsl</code> matches files in the Logs directory with any date (YYYYMMDD) and a single character after SENSOR, as in SENSOR1.</p> <p>If you want to search all subdirectories of the specified path, you must set the Recursive parameter to true.</p> <p> Note: If the files are to be read from a data workbench server's File Server Unit, then you must enter the appropriate URI(s) in the Log Paths parameter. For example, the URI <code>/Logs/*-*.*.vsl</code> matches any .vsl file in the Logs directory. See Configuring a Data Workbench Server File Server Unit on page 540.</p>
Log Server	Information (Address, Name, Port, and so on) necessary to connect to a file server. If there is an entry in the Log Server parameter, the Log Paths are interpreted as URIs. Otherwise, they are interpreted as local paths. See Configuring a Data Workbench Server File Server Unit on page 540.
Log Source ID	This parameter's value can be any string. If a value is specified, this parameter enables you to differentiate log entries from different log sources for source identification or targeted processing. The x-log-source-id field is populated with a value identifying the log source for each log entry. For example, if you want to identify log entries from a Sensor named VSensor01, you could type <code>from VSensor01</code> , and that string would be passed to the x-log-source-id field for every log entry from that source.
	For information about the x-log-source-id field, see Event Data Record Fields on page 646.
Recursive	True or false. If set to true, all subdirectories of each path specified in Log Paths are searched for files matching the specified file name or wildcard pattern. The default value is false.

Parameter	Description
Use Start/End Times	<p>True or false. If set to true and Start Time or End Time is specified, then all files for this log source must have file names starting with dates in ISO format (YYYYMMDD). It is assumed that each file contains data for one GMT day (for example, the time range starting at 0000 GMT on one day and ending at 0000 GMT the following day). If the log sources files contain data that do not correspond to a GMT day, then this parameter must be set to false to avoid incorrect results.</p> <p> Note: By default, .vs1 files containing data collected by Sensor automatically meet the naming and time range requirements described above. If you set this parameter to true, the data workbench server always processes data from files whose names include ISO dates that fall between the specified Start Time and End Time. If you set this parameter to false, the data workbench server reads all of the .vs1 files during log processing to determine which files contain data within the Start Time and End Time range.</p> <p>For information about the Start Time and End Time parameters, see Data Filters on page 546.</p>



Note: Do not use the configuration parameters for **Sensor** data sources to determine which log entries within a log file should be included in a dataset. Instead, set up the data source to point to all of the log files within a directory. Then use the Start Time and End Time parameters of `Log Processing.cfg` to determine which log entries should be used in constructing the dataset. See [Data Filters](#) on page 546.

Log Files

Information about integrating event data from flat files that are not .vs1 files.

The file containing the event data must meet the following requirements:

- Each event data record in the file must be represented by one line.
- The fields within a record must be separated, whether empty or not, by an ASCII delimiter. The data workbench server does not require you to use a specific delimiter. You may use any character that is not a line-ending character and does not appear anywhere within the event data itself.
- Each record in the file must contain:
 - A tracking ID
 - A time stamp
- To specify start and end times for data processing, each file name must be of the form:
 - YYYYMMDD-SOURCE.log

where `YYYYMMDD` is the Greenwich Mean Time (GMT) day of all of the data in the file, and `SOURCE` is a variable identifying the source of the data contained in the file.



Note: Please contact Adobe Consulting Services for a review of the log files that you plan to incorporate into the dataset.

Parameters

For log files log sources, the parameters in the following table are available.



Note: The processing of log file log sources requires additional parameters that are defined in a **Log Processing Dataset Include** file, which contains a subset of the parameters included in a `Log Processing.cfg` file as well as special parameters for defining decoders for extracting data from the log file. For information about defining decoders for log file log sources, see [Text File Decoder Groups](#) on page 561.

Log Processing.cfg: Log Files

Parameter	Description
Name	The identifier for the log file source.
Log Paths	<p>The directories where the log files are stored. The default location is the Logs directory. A relative path refers to the installation directory of the data workbench server.</p> <p>You can use wildcard characters to specify which log files to process:</p> <ul style="list-style-type: none"> • * matches any number of characters. • ? matches a single character. <p>For example, the log path <code>Logs* .log</code> matches any file in the Logs directory ending in <code>.log</code>.</p> <p>If you want to search all subdirectories of the specified path, then you must set the Recursive parameter to true.</p> <p>If the files are to be read from a data workbench server's File Server Unit, then you must enter the appropriate URI(s) in the Log Paths parameter. For example, the URI <code>/Logs/* .log</code> matches any <code>.log</code> file in the Logs directory. See Configuring a Data Workbench Server File Server Unit on page 540.</p>
Log Server	Information (Address, Name, Port, and so on) necessary to connect to a file server. If there is an entry in the Log Server parameter, the Log Paths are interpreted as URIs. Otherwise, they are interpreted as local paths. See Configuring a Data Workbench Server File Server Unit on page 540.
Compressed	True or false. This value should be set to true if the log files to be read by the data workbench server are compressed gzip files.
Decoder Group	The name of the text file decoder group to be applied to the log file log source. This name must match exactly the name of the corresponding text file decoder group specified in the Log Processing Dataset Include file. See Text File Decoder Groups on page 561.
Log Source ID	<p>This parameter's value can be any string. If a value is specified, this parameter enables you to differentiate log entries from different log sources for source identification or targeted processing. The <code>x-log-source-id</code> field is populated with a value identifying the log source for each log entry. For example, if you want to identify log entries from a log file source named <code>LogFile01</code>, you could type <code>fromLogFile01</code>, and that string would be passed to the <code>x-log-source-id</code> field for every log entry from that source.</p> <p>For information about the <code>x-log-source-id</code> field, see Event Data Record Fields on page 646.</p>
Mask Pattern	<p>A regular expression with a single capturing subpattern that extracts a consistent name used to identify the source of a series of log files. Only the file name is considered. The path and extension are not considered for the regular expression matching. If you do not specify a mask pattern, then a mask is generated automatically.</p> <p>For the files <code>Logs\010105server1.log</code> and <code>Logs\010105server2.log</code>, the mask pattern would be <code>[0-9]{6}(.*).</code> This pattern extracts the string "server1" or "server2" from the file names above.</p> <p>See Regular Expressions on page 642.</p>
Recursive	True or false. If this parameter is set to true, all subdirectories of each path specified in Log Paths are searched for files matching the specified file name or wildcard pattern. The default value is false.
Reject File	The path and file name of the file containing the log entries that do not meet the conditions of the decoder.
Use Start/End Times	True or false. If this parameter is set to true and Start Time or End Time is specified, then all files for this log source must have file names starting with dates in ISO format (YYYYMMDD). It is assumed that each file contains data for one GMT day (for example, the time range starting at 0000 GMT on one day and ending at 0000 GMT the following day). If the log sources file names do not begin with ISO dates, or if the files contain data that do not correspond to a GMT day, then this parameter must be set to false to avoid incorrect results.

Parameter	Description
	<p> Note: If the naming and time range requirements described above are satisfied for the log files and you set this parameter to true, the specified text file decoder group limits the files read to those whose names have ISO dates that fall between the specified Start Time and End Time. If you set this parameter to false, the data workbench server reads all of the log files during log processing to determine which files contain data within the Start Time and End Time range.</p> <p>For information about the Start Time and End Time parameters, see Data Filters on page 546.</p>

In this example, the dataset is constructed from two types of log sources.

Log Source 0 specifies log files generated from event data captured by **Sensor**. This data source points to a directory called Logs and to all of the files in that directory with a .vs1 file name extension.

Log Source 1 points to all of the files in the Logs directory with a .txt file name extension. The decoder group for this log source is called “Text Logs.”

-Log Processing.cfg (modified)	
-Log Processing	LogProcessing
-Log Sources	
-0	
Name	VisualSensor
-Log Paths	VSL Files
0	Logs*.vs1
+Log Server	serverInfo
Log Source ID	false
Recursive	
-1	
Name	LogFile
-Log Paths	Log Files
0	Logs*.txt
+Log Server	serverInfo
Compressed	false
Decoder Group	Text Logs
Log Source ID	from FSU01
Mask Pattern	
Recursive	false
Reject File	
Use Start/End Times	false

You should not delete or move log files after the data sources for a dataset have been defined. Only newly created log files should be added to the directory for the data sources.

XML Log Sources

Information on integrating event data from XML log files.

The file containing the event data must meet the following requirements:

- Event data must be included in a properly formatted XML file with appropriate parent-child relationships.
- A unique decoder group must exist for each XML file format. For information about constructing a decoder group, see [XML Decoder Groups](#) on page 563.
- Each visitor record in the file must contain:

- A tracking ID
- A time stamp
- To specify start and end times for data processing, each file name must be of the form

YYYYMMDD-SOURCE.log

where *YYYYMMDD* is the Greenwich Mean Time (GMT) day of all of the data in the file, and *SOURCE* is a variable identifying the source of the data contained in the file.

For an example of an XML file that meets these requirements, see [XML Decoder Groups](#) on page 563.



Note: Please contact Adobe Consulting Services for a review of the XML log files that you plan to incorporate into the dataset.

Parameters

For XML log sources, the parameters in the following table are available.



Note: The processing of XML log sources requires additional parameters that are defined in a **Log Processing Dataset Include** file, which contains a subset of the parameters included in a **Log Processing.cfg** file as well as special parameters for defining decoders for extracting data from the XML file. For information about defining decoders for XML log sources, see [XML Decoder Groups](#) on page 563.

Log Processing.cfg: XML Log Sources

Field	Description
Name	The identifier for the XML log source.
Log Paths	<p>The directories where the XML log sources are stored. The default location is the Logs directory. A relative path refers to the installation directory of the data workbench server.</p> <p>You can use wildcard characters to specify which XML log sources to process:</p> <ul style="list-style-type: none"> • * matches any number of characters • ? matches a single character <p>For example, the log path <code>Logs* .xml</code> matches any file in the Logs directory ending in <code>.xml</code>.</p> <p>If you want to search all subdirectories of the specified path, you must set the Recursive field to true.</p> <p> Note: If the files are to be read from a data workbench server's File Server Unit, you must enter the appropriate URI(s) in the Log Paths field. For example, the URI <code>/Logs/* .xml</code> matches any <code>.xml</code> file in the Logs directory. See Configuring a Data Workbench Server File Server Unit on page 540.</p>
Log Server	Information (Address, Name, Port, and so on) necessary to connect to a file server. If there is an entry in the Log Server field, the Log Paths are interpreted as URIs. Otherwise, they are interpreted as local paths. See Configuring a Data Workbench Server File Server Unit on page 540.
Compressed	True or false. This value should be set to true if the XML log sources to be read by the data workbench server are compressed gzip files.
Decoder Group	The name of the XML decoder group to be applied to the XML log source. This name must match exactly the name of the corresponding XML decoder group specified in the Log Processing Dataset Include file. See XML Decoder Groups on page 563.

Field	Description
Log Source ID	<p>This field's value can be any string. If a value is specified, this field enables you to differentiate log entries from different log sources for source identification or targeted processing. The x-log-source-id field is populated with a value identifying the log source for each log entry. For example, if you want to identify log entries from a log file source named XMLFile01, you could type <code>from XMLFile01</code>, and that string would be passed to the x-log-source-id field for every log entry from that source.</p> <p>For information about the x-log-source-id field, see Event Data Record Fields on page 646.</p>
Mask Pattern	<p>A regular expression with a single capturing subpattern that extracts a consistent name used to identify the source of a series of log files. Only the file name is considered. The path and extension are not considered for the regular expression matching. If you do not specify a mask pattern, then a mask is generated automatically.</p> <p>For the files <code>Logs\010105server1.xml</code> and <code>Logs\010105server2.xml</code>, the mask pattern would be <code>[0-9]{6}(.*).</code> This pattern extracts the string "server1" or "server2" from the file names above.</p> <p>See Regular Expressions on page 642.</p>
Recursive	True or false. If this parameter is set to true, all subdirectories of each path specified in Log Paths are searched for files matching the specified file name or wildcard pattern. The default value is false.
Reject File	The path and file name of the file containing the log entries that do not meet the conditions of the decoder.
Use Start/End Times	<p>True or false. If this parameter is set to true and Start Time or End Time is specified, then all files for this log source must have file names starting with dates in ISO format (YYYYMMDD). It is assumed that each file contains data for one GMT day (for example, the time range starting at 0000 GMT on one day and ending at 0000 GMT the following day). If the log sources file names do not begin with ISO dates, or if the files contain data that do not correspond to a GMT day, then this parameter must be set to false to avoid incorrect results.</p> <p> Note: If the naming and time range requirements described above are satisfied for the XML files and you set this parameter to true, the specified XML decoder group limits the files read to those whose names have ISO dates that fall between the specified Start Time and End Time. If you set this parameter to false, the data workbench server reads all of the XML files during log processing to determine which files contain data within the Start Time and End Time range.</p> <p>For information about the Start Time and End Time parameters, see Data Filters on page 546.</p>



Note: You should not delete or move XML log sources after the data sources for a dataset have been defined. Only newly created XML files should be added to the directory for the data sources.

Avro Data Feed

The Apache Avro data feed is a new format used by Data Workbench to more efficiently integrate data and provide access to new fields with the updated evars, custom events, and solution variables used in Adobe Analytics Premium.

The Avro data feed provides a more efficient way to integrate data into Data Workbench:

- Avro provides a single-source format for traffic and commerce data.
- The Avro feed is compressed data of multiple source chunks provided per day. It provisions only populated fields and provides monitoring and notification features, access to historical data, and auto-recovery.
- The schema, a self-defining layout of Avro log files, is included at the beginning of each file.
- New fields are added with supporting information to ingest Data Workbench data without any changes required to the decoder. These include:
 - Evars: 1-250 (previously 1-75)

-
- Custom Events: 1-1000 (versus 1-100)
 - Access to solution variables for mobile, social, and video data



Note: In addition, using the Avro feed allows immediate access to any new fields in the feed without a shutdown, allowing the fields to be updated with no service hour requirements.

The Avro data feed is set up in separate files:

- An **Avro Log file**: This is the Avro log format generated from the decoder to format traffic and commerce data.
- An **Avro Decoder file**: This file lets you map values into the new Avro format. You can set up the decoder using the Avro Decoder Wizard.

Avro Decoder Wizard

This wizard sets up the Avro decoder log file.

To open, right-click in a workspace and select **Admin > Wizards > Avro Decoder Wizard**.

Step 1: Select an Avro Log File.

In this step, you can select a source file for the Avro schema. Schemas can be accessed from a log file (.log) or an existing decoder file (.avro). Schemas can be pulled from either file.

Avro Log File	Click to open a log (.log) file to view the schema at the top of the log file and generate decoder file.
Avro Decoder File	Click to open and edit the schema of an existing decoder (.avro) file.

Step 2: Select Input Fields.

Select the input fields to be used in the data set to pass through log processing. All fields in the file will be displayed, allowing you to select fields for the feed.

Select Defaults	Select fields to identify as a standard set of default fields .
Select All	Select all fields in the file.
Deselect All	Clear all fields in the file.



Note: A `x-product` (`Generates row`) field is provided if an array is encountered in the data. This field generates new rows for the nested data in an array as input fields. For example, if you have a `Hit` row with many `Product` values in an array, then rows will be generated in the input file for each product.

Step 3: Select fields that get copied to generate rows.

Because new rows can be created from nested values in an array, every new row created must have a tracking ID and a timestamp. This step allows you to select the fields to be copied to rows from the parent record, such as a tracking ID and timestamp. You can also select other values you want added to each row.

Select Defaults	Select a standard set of default fields that require new column values added to each row, such as a tracking ID and timestamp. For example, a <code>hit_source</code> field is a default value required to be added to each new row (it is defined as a default value in the list). You can add other column values to each row as needed.
Select All	Select all fields in the file.
Deselect All	Clear all fields in the file.

Use the **Search** box to find values in the list.

Step 4:Specify the decoder name

Assign a name for the group of fields and save as a decoder file. The name should match the Decoder group name specified in your log source.

Step 5: Save the decoder file.

The file menu will open to name the decoder file and save as a .cfg file in the **Logs** folder.

Configuring a Data Workbench Server File Server Unit

Information about Insight Server file server units and the file server configuration process.

About Data Workbench Server File Server Units

Information about the data workbench server FSU.

You can configure the data workbench server (InsightServer64.exe) to run as a File Server Unit (FSU) by completing the parameters in the **Log Sources > Log Server** node of the **Log Processing.cfg** file. When the data workbench server is configured to run as an FSU, it stores source files (.vs1 files, text files, or XML files) that can be accessed quickly by multiple processing servers (DPUs). When the DPUs in a data workbench server cluster access the FSU to read the log files, they divide the log files among them and guarantee that the same file is not processed more than once.



Note: When setting up an FSU that serves a data workbench server cluster consisting of five to ten DPUs, you should make the cluster's master server the FSU.

For information about installing a data workbench server cluster, see the *Server Products Installation and Administration Guide*.

The File Server Configuration Process

The **Log Processing.cfg** file specifies the location of a dataset's log sources.

If the location is a remote location, the data workbench server machine that is processing the data connects to the designated remote machine to read the logs.

On the data workbench server machine designated to run as an FSU, the **Access Control.cfg** file lets the DPUs connect to the FSU, and the **Communications.cfg** file maps the location of the remote data files. The process steps to configure an FSU are as follows:

1. In the **Log Processing.cfg** file on your master data workbench server, specify the type of data source and the location of the source. See [Specifying the Data Source](#) on page 541.
2. In the **Access Control.cfg** file on the FSU, edit the permissions to allow the DPUs to connect to the FSU to read the log data. See [Editing the Permissions on the File Server Unit](#) on page 542.
3. In the **Communications.cfg** file on the FSU, edit the settings for the **LoggingServer** and **FileServer** entries to specify the location of the log files. See [Specifying the Location of the Log Files](#) on page 543.
4. If you are configuring your dataset profile to run on a data workbench server cluster, you also must make the cluster's FSU the server where all of the profile's dimensions are constructed:
(For data workbench server clusters only) In the **Communications.cfg** and **cluster.cfg** files on the FSU, add entries for a "normalize server" to make the FSU the server within the cluster where all dimensions are constructed. See [Creating a Centralized Normalization Server for a Cluster](#) on page 544.

For instructions to configure a dataset profile to be processed by a data workbench server cluster, see the *Server Products Installation and Administration Guide*.



Note: The following instructions assume that all of the log files reside in the default directory. If you want to store logs in another directory or create multiple log paths, contact Adobe Consulting Services to discuss your specific configuration.

Specifying the Data Source

When specifying remote data sources for a dataset, you must specify the type of data source and the location of the log files on your master data workbench server.

To specify the data source and its location

1. Open the `Log_Processing.cfg` file. See [Editing the Log Processing Configuration File](#) on page 529.
2. Add a **Sensor**, log file, or XML data source. See [Log Files](#) on page 534.
3. Complete the Log Paths parameter. See [Sensor Files](#) on page 533, [Log Files](#) on page 534, or [XML Log Sources](#) on page 536. Be sure to specify a valid URI.
4. Complete the Log Server parameters defined in the following table:

Log Server Parameters

Parameter	Description
Name	Name identifying the remote file server.
SSL Server Common Name	Server Common Name listed on the file server's SSL certificate. This parameter is optional if Use SSL is set to false.
Address	Address of the file server machine. Can be left blank if Name matches SSL Server Common Name . For example: <code>visual.mycompany.com</code> or <code>192.168.1.90</code> .
Port	Port through which the data workbench server machine communicates with the file server.
SSL Client Certificate	Name of the SSL certificate file for the data workbench server (<code>server_cert.pem</code>).
Use SSL	True or false. True indicates that the file server uses SSL .

If a proxy server is required for the DPUs to connect to the FSU, you need to complete the following parameters:

Log Server Parameters for Proxy Servers

Parameter	Description
Proxy Address	The address of a proxy server that the data workbench server must use to access the file server.
Proxy Password	Optional. The password to the proxy server.
Proxy Port	The port of the proxy server. The default is 8080.
Proxy User Name	Optional. The user name for the proxy server.

Following is an example of a defined **Log Server** in the `Log_Processing.cfg` file. Log Source #1 is a LogFile source that points to a directory called Logs (note the URI specified in the Log Paths parameter) on the machine named FSU01.

-Log Processing.cfg	
-Log Processing	LogProcessing
-Log Sources	VisualSensor
+0	
-1	LogFile
Name	Logs
-Log Paths	/Logs/*.txt
0	
-Log Server	serverInfo
Name	FSU01
SSL Server Common Name	fsu01.mycompany.com
Address	fsu01
Port	80
Proxy Address	
Proxy Password	
Proxy Port	8080
Proxy User Name	
SSL Client Certificate	Certificates\server_cert.pem
Use SSL	false

Editing the Permissions on the File Server Unit

In the previous process, you configured a profile for a given dataset to read log files from an FSU. Now you must edit the permissions on the FSU to allow connections from the DPUs that are running the profile. The following steps walk you through editing the permissions file `Access_Control.cfg`.

To edit permissions on the FSU

1. Open the **Server Files Manager** for the data workbench server machine that you are setting up as your FSU and click **Access Control** to show its contents.
For information about opening and working with the **Server Files Manager**, see the *Data Workbench User Guide*.
2. In the **Server Files Manager** window, click **Access Control** to show its contents. The `Access_Control.cfg` file is located within this directory.
3. Right-click the check mark in the server name column for `Access_Control.cfg`, then click **Make Local**. A check mark appears in the **Temp** column for `Access_Control.cfg`.
4. Right-click the newly created check mark under the **Temp** column and click **Open > in Workstation**.
5. In the **Access Control** window, click **Access Control Groups** to show its contents.
6. Right-click the numeric label for the final **AccessGroup** in the list and click **Add new > Group**.
7. Enter a **Name** for the new **AccessGroup**. Example: Connecting Servers.
8. Right-click **Member** under the new **AccessGroup**, then click **Add new > Member**.
9. Enter the IP address for the data workbench server's DPU that connects to this file server.
10. Repeat steps 4 and 5 for any other data workbench server DPUs that connect to this FSU, including the data workbench server DPUs in a cluster that must access the log files.
11. Right-click **Read-Only Access** under the new **AccessGroup**, then click **Add new > URI**.
12. Enter the location of the stored log files on the file server machine. Use forward slashes (/) in the path specification. The default location is `/Logs/`.

-
13. Right-click (**modified**) at the top of the window, then click **Save**.
 14. In the **Server Files Manager** window, right-click the check mark for **Access Control.cfg** in the **Temp** column, then click **Save to > <server name>** to save the locally made changes to the data workbench server's FSU.

Specifying the Location of the Log Files

You must edit the **Communications.cfg** file on the FSU to specify the location of the log files.

To specify the location of the log files

1. In the **Server Files Manager** window, click **Components** to show its contents. The **Communications.cfg** file is located within this directory.
2. Right-click the check mark in the server name column for **Communications.cfg**, then click **Make Local**. A check mark appears in the **Temp** column for **Communications.cfg**.
3. Right-click the newly created check mark under the **Temp** column and click **Open > in Workstation**.
4. In the **Communications.cfg** window, click **component** to show its contents.
5. In the **Communications.cfg** window, click **Servers** to show its contents. Several servers may appear: File Servers, Logging Servers, Init Servers, Status Servers, Send Servers, or Replicate Servers.
6. (For **Sensor** log sources only) Find the **LoggingServer**, which is where **Sensor** writes its log files to be processed by the data workbench server, then click its number to view the menu. Edit the **Log Directory** parameter to reflect the desired location of the log files. The default log directory is the **Logs** folder within the data workbench server's installation directory.

Do not modify any other parameters for the **LoggingServer**.

-Communications.cfg	
-component	CommServer
-Servers	
+0	InitServer
-1	LoggingServer
Flush Interval	10000
Log Directory	Logs\
URI	/Submit.vsp
+2	SendServer
+3	FileServer
+4	FileServer
+5	FileServer

7. Find the **FileServer** that specifies the location of log files. There may be several **File Servers** listed under **Servers**, so you may need to view the contents for many of them (by clicking the server number) to find the desired server.
8. Edit the **Local Path** and **URI** parameters for the **FileServer** to reflect the location of the log files. The following example shows that the log files reside in the **Logs** folder within the data workbench server's installation directory:

-Communications.cfg	
-component	CommServer
-Servers	
+0	InitServer
+1	LoggingServer
+2	SendServer
+3	FileServer
+4	FileServer
+5	FileServer
+6	FileServer
+7	FileServer
+8	FileServer
-9	FileServer
Local Path	Logs\
URI	/Logs/
+10	StatusServer
+11	FileServer



Note: If the **Local Path** and **URI** parameters are populated as shown, you can access the log files on the FSU from any data workbench server by clicking **Logs** in the **Server Files Manager**.

9. Right-click (**modified**) at the top of the configuration window, then click **Save**.
10. In the **Server Files Manager** window, right-click the check mark for **Communications.cfg** in the **Temp** column, then click **Save to > <server name>** to save the locally made changes to the data workbench server's FSU.

Creating a Centralized Normalization Server for a Cluster

If you are configuring your dataset profile to run on a data workbench server cluster, then you should make the cluster's FSU the server where all of the profile's dimensions are constructed.

Adobe strongly recommends that the cluster's FSU serves as the cluster's master server and its centralized normalization server.

To make the FSU the centralized normalization server, you must open and edit the **Communications.cfg** and **Cluster.cfg** files on the FSU.

To make the FSU the centralized normalization server

1. Add a **NormalizeServer** entry to the **Communications.cfg** file on the FSU.



Note: If you have installed the complete release package for data workbench server v5.0 or later, the **Communications.cfg** file on your FSU should have a **NormalizeServer** entry already. You can follow the steps below to confirm that the entry exists.

- a. Open the **Communications.cfg** file in data workbench as described in [*#unique_385/unique_385_Connect_42_section_F9A649BF1B2544FEB10AD8820384EDB0*](#).
- b. Click **component** to show its contents.
- c. Right-click **Servers** and click **Add New > Centralized Normalization Server**.
- d. In the **URI** parameter for the **NormalizeServer**, type **/Cluster/**.

-Communications.cfg	
-component	CommServer
-Servers	
+0	InitServer
+1	LoggingServer
+2	SendServer
+3	FileServer
+4	FileServer
+5	FileServer
+6	FileServer
+7	FileServer
+8	FileServer
+9	FileServer
+10	StatusServer
+11	FileServer
+12	FileServer
+13	FileServer
+14	FileServer
+15	FileServer
+16	FileServer
-17	NormalizeServer
URI	/Cluster/

- e. Right-click **(modified)** at the top of the window, and click **Save**.
 - f. In the **Server Files Manager** window, right-click the check mark for **Communications.cfg** in the **Temp** column, then click **Save to > <server>** name to save the locally made changes to the data workbench server FSU.
2. Define the centralized normalization server in the **Cluster.cfg** file on the master server in your data workbench Server cluster.



Note: If the FSU on which you are setting up your centralized normalization server is not the master data workbench Server in your cluster, you must add the IP addresses of the DPUs in the cluster to the **Cluster Servers** access group in the FSU's **Access Control.cfg** file. For instructions to add servers to the **Cluster Servers** group, see Updating the Access Control File for a Cluster section in the *Server Products Installation and Administration Guide*.

- a. Open the **Profile Manager** within your dataset profile, then click **Dataset** to show its contents. The **Cluster.cfg** file is located within this directory.
- b. Right-click the check mark next to **Cluster.cfg**, then click **Make Local**. A check mark for this file appears in the **User** column.
- c. Right-click the newly created check mark and click **Open > in Notepad**.
- d. Add the text that is highlighted in the following file fragment:

```
Cluster = ClusterConfig:  
Normalize Server = serverInfo:  
Address = string:  
Port = int: 80
```

```
SSL Server Common Name = string: server common name
Use SSL = bool: false
```



Note: When you enter the common name of FSU for the SSL Server Common Name parameter, the FSU uses its .address file to resolve the common name. For information about the .address file, see the *Server Products Installation and Administration Guide*.

- e. Save the file.
- f. In the **Profile Manager**, right-click the check mark for `Cluster.cfg` in the **User** column, then click **Save to > <dataset profile name>** to save the locally made changes to the dataset profile.

Log Processing Parameters

Navigation title:

Links to additional information about specific parameters in the `Log Processing.cfg` file.

Data Filters

Data filters in the `Log Processing.cfg` file limit the log entries that are considered in the dataset construction process.

The filters defined in the `Log Processing.cfg` file include the following:

- End Time
- Hash Threshold
- Start Time

The filtering defined by these parameters occurs after log entries leave the decoders and after transformations but before their evaluation by the **Log Entry Condition**. In general, changing any of these parameters results in changes to the composition of the dataset.

The recommended technique for using **Sensor** data sources to construct a dataset that covers a specific period of time is to use the Start Time and End Time parameters for the dataset.

Using the Start Time and End Time parameters is preferred to other techniques, such as moving log files to separate them by directory. By setting the start and end times for the dataset, the data workbench server automatically uses only those log entries that occurred within the given interval. Assuming that the End Time is in the past, the data workbench server typically updates the dataset using the same set of log entries, even if the dataset is updated by, for example, adding a new transformation.

Log Entry Condition

The **Log Entry Condition** determines which log entries from the total set of entries available in the dataset's data sources should continue to flow through the configuration process.

In essence, it is a filtering process on the available log entries. If the **Log Entry Condition** returns a value of false, the log entry is filtered out of the available set of log entries.

The **Log Entry Condition** is described through the use of condition operations (see [Conditions](#) on page 634) and can use any of the input fields collected by **Sensor** (see the *Data Workbench Sensor Guide*) or any extended fields produced by transformations contained within the `Log Processing.cfg` file to define the test conditions. **Log Entry** conditions are applied during log processing and optionally can be applied during transformation.

This example demonstrates the use of the **log entry condition** for website data. You can use the **Log Entry Condition** to create datasets that focus on a specific portion of the website or visitors performing some specific action on the site.

The **Log Entry Condition** in this example creates a dataset that includes only those log entries that are part of the site's store. By using the **RECondition test** with the matching pattern "`/store/.*`" and the **cs-uri-stem**

field as input to the regular expression, only web pages that start with the string "/store/" are included in the dataset.

-Log Entry Condition	AndCondition
-0	RECondition
-Matches	/store/.*
0	true
Case Sensitive	Comment
Comments	cs-uri-stem
Input	

Key Splitting

The key splitting capability of the data workbench server enables tracking IDs associated with a large number of log entries to be split into multiple tracking IDs.

The number of tracking IDs in the dataset is artificially increased, but the total number of log entries processed by the data workbench server is not artificially increased, thereby preserving the total number of countable events in the dataset. After the data for a single element is split, the data is forever associated with two different tracking IDs and cannot be related.

For example, if you are working with web data, each tracking ID represents a unique visitor. If you enable key splitting, the visitors in your dataset with large amounts of event data are split into multiple visitors. While the number of visitors in the dataset is artificially increased, the total number of countable events such as page views or bookings is not artificially increased. After splitting occurs, the data for the sub-visitors cannot be related.

Key splitting uses a probabilistic algorithm. As a result, there is a trade-off between memory usage, the failure probability, the key splitting threshold (**Split Key Bytes**), and the dataset size. With the recommended settings (as listed below), the failure rate is low. Of those elements whose event data exceeds the key splitting threshold, approximately 1 in 22,000 (usually less than 1 per dataset) will have some of their data truncated rather than split.

The recommended values for each parameter (without and with key splitting) are shown in the following table.

Log Processing.cfg: Key Splitting

Parameter	No Key Splitting	Key Splitting
Group Maximum Key Bytes	1e6	2e6
Split Key Bucket Space	6e6	6e6
Split Key Bytes	0	1e6
Split Key Space Ratio	10	10

Group Maximum Key Bytes specifies the maximum amount of event data that can be processed for a single tracking ID. Data exceeding this limit is filtered from the dataset construction process. **Split Key Bytes** represents the number of bytes at which a single tracking ID is split into multiple elements. Elements are split at approximately this number of bytes according to a probability distribution. **Split Key Space Ratio** and **Split Key Bucket Space** control the memory utilization and failure rate of key splitting.



Note: **Group Maximum Key Bytes**, **Split Key Bytes**, **Split Key Space Ratio**, and **Split Key Bucket Space** all must be declared for key splitting to work properly. Do not change the values of these parameters without consulting Adobe.

ODBC Data Sources

The data workbench server (InsightServer64.exe) can read event data from any SQL database (for example, Oracle or Microsoft SQL Server) that has an ODBC 3.0 compliant driver.

The data workbench server's ODBC support is similar to existing support for loading data from Sensors or from log files generated by external processes. There are, however, some additional considerations and limitations:

- The data workbench server's ODBC support is compatible with the clustering capabilities. Data is distributed among all processing servers, and all subsequent processing (including query processing) benefits fully from clustering.
- ODBC support depends on third-party ODBC drivers. For ODBC support to work, these drivers must be configured on the machine on which the data workbench server runs, using tools external to the Adobe platform. Data workbench machines do not require any additional configuration.
- The table or view from which data is loaded must have an increasing ID column. For any row, the value in this column (which may be an actual column in the table or any SQL column expression) must not decrease as new rows are inserted into the database. If this constraint is violated, data is lost. For adequate performance, an index is required on this column or column expression.



Note: It is possible for multiple rows to have the same value in the **Increasing ID** column. One possibility is a timestamp column with less than perfect precision.

- The data workbench server cannot load columns with long data (data above a certain length as determined by the specific database application in use).
- Retrieving data from a database is slower than reading it from a disk file. Datasets that load data from an ODBC source take much longer to process (particularly when reprocessing) than equivalently sized datasets whose data comes from Sensors or other disk files.

For information about reprocessing your data, see [Reprocessing and Retransformation](#) on page 627.

To configure Insight Server for ODBC event data

Configuring the data workbench server to load data from an SQL database requires that you first perform the following steps in order:

1. Install the appropriate database client software, including an ODBC driver, on the data workbench server machine on which the dataset is processed.



Note: If you are loading ODBC event data for processing on a data workbench server cluster, you must install the database client software on all of the processing servers in the cluster. For information about specifying processing servers in a cluster, see the *Server Products Installation and Administration Guide*.

2. Configure a Data Source using the ODBC Data Source Administrator for Windows.

It is important to note that the data workbench server (InsightServer64.exe) runs as a Windows service. Therefore, the Data Source ordinarily must be configured as a System DSN rather than a User DSN for the data workbench server to be able to use it. You can find more information about this configuration step in the documentation for your database software.

After installing the database client software on the appropriate data workbench server machine, you can configure the dataset to use the ODBC data source by editing the appropriate parameters in the **Log Processing** configuration file for the desired profile.

Parameters

For data from databases using the Open Database Connectivity (ODBC) standard, the following parameters are available:

Log Processing.cfg: ODBC Data Sources

Parameter	Description
Name	The identifier for the ODBC source.
Data Source Name	A DSN, as provided by an administrator of the data workbench server machine on which the dataset is processed, that refers to the database from which data is to be loaded.
Database Password	The password to be used when connecting to the database. If a password has been configured for the DSN in the Data Source Administrator , this may be left blank. Any password supplied here overrides the password configured for the DSN in the Data Source Administrator .
Database User ID	The user ID to be used when connecting to the database. If a user ID has been configured for the DSN in the Data Source Administrator , this may be left blank. Any user ID supplied here overrides the user ID configured for the DSN in the Data Source Administrator .
Fields	A vector of column objects that specifies a mapping from data columns in the database to data fields in the data workbench server execution engine. Each column has entries Column Name and Field Name . Column Name is an SQL column expression that must be valid in the context of the table identified by Table Identifier described above. It may be a column name or any SQL expression based on any number of columns in the table. A formatting function may be necessary to convert values of certain data types to strings in a way that does not lose precision. All data is implicitly converted to strings using the database's default formatting method, which may cause data loss for some column data types (such as date/time data types) if explicit formatting expressions are not used.
Increasing ID Column	<p>A column name or SQL column expression that meets the criterion that it increases (or at least does not decrease) as new rows are added. That is, if Row B is added to the table at a later time than Row A, the value of this column (or column expression) in Row B must be greater (according to the database's native sorting order) than the corresponding value in Row A.</p> <ul style="list-style-type: none"> The Increasing ID Column name may be the same as the name of an existing column, but is not required to be. This expression is assumed to have an SQL character data type. If the actual increasing ID column is of some other data type, this value must be a column expression to convert it to a string. Because this usually means comparisons are lexicographic (character by character), it is important to format the value carefully. The expression is used in SQL ORDER BY clauses and compared to in SQL WHERE clauses. It is critically important to have an index built on the exact column expression that is used.
Log Source ID	<p>This parameter's value can be any string. If a value is specified, this parameter enables you to differentiate log entries from different log sources for source identification or targeted processing. The x-log-source-id field is populated with a value identifying the log source for each log entry. For example, if you want to identify log entries from an ODBC source named ODBCSOURCE01, you could type <code>from ODBCSOURCE01</code>. and that string would be passed to the x-log-source-id field for every log entry from that source.</p> <p>For information about the x-log-source-id field, see Event Data Record Fields on page 646.</p>
Run On Server	Index value in the <code>profile.cfg</code> file of the processing server that makes the ODBC queries to get data from the database. (The Processing Servers parameter in the <code>profile.cfg</code> file lists all of the processing servers for the dataset, and each server has an index value, the first being 0.) The default value is 0.
Table Identifier	An SQL expression that names the table or view from which data is to be loaded. A typical table identifier is of the form SCHEMA.TABLE.

This example shows the **Log Processing** configuration window in data workbench with an ODBC data source. This Data Source takes data from a table called `VISUAL.VSL` in a database with **Data Source Name** "VSTestO." Five (5) column objects (**Fields**) map data from the data columns in the database to the data workbench server.

-Log Processing.cfg	
-Log Processing	LogProcessing
+Log Entry Condition	AndCondition
-Log Sources	
-0	ODBCLogSource
Name	Sample ODBC Data
-Fields	
-0	Column
Column Name	"TIMESTAMP"
Field Name	x-timestamp
-1	Column
Column Name	"X-TRACKINGID"
Field Name	x-trackingid
-2	Column
Column Name	"C-IP"
Field Name	c-ip
-3	Column
Column Name	"CS(REFERRER)"
Field Name	cs(referrer)
-4	Column
Column Name	"CS(COOKIE)"
Field Name	cs(cookie)
Data Source Name	VSTest0
Database Password	visual
Database User ID	visual
Increasing ID Column	to_char("TIMESTAMP",'YYYY_MM_DD HH24:MI:SS')
Log Source ID	from ODBCSOURCE01
Run On Server	0
Table Identifier	VISUAL.VSL

Considerations for the Log Processing Configuration File

Navigation title: Considerations for the Log Processing Configuration File

Conceptual information to consider when editing the Log Processing.cfg file.

- Data files should not be moved between directories after the sources for a dataset have been defined. The only additional files a directory should receive are newly created ones that result from the data workbench server receiving data from Sensor(s).
- Changing any of the parameters in this file requires reprocessing of all the data. The Pause parameter in the Log Processing Mode.cfg file must be set to false for reprocessing to occur. (Note that this parameter's default value is false, so changing the parameter may not be required.) For information about the Log Processing Mode.cfg file, see [Additional Configuration Files](#) on page 570.
- If you reprocess the data, you can check the Log Processing Progress parameter in data workbench's **Processing Legend**.

For information about reprocessing your data, see [Reprocessing and Retransformation](#) on page 627. See [Processing Legend](#) on page 451.

- The Log Processing.cfg file can be shared by multiple dataset profiles. Transformations defined in the Log Processing.cfg file are applied to all dataset profiles that share this configuration file.



Note: Adobe recommends defining transformations for the log processing in one or more log processing **dataset include** files. For information, see [Log Processing Dataset Include Files](#) on page 559.

- You can add any of the parameters described above to the `Log_Processing.cfg` file by opening and editing the file in Notepad. Any changes you make and save appear when you reopen the file in data workbench. When adding a new parameter, use the Space key (not the Tab key) to indent two (2) spaces to the right of the previous heading level.

Any errors that occur during the log processing phase of the dataset construction process for a dataset profile are shown in the **Profiles** node of the **Detailed Status** interface in data workbench. For information about the **Detailed Status** interface, see the *Data Workbench User Guide*.

Transformation Configuration File

Navigation title: Transformation Configuration File

Information about the `Transformation.cfg` file and how its parameters affect the transformation phase of the dataset construction process.

About the Transformation Configuration File

Navigation title: About the Transformation Configuration File

The `Transformation.cfg` file controls the transformation phase of dataset construction during which additional data transformations are applied to data already processed during log processing to create extended dimensions for use in analysis.

You must edit the `Transformation.cfg` file to perform any of the following dataset configuration tasks:

- Filtering data from log processing by defining the **log entry condition** for transformation.
- Setting the time zone to be used for creating time dimensions and making time conversions. See [Time Zones](#) on page 554.



Note: **Transformation Dataset Include** files can contain additional instructions for the transformation phase of dataset construction. These files exist within the `Dataset\Transformation` directory for any inherited profile, and they typically define application-specific parameters (such as web-specific configuration parameters for the **Site** application). For information about **Transformation Dataset Include** files, see [Dataset Include Files](#) on page 557. For information about web-specific configuration parameters for Site, see [Configuration Settings for Web Data](#) on page 649.

Editing the Transformation Configuration File

Navigation title: Editing the Transformation Configuration File

Steps for editing the `Transformation.cfg` file for a dataset profile.

1. While working in your dataset profile, open the **Profile Manager** and click **Dataset** to show its contents.
For information about opening and working with the **Profile Manager**, see the *Data Workbench User Guide*.



Note: A Transformation subdirectory may exist within the `Dataset` directory. This subdirectory contains the **Transformation Dataset Include** files that have been created for one or more inherited profiles. For information about **Transformation Dataset Include** files, see [Dataset Include Files](#) on page 557.

2. Right-click the check mark next to `Transformation.cfg` and click **Make Local**. A check mark for this file appears in the **User** column.
3. Right-click the newly created check mark and click **Open > in Workstation**. The `Transformation.cfg` window appears.

You also can open the `Transformation.cfg` file from a **Transformation Dependency Map**. For information about **transformation dependency maps**, see [Dataset Configuration Tools](#) on page 628.

4. Edit the parameters in the configuration file using the following table as a guide.

When editing the `Transformation.cfg` file within a data workbench window, you can use shortcut keys for basic editing features, including cut (Ctrl+x), copy (Ctrl+c), paste (Ctrl+v), undo (Ctrl+z), redo (Ctrl+Shift+z), select section (click+drag), and select all (Ctrl+a). In addition, you can use the shortcuts to copy and paste text from one configuration file (`.cfg`) to another.



Note: A **Transformation Dataset Include** files for an inherited profile contains a subset of the parameters described in the following table as well as some additional parameters. For information about **Transformation Dataset Include** files, see [Dataset Include Files](#) on page 557

Transformation.cfg

Parameter	Description
End Time	<p>Optional. Filter data to include log entries with timestamps up to, but not including, this time. Adobe recommends using one of the following formats for the time:</p> <ul style="list-style-type: none">• January 1 2013 HH:MM:SS EDT• Jan 1 2013 HH:MM:SS GMT <p>For example, specifying "July 29 2013 00:00:00 EDT" as the End Time includes data through July 28, 2013, at 11:59:59 PM EDT.</p> <p>You must specify a time zone. The time zone does not default to GMT if not specified. For a list of time zone abbreviations supported by the data workbench server, see Time Zone Codes on page 658.</p> <p> Note: If you specify a value for End Time, a parameter named End Time is set and applied throughout the transformation phase of dataset construction. For information about parameters, see Defining Parameters in Dataset Include Files on page 568.</p>
Extended Dimensions	Optional. Adobe recommends defining extended dimensions in one or more Transformation Dataset Include files. For information, see Transformation Dataset Include Files on page 567.
Hash Threshold	Optional. A sampling factor for random sub-sampling of rows. If set to a number n, then only one out of each n tracking IDs enters the dataset, reducing the total number of rows in the dataset by a factor of n. To create a dataset that requires 100 percent accuracy (that is, to include all rows), you would set Hash Threshold to 1. If Hash Threshold is specified in both the <code>Log Processing.cfg</code> and <code>Transformation.cfg</code> files, it is not applied in sequence; the maximum of the values set in either configuration file applies.
Log Entry Condition	Optional. Defines the rules by which log entries output from log processing are considered for inclusion in the dataset profile. See Log Entry Condition on page 546.
New Visitor Condition	Optional. For use with web data. Defines the rules by which visitors are considered for inclusion in the data. The New Visitor Condition defines the first log entry for a visitor (ordered by time) that is to be used in the dataset. All subsequent log entries for this visitor are included in the dataset regardless of whether they meet this condition. See New Visitor Condition on page 554.
Reprocess	Optional. Any character or combination of characters can be entered here. Changing this parameter and saving the file initiates data retransformation. For information about reprocessing your data, see Reprocessing and Retransformation on page 627.
Schema Checking	True or false. If true, then the data workbench server identifies dataset corruption problems and records information about the problems in log files in the data workbench server's Trace directory. The default value is true. Adobe recommends leaving this parameter set to true at all times.
Stages	Optional. The names of the processing stages that can be used in Transformation Dataset Include files. Processing stages provide a way to order the transformations that are defined in Transformation Dataset

Parameter	Description
	<p>Include files. This parameter is very helpful if you have defined one or more transformations within multiple Transformation Dataset Include files and you want specific transformations to be performed at specific points during transformation.</p> <p>The order in which you list the stages here determines the order in which the transformations in the Transformation Dataset Include files are executed during transformation. Preprocessing and Postprocessing are built-in stages; Preprocessing is always the first stage, and Postprocessing is always the last stage. By default, there is one named stage called Default.</p> <p>To add a new processing stage</p> <ul style="list-style-type: none"> • In the <code>Transformation.cfg</code> window, right-click Stages and click Add New > Stage. • Enter a name for the new stage. <p>To delete an existing processing stage</p> <ul style="list-style-type: none"> • Right-click the number corresponding to the stage that you want to delete and click Remove<#stage_number>. <p> Note: When you specify a Stage in a Transformation Dataset Include files the name of the stage must match exactly the name that you enter here. For more information about dataset include files, see Dataset Include Files on page 557.</p>
Start Time	<p>Optional. Filter data to include log entries with timestamps at or after this time. Adobe recommends using one of the following formats for the time:</p> <ul style="list-style-type: none"> • January 1 2013 HH:MM:SS EDT • Jan 1 2013 HH:MM:SS GMT <p>For example, specifying July 29 2013 00:00:00 EDT as the Start Time includes data starting from July 29, 2013, at 12:00:00 AM EDT.</p> <p>You must specify a time zone. The time zone does not default to GMT if not specified. For a list of time zone abbreviations supported by data workbench Server, see Time Zone Codes on page 658.</p> <p> Note: If you specify a value for Start Time, a parameter named Start Time is set and applied throughout the transformation phase of dataset construction. For information about parameters, see Defining Parameters in Dataset Include Files on page 568.</p>
Transformations	<p>Optional. Adobe recommends defining transformations for the transformation phase of dataset construction in one or more Transformation Dataset Include files. For information, see Transformation Dataset Include Files on page 567.</p>
Time Zone	<p>Time zone of the dataset profile. Time zones are used for time conversions and for creating time dimensions. See Time Zones on page 554.</p> <p> Note: When defined in the <code>Log Processing.cfg</code> file, the Time Zone parameter is used for time conversions only.</p>

5. Right-click (**modified**) at the top of the window and click **Save**.
6. In the **Profile Manager**, right-click the check mark for `Transformation.cfg` in the **User** column, then click **Save to > <dataset profile name>** to make the locally made changes take effect. Retransformation of the data begins after synchronization of the dataset profile.

 **Note:** Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

For information about reprocessing or retransforming your data, see [Reprocessing and Retransformation](#) on page 627.

Information About Specific Transformation Parameters

Navigation title:Information About Specific Transformation Parameters

Information about specific parameters that you can define in the Transformation.cfg file.

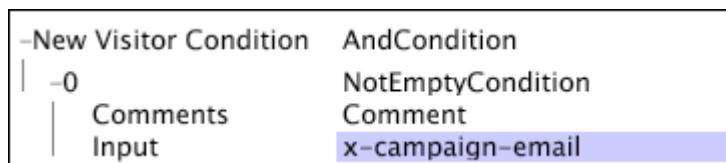
New Visitor Condition

Navigation title:New Visitor Condition

The **New Visitor Condition** is a **Condition Operation** that is used with website data to determine which visitors are considered for inclusion in the dataset.

The **New Visitor Condition** defines the first log entry (ordered by time) for a visitor that is to be used in the dataset, and all subsequent log entries for this visitor are included in the dataset regardless of whether they meet this condition. Because the **New Visitor Condition** requires that data is ordered by visitor and time, it is applied only during the transformation phase of dataset construction.

The **New Visitor Condition** shown in this example creates a dataset that includes only those log entries for visitors who respond to email campaigns. This is accomplished by using the **NotEmptyCondition** test (see [Not Empty](#) on page 636) and the **x-campaign-email** field as input to the regular expression. After the new visitors meeting the condition are identified, all log entries for those visitors are captured.



Time Zones

Navigation title:Time Zones

The Time Zone parameter in the Transformation.cfg file controls time dimensions, time conversions (for example, defining the **x-local-timestring** field), and formatting of all local times in the dataset profile.



Note: The Time Zone parameter does not affect system-level functionality such as timestamps in status and event logs, which are expressed in system local time.

The Time Zone parameter supports a system-independent time zone format ("Coordinated Universal Time") of the following format:

Time Zone = string: UTC +hhmm dstrules

The sign (+) can be either a plus (+) or a minus (-) sign, and *hhmm* is the offset from UTC in hours and minutes. The optional variable *dstrules* specifies a set of rules to implement Daylight Saving Time or a similar clock-shifting policy.

If you specify *dstrules*, a tab-delimited file named *dstrules.dst* must be present within the dataset profile's *Dataset\TimeZone* subdirectory. The file specifies a time-zone independent set of rules for Daylight Saving Time. You can have different sets of rules for different years. The *DST.dst* file provided by Adobe in the Base profile specifies the standard U.S. rules established by the Energy Policy Act of 2005 (in effect starting 2007) and the U.S. rules for prior years.

Sample Time Zone entries are listed below:

- U.S. Eastern Daylight Time: Time Zone = string: UTC -0500 DST
- UTC time with no offset and no dstrules : Time Zone = string: UTC -0000

When this format is used, the system time zone of data workbench server, data workbench, and **Report** machines need not be the same as the specified time zone. In addition, all active dataset profiles on a data workbench server machine need not have the same time zone setting.

Adobe does not recommend running more than one dataset profile on a single data workbench server machine or data workbench server cluster.

Data workbench users will see data in the dataset profile's time zone instead of their system time zone. Adobe recommends that the system time zone for a data workbench server machine be the same as the time zone used in its datasets.



Note: You can specify a Time Zone parameter in the `Log_Processing.cfg` file, where it is used for time conversions during log processing. For information about the Time Zone parameter in the `Log_Processing.cfg` file, see [Log Processing Configuration File](#) on page 528.

Considerations for the Transformation Configuration File

Navigation title: Considerations for the Transformation Configuration File

Important information to consider when editing the `Transformation.cfg` file.

- Changing any of the parameters in this file requires retransformation of the data.
- If you reprocess the data, you can check the **Transformation Progress** parameter in data workbench's **Processing Legend**.

For information about reprocessing your data or the **Processing Legend**, see [Reprocessing and Retransformation](#) on page 627.

- **CrossRows**, **ODBCLookup**, **Sessionize**, and **AppendURI** transformations work only when defined in a **Transformation Dataset Configuration** file. For information about these transformations, see [Data Transformations](#) on page 573.



Note: Adobe recommends defining transformations for the transformation phase of dataset construction in one or more **Transformation Dataset Include** files. For information, see [Transformation Dataset Include Files](#) on page 567.

- You can add any of the parameters described above to the `Transformation.cfg` file by opening and editing the file in Notepad. Any changes you make and save appear when you reopen the file in data workbench. When adding a new parameter, use the Space key (not the Tab key) to indent two (2) spaces to the right of the previous heading level.

Any errors that occur during the transformation phase of the dataset construction process for a dataset profile are shown in the **Profiles** node of the **Detailed Status** interface in data workbench. For information about the **Detailed Status** interface, see the *Data Workbench User Guide*.

DeviceAtlas Distribution

The DeviceAtlas JSON file will now be distributed in a .bundle file (a renamed .tar.gz) along with DeviceAtlas.dll and DeviceAtlas64.dll files.

When the administrator upgrades the Insight Server to version 6.0, the DeviceAtlas.bundle file is included with the upgrade package in the Software and Docs profile (softdocs profile) located at:

Server Packages > v6.00 > Server_6.00.zip

The DeviceAtlas.bundle file is extracted to Server\Lookups\DeviceAtlas.

The DeviceAtlas.bundle file should be placed in a directory that is synchronized to the DPUs, and the DeviceAtlas.cfg file corresponding to the new DeviceAtlasComponent should be placed in the "Components for Processing Servers" directory on the synchronization master. When the DeviceAtlas.bundle file is changed, the very next DeviceAtlas lookup call will get results based on the updated API and/or JSON file.

Modify the Transformation.cfg file

The DeviceAtlas Transformations will no longer need to specify the path to the JSON file. Any previous DeviceAtlasTransformation that is defined in the transformation.cfg file should no longer include the File parameter that points to the obfuscated JSON file.

This example Transformation.cfg file shows the File argument that should be deleted to avoid confusion. (Leaving it there will not cause harm, but only potential confusion because it will be ignored.)

```
6 = DeviceAtlasTransformation:  
  Comments = Comment: 0 items  
  Condition = AndCondition: 0 items  
  
File = string: Lookups\\DeviceAtlas\\20110106_private.json.obfuscated  
^^ DELETE THE ABOVE LINE FROM ALL PREVIOUS TRANSFORMATIONS ^^  
  
Name = string: DeviceAtlas Lookup  
Outputs = vector: 4 items  
  0 = Column:  
    Column Name = string: vendor  
    Field Name = string: x-vendor  
  1 = Column:  
    Column Name = string: model  
    Field Name = string: x-model  
  2 = Column:  
    Column Name = string: isBrowser  
    Field Name = string: x-isbrowser  
  3 = Column:  
    Column Name = string: usableDisplayHeight  
    Field Name = string: x-usable-display-height  
User Agent = string: x-ua
```

Modify the DeviceAtlas.cfg file

This is an example of the component argument required in the DeviceAtlas.cfg file.

```
component = DeviceAtlasComponent:  
  DeviceAtlas Bundle File = string:Lookups\\DeviceAtlas\\DeviceAtlas.bundle  
  
  Unsyncronized Bundle Extraction Path = string: Temp\\DeviceAtlas\\
```

This DeviceAtlas.bundle file will be treated just like a configuration file from the perspective of the Profile Synchronization feature. In addition, the JSON data and DLL will be used at the Component level rather than at the individual Transformation level.

A new DeviceAtlasComponent, upon startup, finds the .bundle conglomeration, de-obfuscates the JSON file into memory, extracts the files into a temporary directory, and loads the appropriate DLL for the running platform. This component also monitors changes to the bundle file, and reloads the DLL and .cfg file automatically if it changes.

Running DeviceAtlas

Proper configuration makes a big difference in the time required for transformation. The transformation can be configured to run only once per visitor per session to allow DeviceAtlas to speed up the process.

If deployed using Log Processing.cfg:

Run the transformations twice.

1. Look up only the mobile_id field, then
2. Create conditions to ignore the mobile_id and then look up the rest of the fields.

If deployed using Transformation.cfg:

Deploy as in Step 1 in Log Processing above, or use cross-rows to support a conditional setting.

- Cross-Rows—Grab the previous session key. Then identify if the current session key is different from the one found with cross-rows. If so, then the DeviceAtlas transformation will only run on one record per session.

Dataset Include Files

Navigation title:Dataset Include Files

Information about dataset include files and how their parameters affect the log processing and transformation phases of the dataset construction process.

About Dataset Include Files

Navigation title:About Dataset Include Files

Many of the internal profiles that you received with your Adobe application come with their own dataset configuration files.

Because the internal profiles are sub-profiles of the dataset profile, their dataset configuration files contain rules that provide additional parameters for the log processing or transformation phases of dataset construction. The dataset configuration files for internal profiles and for any inherited profiles that you create are called dataset include files.

A dataset include file contains a subset of the parameters contained in the main dataset configuration files (`Log Processing.cfg` or `Transformation.cfg`) for the dataset profile. Dataset include files containing parameters associated with log processing are called **Log Processing Dataset Include** files (see [Log Processing Dataset Include Files](#) on page 559), while dataset include files associated with transformation are called **Transformation Dataset Include** files. See [Transformation Dataset Include Files](#) on page 567. You can create multiple dataset include files for use in the dataset construction process. The complete dataset includes all of the fields, transformations, and extended dimensions defined in all of the dataset configuration files for the dataset profile and any inherited profiles.

Working with Dataset Include Files

Navigation title:Working with Dataset Include Files

Dataset include files provide a flexible way to configure your dataset.

Within each file, you can define as few or as many fields, transformations, or dimensions as you desire, and you can organize the include files based on the inherited profile to which they belong. When configuring your dataset, you have the option of editing the dataset include files provided with the internal profiles for your Adobe application or creating new dataset include files for any inherited profiles that you create.

When you edit the parameters of a dataset include file for an internal profile and save the updated file to your dataset profile or an inherited profile that you create, you are, in effect, overriding the file's original settings. Adobe recommends editing a dataset include file for an internal profile whenever you need to make minor changes to the dataset's contents, such as changing a Condition parameter or a parameter's default setting. See [Editing Existing Dataset Include Files](#) on page 557. However, when you want to specify a new field to be passed from log processing to transformation, update or create new fields using transformations, or define extended dimensions, it is best to create a new dataset include file. See [Creating New Dataset Include Files](#) on page 558. You can edit the file that you create whenever or however you see fit.

Editing Existing Dataset Include Files

Navigation title:Editing Existing Dataset Include Files

Steps to editing existing dataset include files.

You open an existing dataset include file using the **Profile Manager** in data workbench.

For information about opening and working with the **Profile Manager**, see the *Data Workbench User Guide*.

-
1. While working in your dataset profile, open the **Profile Manager** and click **Dataset** to show the contents of the directory.
 - To open a **Log Processing Dataset Include** file, click **Log Processing** to show the contents of the directory.
 - To open a **Transformation Dataset Include** file, click **Transformation** to show the contents of the directory.
 2. Right-click the check mark next to the desired dataset include file and click **Make Local**. A check mark for this file appears in the **User** column.
 3. Right-click the newly created check mark and click **Open > from the workbench**. The configuration window appears.

You can also open a dataset include file from a **Transformation Dependency Maps**. For information about **Transformation Dependency Maps**, see *Reprocessing and Retransformation* on page 627.

4. Edit the parameters in the configuration file as appropriate. See *Log Processing Dataset Include Files* on page 559 or *Transformation Dataset Include Files* on page 567 for descriptions of the parameters.

When editing a dataset include file within a data workbench window, you can use shortcut keys for basic editing features, including cut (Ctrl+x), copy (Ctrl+c), paste (Ctrl+v), undo (Ctrl+z), redo (Ctrl+Shift+z), select section (click+drag), and select all (Ctrl+a). In addition, you can use the shortcuts to copy and paste text from one configuration file (.cfg) to another.

5. To save your changes, right-click (**modified**) at the top of the window and click **Save**.
6. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs. Reprocessing or retransformation of the data begins after synchronization of the dataset profile.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

Creating New Dataset Include Files

Navigation title:Creating New Dataset Include Files

Steps to create a new dataset include file.

You should create a new dataset include file to perform any of the following dataset configuration tasks:

- Specifying new fields of data to be passed from log processing to transformation.
- Defining transformations that do either of the following:
 - Update existing log fields.
 - Produce new fields that are to be passed from log processing to transformation or that are used to define extended dimensions.

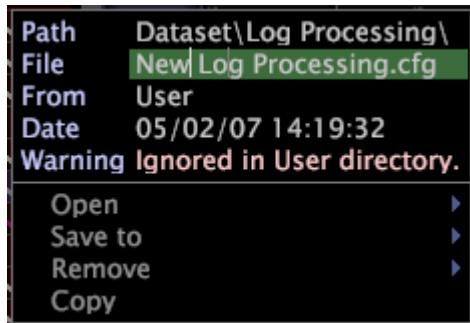
For information about the available transformation types, see *Data Transformations* on page 573.



Note: If you are defining transformations in a new dataset include file, be sure to keep the order of the inputs and outputs in mind. For information about the ordering of transformations, see *Conventions for Constructing Transformations* on page 574.

- Creating extended dimensions. For information about the available dimension types, see *Extended Dimensions* on page 606.
1. While working in your dataset profile, open the **Profile Manager** and click **Dataset** to view the existing dataset include files.

- To view the **Log Processing Dataset Include** files, click **Log Processing**.
 - To view the **Transformation Dataset Include** files, click **Transformation**.
2. Create a new **Log Processing** or **Transformation Dataset Include** files by performing one of the following steps:
 - In the **User** column for the Log Processing directory, click **Create > New Log Processing**. A file named **New Log Processing.cfg** appears in the directory.
 - In the **User** column for the Transformation directory, click **Create > New Transformation**. A file named **New Transformation.cfg** appears in the directory.
 3. Rename the new file by right-clicking its check mark in the **User** column and typing the new name in the **File** parameter.



4. Right-click the check mark for the renamed file and click **Open > from the workbench**. The configuration window appears.
5. Edit the parameters in the configuration file as appropriate. See *Log Processing Dataset Include Files* on page 559 or *Transformation Dataset Include Files* on page 567 for descriptions of the available parameters.
6. To save your changes, right-click (**modified**) at the top of the window and click **Save**.
7. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs. Reprocessing or retransformation of the data begins after synchronization of the dataset profile.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

To edit a dataset include file that you created, see *Editing Existing Dataset Include Files* on page 557.

Types of Dataset Include Files

Navigation title:Types of Dataset Include Files

Information about **Log Processing** and **Transformation dataset include** files.

- *Log Processing Dataset Include Files* on page 559
- *Transformation Dataset Include Files* on page 567

Log Processing Dataset Include Files

Navigation title:Log Processing Dataset Include Files

The **Log Processing Dataset Include** file for an inherited profile contains parameters associated with the log processing phase of dataset construction.

The first line of a **Log Processing Dataset Include** file defines a type **LogProcessingInclude** that supports the Decoder Groups, Fields, Log Entry Condition, Parameters, Reprocess, Stage, and Transformations parameters.

All other parameters for log processing must be defined in the `Log Processing.cfg` file in the dataset profile's Dataset directory. You can name a **Log Processing Dataset Include** file anything you want, but its file extension must be `.cfg`. The file must be stored within the *inherited profile name\Dataset\Log Processing* directory. Because the files are loaded recursively during the log processing phase of dataset construction, you can store the **Log Processing Dataset Include** files at any level within the directory (for example, *inherited profile name\Dataset\Log Processing\folder name\include file name.cfg*).



Note: Many web-specific configuration parameters for Site are defined in **Log Processing Dataset Include** files. For information about these parameters, see [Configuration Settings for Web Data](#) on page 649.

Log Processing Dataset Include File Parameters

Parameter	Description
Decoder Groups	<p>Required if you have defined log file or XML file log sources in the <code>Log Processing.cfg</code> file. The text file or XML decoders that you define to extract fields of data from log file and XML log sources.</p> <p>To add a new decoder group</p> <ul style="list-style-type: none"> Right-click Decoder Group and click Add new > TextFileDecoderGroup or XMLDecoderGroup. In the Name parameter for the new group, enter the desired name of the decoder group. <p> Note: When you specify a Decoder Group in the <code>Log Processing.cfg</code> file for the dataset profile, the name must match exactly the name that you enter here. For more information, see Log Files on page 534.</p> <p>For information about the decoders that you can define for each group, see Text File Decoder Groups on page 561 or XML Decoder Groups on page 563.</p>
Fields	<p>Lists fields that are defined in Log Sources or Transformations in a Log Processing Dataset Configuration file but used in transformations, conditions, or extended dimensions in a Transformation Dataset Configuration file must be listed here.</p> <p>Each field below must be listed in some Log Processing Dataset Include file:</p> <ul style="list-style-type: none"> x-trackingid x-timestamp
Log Entry Condition	<p>Optional. Defines the rules by which log entries are considered for inclusion in the dataset. See Log Entry Condition on page 546.</p> <p> Note: To be included in the dataset, a log entry must satisfy the Log Entry Condition in the <code>Log Processing.cfg</code> file and in every Log Processing Dataset Include file.</p>
Parameters	<p>Optional. A variable that you can reference in other configuration parameters. For more information, see Defining Parameters in Dataset Include Files on page 568.</p>
Reprocess	<p>Optional. Any character or combination of characters can be entered here. Changing this parameter and saving the file to the data workbench server initiates data reprocessing.</p> <p>For information about reprocessing your data, see Reprocessing and Retransformation on page 627.</p>
Stage	<p>Optional. The name of the processing stage that applies to this Log Processing Dataset Include file. The processing stages are defined in the Stages parameter in the <code>Log Processing.cfg</code> file.</p> <p> Note: When you specify a Stage, the name of the Stage must match exactly the name that is listed in the Stages parameter in the <code>Log Processing.cfg</code> file for the dataset profile.</p>
Transformations	<p>Optional. Defines the data transformations that need to be applied during log processing. For information about the available transformation types, see Data Transformations on page 573.</p>



Note: For descriptions of the parameters in the `Log_Processing.cfg` file, see [Log Processing Configuration File](#) on page 528.

You should keep the following points in mind whenever you are working with **Log Processing Dataset Include** files:

- Changing any of the parameters in this file requires reprocessing of all of the data.
- You can add any of the parameters described above to the **Log Processing Dataset Include** file by opening and editing the file in Notepad. Any changes you make and save appear when you reopen the file in data workbench. When adding a new parameter, use the Space key (not the Tab key) to indent two (2) spaces to the right of the previous heading level.

Text File Decoder Groups

Navigation title:Text File Decoder Groups

The processing of log files as log sources requires the definition of a decoder within the **Log Processing Dataset Include** file to extract fields of data from the log entries.

Defining text file decoder groups for log file log sources requires knowledge of the log file's structure and contents, the data to be extracted, and the fields in which that data is stored. This section provides basic descriptions of the parameters that you can specify for decoders, but the manner in which you use any decoder depends on the log file that contains your source data.

For information about format requirements for log file log sources, see [Log Files](#) on page 534. For assistance with defining text file decoders, contact Adobe.

A text file decoder group can include:

- [Regular Expression Decoders](#) on page 561
- [Delimited Decoders](#) on page 562

Regular Expression Decoders

A regular expression decoder identifies complex string patterns within the log entries in a log file and extracts these patterns as fields of data. For each decoder, the number of fields must equal the number of capturing sub-patterns in the regular expression. The portion of the line matching the nth capturing sub-pattern is assigned to the nth field for that line.

To add a regular expression decoder to a text file decoder group

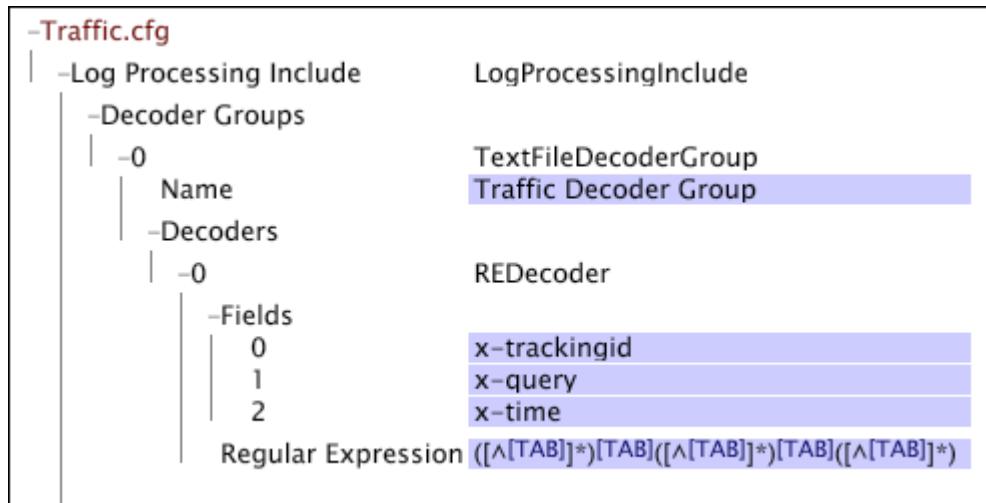
1. Open the **Log Processing Dataset Include** file as described in [Editing Existing Dataset Include Files](#) on page 557 and add a text file decoder group. See the table entry [Log Processing Dataset Include Files](#) on page 559.
2. Right click **Decoders** under the newly created decoder group, then click **Add new > Regular Expression**.
3. Specify the following information:
 - **Fields:** List of the fields in the log file. If any of the fields defined here are to be passed to the transformation phase of dataset construction, those fields must be listed in the `Fields` parameter of one of the **Log Processing Dataset Include** files for the dataset. Custom field names must begin with "x-".
 - **Name:** Optional identifier for the decoder.
 - **Regular Expression:** Used to extract the desired fields from each line in the file.
4. Repeat steps 4 and 5 for any other decoders that you want to add to the group.
5. To save the **Log Processing Dataset Include** file, right-click (**modified**) at the top of the window and click **Save**.
6. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column. Click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs.

Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.



Note: A given log file can have multiple regular expression decoders. The order in which you define the decoders is important: the first decoder to match a line in the log file is the one used to decode that line.

This example illustrates the use of a regular expression decoder to extract fields of data from a tab-delimited text file. You can achieve the same result by defining a delimited decoder with a tab delimiter.



For more information about regular expression decoders, including terminology and syntax, see [Regular Expressions](#) on page 642.

Delimited Decoders

A delimited decoder decodes a log file whose fields are delimited by a single character. The number of fields must correspond to the number of columns in the delimited file; however, not all fields need to be named. If a field is left blank, the column is still required in the log file, but the decoder ignores it.

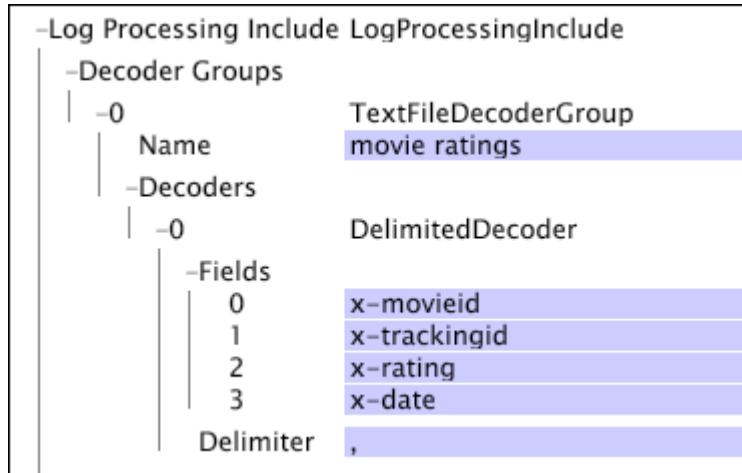
To add a delimited decoder to a text file decoder group

1. Open the **Log Processing Dataset Include** file as described in [Editing Existing Dataset Include Files](#) on page 557 and add a text file decoder group. See the table entry [Log Processing Dataset Include Files](#) on page 559.
2. Right click **Decoders** under the newly created decoder group, then click **Add new > Delimited**.
3. Specify the following information:
 - **Fields:** List of the fields in the log file. If any of the fields defined here are to be passed to the transformation phase of dataset construction, those fields must be listed in the **Fields** parameter of one of the **Log Processing Dataset Include** files for the dataset. Custom field names must begin with "x-".
 - **Delimiter:** Character that is used to separate fields in the output file.
4. Repeat steps 4 and 5 for any other decoders that you want to add to the group.
5. To save the **Log Processing Dataset Include** file, right-click (**modified**) at the top of the window and click **Save**.
6. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

This example illustrates the use of a delimited decoder to extract fields of data from a comma-delimited text file containing data about movies.



XML Decoder Groups

Navigation title: XML Decoder Groups

The processing of XML files as log sources to define decoders for extracting data from the XML file.



Note: Defining XML decoder groups for XML log sources requires knowledge of the XML file's structure and contents, the data to be extracted, and the fields in which that data is stored. This section provides basic descriptions of the parameters that you can specify for decoders. The manner in which you use any decoder depends on the XML file that contains your source data.

For information about format requirements for XML log sources, see [Log Sources](#) on page 532. For assistance with defining XML decoders, contact Adobe.

The top level of an XML decoder is a decoder group (XMLDecoderGroup), which is a set of decoder tables that you use to extract data from an XML file of a particular format. If you have XML files of different formats, then you must define a decoder group for each format. Each decoder group consists of one or more decoder tables.

The following table describes the Tables parameter and all of the sub-parameters that you must specify to define an XML decoder group.

XMLDecoderGroup

Parameter	Description
Tables	<p>Each table in a decoder group represents one level of data to be extracted from the XML file. For example, if you want to extract data about visitors, then you would create a decoder table that consists of the information you want to extract for each visitor. You also can create decoder tables within decoder tables (see Children).</p> <p>To add a table to a decoder group</p> <ul style="list-style-type: none">Right-click Tables and click Add new > XMLDecoderTable.
Fields	<p>The extended fields (for example, x-trackingid, x-email) in which the data is stored. The data to be stored in the field is determined by the Path and/or Operation subfields.</p> <p>The Path is the field's level within the structured XML file. A field's path is relative to the path of the table in which it is defined. Examples include <code>tag.tag.tag</code> or <code>tag.tag.tag.@attribute</code>. Note that paths are case-sensitive.</p>

Parameter	Description
	<p>An Operation is applied to each line in the specified path to produce an output. The following operations are available:</p> <ul style="list-style-type: none"> • LAST: The field takes the value of the path's last occurrence in the XML file. • RANDOM: Assigns a random value to the field. This operation is useful if you need to generate a unique id, such as for the x-trackingid field. • INHERIT: The defined field inherits its value from the parent table's corresponding field. • "constant": The constant must be enclosed in quotation marks. You can use a constant operation to check for the existence of a particular path; if the path exists, then the field is assigned the constant's value. <p>To add a field to a decoder table</p> <ul style="list-style-type: none"> • Right-click Fields, then click Add new > XMLDecoderField. Define Field, Operation and Path as appropriate.
Path	<p>The level within the structured XML file for which the decoder table contains information. For a child XML decoder table, the path is relative to the parent table's path. Note that paths are case-sensitive.</p> <p>For example, if your XML file contains the structure:</p> <pre><logdata> <visitor> ... </visitor> </logdata></pre> <p>then the path would be <code>logdata.visitor</code>.</p>
Table	<p>The value of this parameter should always be "Log Entry."</p> <p> Note: Do not change this value without consulting Adobe.</p>
Children	<p>Optional. One or more embedded decoder tables. Each child includes the Fields, Path, and Table parameters described above.</p> <p>To add a child to a decoder table</p> <ul style="list-style-type: none"> • Right-click Children and click Add new > XMLDecoderTable. Define Field, Operation and Path as appropriate.

To use an XML file as a log source for a dataset, XML decoder groups and tables must be defined to extract the information that is to be processed into the dataset. In this example, you can see how to define decoder groups and tables for a sample XML log source for a web dataset.

The following XML file contains information about a website visitor, including a visitor ID, email address, physical address, and information about the visitor's page views.

```
<?xml version="1.0" encoding="utf-8" ?>
- <logdata>
  - <visitor id="1">
    - <contact>
      <email>foo@bar.com</email>
      <street>123 Nowhere Street</street>
      <city>Nowheresville</city>
      <state>AK</state>
    </contact>
    <registered />
  - <pageview>
    <timestamp>2006-01-01 08:00:00</timestamp>
    <uri>/index.html</uri>
  </pageview>
  - <pageview>
    <timestamp>2006-01-01 08:00:30</timestamp>
    <uri>/</uri>
  </pageview>
</visitor>
</logdata>
```

Since we have a single XML file, we need only one decoder group, which we name "Sample XML Format." This decoder group applies to any other XML files of the same format as this file. To begin constructing XML decoder tables within this decoder group, we must first determine what information we want to extract and the fields in which the data will be stored.

In this example, we extract information about the visitor and the page views associated with that visitor. To do this, we create a top-level (parent) XML decoder table with information about the visitor and an embedded (child) XML decoder table with information about that visitor's page views.

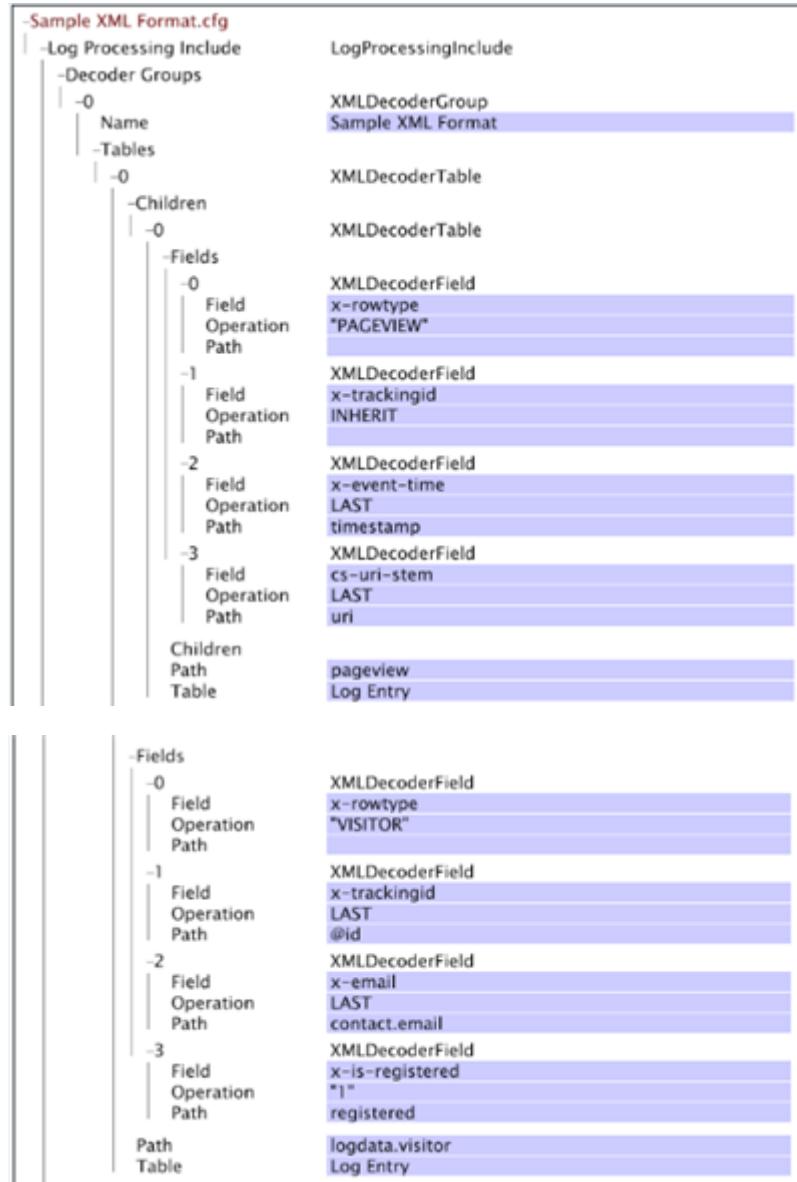
Information for the parent (visitor) table is as follows

- A data type identifier for each row of data in the XML file. We use VISITOR as our identifier so that we can quickly identify rows of data pertaining to the visitor and not to the page views. We can store this value in the x-rowtype field.
- The visitor's ID, which we store in the x-trackingid field.
- The visitor's email address (contact.email), which we store in the x-email field.
- The visitor's registration status. If the visitor is a registered user, then we can store the value "1" in the x-is-registered field.
- The Path value is logdata.visitor, and the Table value is **Log Entry**. For information about these parameters, see the XMLDecoderGroup table above.

Information for the child (page views) table is as follows:

- A data type identifier for each row of data in the XML file. We use "PAGEVIEW" as our identifier so that we can quickly identify rows of data pertaining to the visitor's page views and not to the visitor only. We store this value in the x-rowtype field.
- The visitor's ID. This value is inherited from the parent table and is stored in the x-trackingid field.
- The timestamp of each page view, which is stored in the x-event-time field.
- The URI of each page view, which is stored in the cs-uri-stem field.
- The Path value is pageview, and the Table value is "Log Entry." For information about these parameters, see the XMLDecoderGroup table above.

The following screen capture shows a portion of **Log Processing Dataset Include** file with the resulting XML decoder group for the sample XML file based on the discussed structure of the parent and child XML decoder tables.



A table showing the output of this decoder for our sample XML file looks something like the following:

x-rowtype	cs-uri-stem	x-email	x-is-registered	x-event-time	x-tracking-id
VISITOR		foo@bar.com	1		1
PAGEVIEW	/index.html			2006-01-01 08:00:00	1
PAGEVIEW	/			2006-01-01 08:00:30	1

You can create a table like the one above in data workbench by using a field viewer interface. For information about the field viewer interface, see [Dataset Configuration Tools](#) on page 628.

Using #value on XML element to read its attribute value

You can now use the **#value** tag in XML paths to pull the value of an XML element.

For example, previously specifying a path of `<Hit><Page name="Home Page" index="20">home.html</Page></Hit>` left you unable to read the value of the `<Page>` tag. To read the value of a `<Page>` tag and its attributes, you can use `Hit.Page.@name` and `Hit.Page.@index` respectively.

You can also pull the value of the tag using `Hit.Page.#value` expression.

For example, you can read the value of tag `<varValue>` by adding following field in decoder:

```
7 = XMLDecoderField:  
Field = string: x-varvalue-name-added  
Operation = string: LAST  
Path = string: #value  
Path = string: varValue  
Table = string: Log Entry
```

Similarly, you can read the value of tag `<Rep>` by adding following field in decoder:

```
7 = XMLDecoderField:  
Field = string: x-rep-name-added  
Operation = string: LAST  
Path = string: Rep.#value  
Path = string: Reps  
Table = string: Log Entry
```

In contrast, to read the value of element tag with no attribute, a `<text>` tag under a `<line>` tag and its value can be read directly by giving "text" in a path or using `line.text`, depending on how you have built the decoder.

```
2 = XMLDecoderField:  
Field = string: x-chat-text  
Operation = string: LAST  
Path = string: text  
Path = string: line  
Table = string: Log Entry
```

Transformation Dataset Include Files

Navigation title: Transformation Dataset Include Files

The **Transformation Dataset Include** file for an inherited profile contains parameters associated with the transformation phase of dataset construction.

The first line of the file defines a type **TransformationInclude** that supports the Extended Dimensions, Parameters, Reprocess, Stage, and Transformations parameters. All other parameters must be defined in the `Transformation.cfg` file in the dataset profile's Dataset directory.

Including parameters other than Extended Dimensions, Parameters, Reprocess, Stage, and Transformations in a **Transformation Dataset Include** file generates errors.

You can name a **Transformation Dataset Include** file anything you want, but its file extension must be `.cfg`. The file must be stored within the *inherited profile name\Dataset\Transformation* directory. Because the files are loaded recursively during the transformation phase of dataset construction, you can store the **Transformation Dataset Include** files at any level within the directory (for example, *inherited profile name\Dataset\Transformation\folder name\include file name.cfg*).



Note: Many web-specific configuration parameters for Site are defined in **Transformation Dataset Include** files. For information about these parameters, see [Configuration Settings for Web Data](#) on page 649.

The following table describes the parameters that are available in a **Transformation Dataset Include** file:

Transformation Dataset Include File Parameters

Parameter	Description
Extended Dimensions	Optional. Defines the extended dimensions. See Extended Dimensions on page 606.
Parameters	Optional. A variable that you can reference in other configuration parameters. For more information, see Defining Parameters in Dataset Include Files on page 568.
Reprocess	Optional. Any character or combination of characters can be entered here. Changing this parameter and saving the file initiates data retransformation. For information about reprocessing your data, see Reprocessing and Retransformation on page 627.
Stage	Optional. The name of the processing stage that applies to this Transformation Dataset Include file. The processing stages are defined in the Stages parameter in the Transformation.cfg file.  Note: When you specify a Stage, the name of the Stage must match exactly the name that is listed in the Stages parameter in the Transformation.cfg file for the dataset profile.
Transformations	Optional. Defines the data transformations that need to be applied during transformation. For information about the available transformation types, see Data Transformations on page 573.



Note: For descriptions of the parameters in the Transformation.cfg file, see [Transformation Configuration File](#) on page 551.

You should keep the following points in mind whenever you are working with **Transformation Dataset Include** files:

- Changing any of the parameters in this file requires retransformation of the data.
- **CrossRows**, **ODBCLookup**, **Sessionize**, and **AppendURI** transformations work only when defined in a **Transformation Dataset Configuration** file. For information about these transformations, see [Data Transformations](#) on page 573.
- You can add any of the parameters described above to the **Transformation Dataset Include** file by opening and editing the file in Notepad. Any changes you make and save appear when you reopen the file in data workbench. When adding a new parameter, use the Space key (not the Tab key) to indent two (2) spaces to the right of the previous heading level.

If you subscribe to Adobe's **IP Geo-location** or **IP Geo-intelligence** data service, Adobe provides you with an internal profile consisting of a set of data transformations and extended dimensions that are created specifically for the data service. The transformations and dimensions are defined in **Transformation Dataset Include** files that are included in the Dataset directory of the internal profile. For instructions to install the internal profile for the **IP Geo-location** or **IP Geo-intelligence** data service, see the *Data Workbench User Guide*.

Defining Parameters in Dataset Include Files

Navigation title:Defining Parameters in Dataset Include Files

When configuring your dataset, you can define variables, referred as parameters, to represent meaningful values.

To assign a value to a parameter (that is, to define the parameter), you add the parameter's name and value to the Parameters vector in a log processing or **Transformation Dataset Include** file. After you define parameters, you can reference them in your dataset profile's configuration files. Defining and referencing such parameters is referred to as parameter substitution. Using parameter substitution when configuring your dataset enables you to create a centralized location for your parameter definitions. When you need to update a parameter that is referenced multiple times or in multiple files, you have to make the change only once.



Note: In this guide, the term parameter has been used to refer to the name of any setting in a configuration file (such as Log Entry Condition, Reprocess, or Transformations). However, as used in this section,

parameter refers specifically to a member of the Parameters vector in a dataset include file and not to the name of a setting in a configuration file.

You should consider the following points when defining a parameter:

- A parameter must be defined exactly once. Therefore, you cannot define the same variable in multiple dataset include files.
- Any parameter that you define is local to either the log processing or the transformation phases, but it is global across multiple dataset configuration files for that phase. For example, if you define a parameter in a **Transformation Dataset Include** file, the parameter is defined for the entire transformation phase, and you can reference it in the Transformation.cfg file and all other **Transformation Dataset Include** files for the inherited profiles. The parameter would not be defined for log processing; therefore, any references to the parameter in the Log Processing.cfg file or a **Log Processing Dataset Include** file would generate a processing error.

To define a parameter

You can define string, numeric, and vector parameters in **Log Processing** and **Transformation Include** files.

1. In the data workbench window for the **Log Processing** or **Transformation Dataset Include** file, right-click **Parameters**, then click **Add new > Parameter**.
2. Select **String Parameter**, **Numeric Parameter**, or **Vector Parameter**, and complete the Name and Value parameters as described in the following sections.
3. To save the dataset include file in which you have defined the parameter, right-click (**modified**) at the top of the window and click **Save**.
4. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

To reference a parameter

- When you reference a defined parameter in another dataset configuration file, you must type its name as `$ (parameter name)`.

The following sections describe the types of parameters that you can define.

- [String and Numeric Parameters](#) on page 569
- [Vector Parameters](#) on page 570

String and Numeric Parameters

Navigation title: String and Numeric Parameters

String and numeric parameters take as their values strings and numbers, respectively.

You can use them interchangeably, but numeric parameters must be defined to have a numerical value. You can reference string and numeric parameters when defining transformations, conditions, and extended dimensions, and you can reference more than one parameter in the same line.

You cannot reference string and numeric parameters in **Input** or **Output** fields, but you can use a string parameter to define a constant input field. In addition, you cannot reference string and numeric parameters in decoders or decoder groups.

This example shows a **Log Processing Dataset Include** file that defines a string parameter and a numeric parameter. Note that the string parameter, named "Value Lookups," defines a file location (Lookups\Values) relative to the data workbench server installation directory.

-Log Processing Parameters.cfg (modified)	
-Log Processing Include LogProcessingInclude	
-Parameters	
-0	StringParameterDef
-Name	Value Lookups
-Value	Lookups\Values
-1	NumericParameterDef
-Name	Maximum Count
-Value	32768

Vector Parameters

Navigation title:Vector Parameters

Vector parameters contain multiple values for a single variable.

You can reference vector parameters only as the sole item of a vector. This example shows a **Transformation Dataset Include** file that defines a vector parameter. The vector parameter, "Internal Domains," consists of three values.

-Internal Domains.cfg	
-Transformation Include TransformationInclude	
-Parameters	
-0	VectorParameterDef
-Name	Internal Domains
-Value	
0	abcsite.com
1	123site.com
2	xyzsite.com

Note that the vector parameter is the only item listed for the **Matches** vector in the **String Match** condition.

-Condition	AndCondition
-0	StringCondition
-Matches	
0	\$(Internal Domains)
-Input	cs(referrer-domain)

For more information about internal domains, see [Configuration Settings for Web Data](#) on page 649. For information about the **String Match** condition, see [Conditions](#) on page 634.

Additional Configuration Files

Navigation title:Additional Configuration Files

Information about some of the additional configuration files contained in the Dataset directory for the dataset profile or its inherited profiles.

Although these files are not always required to define a dataset profile, you can edit them to specify additional parameters for processing the data.

Log Processing Mode.cfg

Navigation title: Log Processing Mode.cfg

The configuration file `Log Processing Mode.cfg` enables you to pause processing of data into a dataset, specify offline sources, or specify the frequency at which data workbench server saves its state files.

Making changes to the `Log Processing Mode.cfg` file, including adding or removing sources, does not cause reprocessing of the data.

To edit the `Log Processing Mode.cfg` file for a dataset profile

1. While working in your dataset profile, open the **Profile Manager** and click **Dataset** to show its contents.



Note: If the `Log Processing Mode.cfg` file is not located in the directory for the desired profile, you need to copy this file from the Base directory on the data workbench server machine into the profile's directory.

For information about opening and working with the **Profile Manager**, see the *Data Workbench User Guide*.

2. Right-click the check mark next to the configuration file's name and click **Make Local**. A check mark for this file appears in the **User** column.
3. Right-click the newly created check mark and click **Open > from the workbench**. The configuration window appears.
4. Edit the parameters in the configuration file using the following table as a guide.



Note: Some of the parameters in the `Log Processing Mode.cfg` file have names that include **Fast Input** or **Fast Merge**. **Fast Input** refers to the log processing phase of dataset construction and is responsible for approximately half of the total dataset processing time. **Fast Merge** refers to the transformation phase of dataset construction only when preceded by log processing. **Fast Merge** does not occur during retransformation that results from modifying a **Transformation Dataset Configuration** file. Like **Fast Input**, **Fast Merge** is also responsible for approximately half of the dataset processing time.

Log Processing Mode.cfg

Parameter	Description
Cloud Bytes	A tuning parameter that affects the efficiency of data transformation. The default value is 128000000.  Note: You should not change this value without consulting Adobe.
Fast Input Decision Ratio	A tuning parameter that specifies the ratio of total to unread log bytes at which the system enters Fast Input mode (and subsequently Fast Merge) instead of processing data in real time. The default value is 200, meaning that the system enters Fast Input mode from real-time mode when the unread log data is at 1/200th of the total data. A higher decision ratio makes the system enter Fast Input mode more readily, while a lower ratio makes it less likely to enter Fast Input mode.  Note: Setting the parameter to 0 prevents the system from entering Fast Input mode at all, even for initial processing. Setting the parameter to 1.1 enables the system to enter Fast Input during initial processing but not for subsequent processing. Adobe does not recommend using values between 0 and 1.1. For more information about setting this parameter, contact Adobe.
Fast Input FIFO Bytes	A tuning parameter that balances memory usage and system performance during data processing. The default value is 120000000.

Parameter	Description
	 Note: You should not change this value without consulting Adobe.
Fast Merge Buffer Bytes	A tuning parameter that balances memory usage and system performance during data processing. The default value is 128000000.  Note: You should not change this value without consulting Adobe.
Offline Sources	Mask of the offline log source. To specify an offline source <ul style="list-style-type: none"> • Right-click Offline Sources, then click Add new > Source. • In the parameter for the new source, enter the mask of the log sequence. For Sensor log sources with file names of the format YYYYMMDD-SENSORID.vsl, the mask is <i>SENSORID.SENSORID</i> is case-sensitive. For log file log sources, the mask is the string extracted by the Mask Pattern. See Log Files on page 534. <p>Adding or removing sources from Offline Sources does not cause reprocessing of the dataset.</p> <p>As Of time measurements are maintained for the processing of the profile's online sources. When the offline source is again online, the processing of incoming log files for that source resumes.</p> <p>Whenever a source comes back online, you should remove it from Offline Sources. If you do not do so, data workbench server treats the source as an online source and updates the As Of time as long as the source is sending data. If the source goes offline again, the As Of time measurements stop.</p>
Paused	True or false. If true, new data is not processed into the dataset. The default value is false.
Real Time Delay	The amount of time in seconds that data workbench Server waits between intervals of processing data into the dataset. When this value is set to zero, the system attempts to keep up with incoming data in real time. The default value is zero (0), but you can increase this value to reduce CPU load.
Real Time FIFO Bytes	The amount of memory in bytes used to store data that is waiting to be processed into the dataset. You may need to change this value based on the number of seconds that you specify for Real Time Delay. The default value is 16000000.  Note: You should not change this value without consulting Adobe.
Save Interval (sec)	Frequency at which the data workbench server saves its state files. The default value is 3600.  Note: You should not change this value without consulting Adobe.

When editing the `Log Processing Mode.cfg` file within a data workbench window, you can use shortcut keys for basic editing features, including cut (Ctrl+x), copy (Ctrl+c), paste (Ctrl+v), undo (Ctrl+z), redo (Ctrl+Shift+z), select section (click+drag), and select all (Ctrl+a). In addition, you can use the shortcuts to copy and paste text from one configuration file (.cfg) to another.

5. Right-click (**modified**) at the top of the window and click **Save**.
6. In the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <datasetprofile name>**.

 **Note:** Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

Server.cfg

Navigation title:Server.cfg

The Sample Bytes parameter in the `Server.cfg` file specifies the data cache size (in bytes) for data workbench.

The default value is 250e6. Instructions for opening and saving the `Server.cfg` file are the same as those for Log Processing Mode.`cfg`. See [Log Processing Mode.`cfg`](#) on page 571.



Note: Because this file's parameter affects system performance, please contact Adobe before making any changes.

You can further limit the size of the data cache for data workbench machine that connect to the data workbench server by setting the Maximum Sample Size parameter in the `Insight.cfg` file. For more information, see the *Data Workbench User Guide*.

Other Files

Navigation title:Other Files

The `Dataset` directory includes additional files that are either required for the operation of the software or provide additional functionality for your Adobe implementation.

- **Client.cfg:** The `Client.cfg` file within the `Dataset` directory for the **Base** profile is required for operation of the software. Do not delete or modify any of the parameters in the `Client.cfg` file.
- **Cluster.cfg:** The `Cluster.cfg` file within the `Dataset` directory for the **Base** profile is required for operation of the software. In the `Cluster.cfg` file, you should modify only the Normalize Server parameter if you are configuring a dataset to be processed on an data workbench server cluster. For instructions to modify the Normalize Server parameter, see [Creating a Centralized Normalization Server for a Cluster](#) on page 544.
- **Insight Transform.cfg and Insight Transform Mode.cfg:** If you are using transformation functionality, you have two additional configuration files, data workbench `Transform.cfg` and data workbench `TransformMode.cfg`, in the `Dataset` directory for the **Transform** profile. For information about these files and their parameters, see [Transform Functionality](#) on page 617.
- The `PAServer.cfg` file. If you want to submit Predictive Analytics clustering jobs to Insight Servers, then you will need to configure the `PAServer.cfg` file for handling server-side clustering submissions.

The custom profile should inherit the `PAServer.cfg` from the Predictive Analytics profile (`Server\Profiles\Predictive Analytics\Dataset`).



Important: Set a *Master Server* in this file and save the `PAServer.cfg` to the implementation site.

```
PAServer = PAServerConfig:  
    Master Server = serverInfo:  
        Address = string:  
        Port = int: 80  
        Use SSL = bool: false
```

Data Transformations

Navigation title:Data Transformations

Overview of data transformations, describes each of the available transformation types, and describes how to use the transformations.

About Transformations

Navigation title:About Transformations

Transformations enable you to extract information available in your data files and manipulate it into a more useful form.

Transformations operate on the log entries (you can think of log entries as rows of data) in your log sources. For each row of data, the transformation takes the value of the specified input field, performs a set of processing steps, and records the result in the output field that you specify. You can define transformations to be executed during either the log processing or transformation phase of the dataset construction process:

- **During log processing:** Transformations executed during the log processing phase of dataset construction are applied to each event data record (log entry) to update existing log fields or produce new fields. The results of the transformations are used along with log entry conditions to evaluate which log entries are filtered out of the dataset during log processing.
- **During transformation:** Transformations executed during the transformation phase of dataset construction operate on the fields of data passed from log processing to create extended dimensions that you can use in your analyses. See [Extended Dimensions](#) on page 606.



Note: The data input to transformation from log processing is ordered by time and grouped by the tracking ID in your source data. Several transformations require that the data is in this form and work only when defined in during transformation.

Changes to transformations must be made with care. Transformations do not affect which log entries flow into the dataset construction process, but they do affect the results presented. This permits changes to be made in what is being analyzed without changing the data upon which the analysis is based. However, changes in transformations can fundamentally alter the values produced in analyses.

Conventions for Constructing Transformations

Navigation title:Conventions for Constructing Transformations

Table showing what conventions apply when constructing transformations.

Transformation Conventions

Convention	Description
Sequential execution	<p>The transformations within a dataset configuration file are applied to the log entries sequentially (that is, in the order in which they are listed in the configuration file). Therefore, transformations must be listed in the order their outputs are used as inputs to other transformations. More specifically, if the output of one transformation is used as the input to another transformation, it is important for that former transformation to be listed prior to the latter transformation in the dataset configuration files. Otherwise, the data workbench server generates an error.</p> <p>Processing stages provide a way to order the transformations that are defined within multiple dataset include files. For all of the dataset include files associated with a particular processing stage, transformations are ordered based on their inputs and outputs. In addition, if multiple dataset include files within a stage output data to the same field as a result of a transformation, the data workbench server generates an error.</p> <p>For more information about stages, see Log Processing Configuration File on page 528, Transformation Configuration File on page 551, and Dataset Include Files on page 557.</p> <p>A Transformation Dependency Map can display how a field is modified by a series of transformations. See Dataset Configuration Tools on page 628.</p>
Output names	Most transformations specify an output field. If the output is a user-defined extended field, the name for this field must start with "x-." The output field names cannot contain spaces or special characters. The names of extended fields can be written with mixed-case, such as "x-NewCampaignName," or "x-New-Campaign-Name" for readability, but they are treated by the software as case-insensitive.
Input fields	<p>Input fields refer to one of the baseline fields or a user-created field resulting from the output of a previous transformation. If a constant string is needed, a quoted string can be used instead of a baseline or user-created field.</p> <p>For a list of some of the commonly defined fields of data that the data workbench server can process, see Event Data Record Fields on page 646.</p>
Simple strings and vectors of strings	All transformations operate on strings and/or vectors of strings. Simple strings are literal sequences of characters. String vectors contain zero or more simple strings in a specific order.

Defining a Transformation

Navigation title:Defining a Transformation

You can define data transformations to be applied during either the log processing or transformation phase of dataset construction.



Note: Adobe recommends defining transformations in either **Log Processing** or **Transformation Dataset Include** files instead of in Log_Processing.cfg or Transformation.cfg.

The following transformations work only when defined in the Transformation.cfg file or in a **Transformation Dataset Include** file:

- [AppendURI](#) on page 595I
- [CrossRows](#) on page 577
- [LookupRows](#) on page 584
- [ODBC Data Sources](#) on page 547
- [Sessionize](#) on page 591

To define a transformation

1. Use the **Profile Manager** to open the dataset configuration file in which you want to define the transformation.
 - (Recommended) To open a dataset include file, see [Dataset Include Files](#) on page 557.
 - To open the Log_Processing.cfg file, see [Editing the Log Processing Configuration File](#) on page 529.
 - To open the Transformation.cfg file, see [Editing the Transformation Configuration File](#) on page 551.
2. Right-click **Transformations**, then click **Add new > <Transformation type>**.
3. Input the appropriate information for your transformation. For descriptions of the transformation types and information about their parameters, see the following sections:
 - [Standard Transformations](#) on page 576
 - [URI Transformations](#) on page 595
 - [Integrating Lookup Data](#) on page 599
4. After you have defined your transformation(s) in the configuration file, save the file locally and save it to your dataset profile on the data workbench server.

Tips for defining and editing transformations:

- When editing the configuration of a transformation within a data workbench window, you can use shortcut keys for basic editing features, including cut (Ctrl+x), copy (Ctrl+c), paste (Ctrl+v), undo (Ctrl+z), redo (Ctrl+Shift+z), select section (click+drag), and select all (Ctrl+a). In addition, you can use the shortcuts to copy and paste text or entire transformation definitions from one configuration file (.cfg) to another.
- For any transformation that you define, you can add one or more comment lines to the **Comments** parameter to further describe the transformation or add notes about its use. To add a comment using data workbench, right-click the **Comments** label, then click **Add new > Comment Line**.
- You can open the configuration of any transformation from a **Transformation Dependency Map**. After you open the configuration, you can edit it and save your changes. For information about **Transformation Dependency Maps**, see [Dataset Configuration Tools](#) on page 628.
- An empty string output from a transformation can overwrite a non-empty string in the output field.

Types of Transformations

Navigation title:Types of Transformations

Insight Server provides standard and URI transformations.

- [Standard Transformations](#) on page 576
- [URI Transformations](#) on page 595

Standard Transformations

Navigation title:Standard Transformations

A standard transformation takes a set of inputs, performs some operation, then provides that result in one or more output fields.

The following sections provide a description, a table of available parameters, and an example for each of the standard transformations. Every data workbench server DPU provides all of the standard transformations.

ChangeCase

Navigation title:ChangeCase

The **ChangeCase** transformation changes the case of the string in the Input parameter as specified by the **Action** parameter.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Action	Upper or lower. Specifies whether the case is to be changed to upper or lower.	lower
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Input	The name of the field from the log entry to use as input.	
Output	The name of the output field.	

In this example, which uses fields of data collected from website traffic, the case of the string within the s-dns field is changed to lower case, and the new value is output in the new field, x-lowercase-dns.

-0	ChangeCase
Name	ChangeHostCase
-Comments	Comment
0	Change host to lower case
Action	lower
Condition	AndCondition
Input	s-dns
Output	x-lowercase

[Copy](#)

Navigation title:Copy

The **Copy** transformation simply copies the value in the input field to the given output field. If the input field could be a vector of strings, the output field must start with "x-."

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	Used if the condition test is true and the input value is not available in the given log entry.	
Input	The name of the field from which to copy.	
Output	The name of the output field.	

In this example, which uses fields of data collected from website traffic, the output field, x-purchase-success, is given the literal value of "1" each time cs-uri-stem matches /checkout/confirmed.php. If the **Condition** is not satisfied (that is, cs-uri-stem does not match /checkout/confirmed.php), x-purchase-success is not changed.

-1	Copy
Name	Confirmed Purchase
-Condition	AndCondition
-0	StringCondition
-Matches	
0	/checkout/confirmed.php
Case Sensitive	true
Comments	Comment
Input	cs-uri-stem
Comments	Comment
Default	
Input	"1"
Output	x-purchase-success

CrossRows

Navigation title:CrossRows

Like other transformations, the **CrossRows** transformation is applied to the rows of data (log entries) in your log sources.

For each row of data, the transformation takes the value of the specified input field, performs a set of processing steps, and records the result in the output field that you specify. However, when the **CrossRows** transformation works on one row of data (this row is called the output row), it takes into account that row plus one or more other rows of data (these rows are called input rows) that are associated with the same tracking ID. Therefore, for a given tracking ID, the value of the output field for each output row is based on the values of the input field for one or more input rows.

The transformation provides multiple conditions and constraints that enable you to limit the input rows for the transformation. You can express these limits in terms of the data workbench server's conditions (see [Conditions](#) on page 634), a range of input rows relative to the output row, or a range of times relative to the time of the output row. For those input rows that satisfy the transformation's conditions and constraints, you can apply an operation (such as SUM) that determines the value of output field.



Note: To work, the **CrossRows** transformation requires that the data is ordered in time and grouped by the tracking ID in your source data. Therefore, **CrossRows** works only when defined in the Transformation.cfg file or in a **Transformation Dataset Include** file.

As you review the descriptions of the parameters in the following table, remember the following:

- The output row is the row of data that the transformation is working on at a given point in time.
- Input rows are all of the other rows of data (before, after, or including the output row) whose values of the input field serve as inputs to the transformation. Input rows are subject to the Input Condition, Key, Row Begin, Row End, Time Begin, and Time End parameters.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	

Parameter	Description	Default
Condition	Limits the output of the transformation to certain log entries. If the condition is not met for a particular log entry, the field in Output parameter is left unchanged. The input still may be used to affect other log entries.	
Input	The name of the field from the input row to use as input.	
Input Condition	Accepts input for the transformation from only certain input rows. If the Input Condition is not met for a particular input row, the input field from that row is ignored and does not affect other output rows. However, the output field from that row is still modified per the specified Condition.	
Key	<p>Optional. The name of the field to use as the key.</p> <p>If a key is specified, the input rows for a given output row are limited to the contiguous block of rows having the same Key value as the output row. This restriction is in addition to all other limitations placed on the input rows by other parameters of the CrossRows transformation.</p> <p>For example, if you are working with web data and you make the field x-session-key (which has a unique value for each session) the key, then the input rows for the transformation are limited to those rows having the same x-session-key value as the output row. Therefore, you are considering only those input rows representing page views that occur during the same session as the output row.</p>	
Operation	An operation that, for each output row, is applied to all of the input rows satisfying all of the conditions defined by the Input Condition, Key, Row Begin, Row End, Time Begin, and Time End parameters to produce an output: <ul style="list-style-type: none"> • ALL takes all of the values of the input field from the input rows and outputs them as a vector. • SUM interprets the values of the input field from the input rows as numbers and sums them. • FIRST ROW outputs the value of the input field from the first input row. • LAST ROW outputs the value of the input field from the last input row. 	
Output	The name of the output field.	
Row Begin/Row End	<p>Optional. Specifies a range of input rows relative to the output row. For example, a Row Begin value of "0" excludes all rows before the output row. A row begin value of "1" excludes the output row as well. Common ranges include:</p> <ul style="list-style-type: none"> • Begin 0: This row and all subsequent ones. • Begin 1: All subsequent rows. • End 0: This row and all previous ones. • End -1: All previous rows. • Begin -1, End -1: The previous row. • Begin 1, End 1: The next row. 	All rows
Time Begin/Time End	<p>Optional. Specifies a range of times relative to the time of the output row. For example, a Time End of 30 minutes includes all rows that take place within 30 minutes after the output row. A Time Begin of -30 minutes includes all rows that take place within 30 minutes before the output row.</p> <p>Available time units are days, weeks, hours, minutes, ms (milliseconds), ticks (100 nanoseconds), and ns (nanoseconds).</p>	All times

The **CrossRows** transformation in this example is applied to rows of web data to find for each page view the time of the next page view. Because we know that **CrossRows** is applied only during the transformation phase of the dataset construction process, the rows of data are ordered by visitor (each visitor has a unique tracking ID) and time.

The input field, x-timestamp, is considered for only those input rows in which the x-is-page-view field is populated (indicating the row of data represents a page view). The x-session-key field (which has a unique value for each session) is specified for the Key parameter. Therefore, the input rows (log entries) for the transformation are limited to the contiguous block of rows having the same value of x-session-key as the output row. In other words, to be considered for the transformation, an input row must represent a page view that occurs during the same session as the page view in the output row. The first row operation takes the value of the output field from the first input row satisfying the **Input** Condition and having the same x-session-key value as the output row.

-CrossRows (Time of next pageview)	
Name	Time of next pageview
-Input Condition	AndCondition
-0	NotEmptyCondition
Comments	Comment
Input	x-is-page-view
Comments	Comment
Condition	AndCondition
Default	
Input	x-timestamp
Key	x-session-key
Operation	first row
Output	x-next-timestamp
Row Begin	1
Row End	
Time Begin	
Time End	

CrossRows executes in an amount of time proportional to the size of its inputs plus the size of its outputs. This means that for operations SUM, FIRST ROW, and LAST ROW, it is no less efficient than other transformations. For ALL, the situation is more complex because it is possible to configure **CrossRows** to output an amount of data for each row of data (log entry) that is proportional to the total number of rows (log entries) for a given tracking ID.

ExtractValue

Navigation title:ExtractValue

If you are working with web data, you can use the **ExtractValue** transformation to extract a value from a query string, cookie, or similarly encoded field in your website data.

Note that the name(s) corresponding to the value to be extracted can be different in each log entry.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Input Name	The name(s) of the field(s) to be extracted from the Input Query. Note: If the Input Name is a vector (that is, there are multiple names present), only one value is extracted.	
Input Query	The encoded mapping (query string, cookie, and so forth) from which the value is to be extracted.	
Output Value	The name of the field used to capture the extracted decoded value.	

If you want to extract a search phrase, you can extract the entire phrase and, if desired, split the phrase into search terms using a **Tokenize** transformation. For information about the **Tokenize** transformation, see [Tokenize](#) on page 593.

This example configures an **ExtractValue** transformation to extract values of the x-v-search-querynames field from cs(referrer-query) and store them in the x-search-phrase field.

-ExtractValue (Search Phrase)	
Name	Search Phrase
Comments	Comment
Input Name	x-v-search-querynames
Input Query	cs(referrer-query)
Output Value	x-search-phrase

Flatten

Navigation title: Flatten

The **Flatten** transformation takes a vector of strings and maps each value into its own field.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition is met and the input value is not available for the log entry.	
Input	A vector of string values to map to the output field names.	
Outputs	A set of output field names.	

Considerations for Flatten

- If the input vector contains more values than there are defined output fields, the extra input values are simply dropped.
- If the input vector contains fewer values than there are defined output fields, the extra output fields are given the default value (if defined) or an empty string if no default value is defined.

Here, the **Flatten** transformation is used to take a vector of products (x-products) and separate them into four fields (x-product1, ..., x-product4).

-2	Flatten
Name	Products as Simple Strings
-Outputs	
0	x-product1
1	x-product2
2	x-product3
3	x-product4
Comments	Comment
Condition	AndCondition
Default	
Input	x-products

If the input value contained the strings B57481, C46355, and Z97123, the output fields would have the values shown here:

- x-product1 = B57481
- x-product2 = C46355
- x-product3 = Z97123
- x-product4 = Empty (There are more inputs than outputs, and there is no default value specified.)

Format

Navigation title: Format

The **Format** transformation takes a set of inputs and formats them to create an output matching the given structure.

The transformation works on either simple strings or string vectors and produces output by applying the given format to each input value until all of the input values have been transformed.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Format	A formatting string used to specify how the output will look. %1% refers to a value from the first input vector, %2% refers to a value from the second input vector, and so on.	
Inputs	Fields containing either simple strings or string vectors. In the case of string vectors as inputs, the output will also be a string vector resulting from the application of the Format parameter to each set of input values.  Note: The numbering of inputs starts at 0, but the numbering of the format substitution values starts at %1%.	
Output	The name of the field created to contain the results of the transformation. If the inputs are string vectors, the length of the output string vector will be the length of the longest input vector. If some of the input string vectors are of shorter length, empty strings are used for their position in the format string until the length of the output vector is reached.	

In this example, two vectors, one a vector of strings representing product categories and the other a corresponding string vector representing the quantity of each product purchased, are transformed into a single vector of equivalent length that takes the form: Product %1%, Quantity %2%.

-3	Format
Name	Format Product and Qty
-Inputs	
0	x-prod-category-id
1	x-prod-qty
Comments	Comment
Condition	AndCondition
Format	Product %1%, Quantity %2%
Output	x-product-and-quantity

If the input vectors contained product categories of (683, 918) and quantities of (10, 4), the result would be one final output vector containing the following two strings: ("Product 683, Quantity 10", "Product 918, Quantity 4").

Hash

Navigation title:Hash

The **Hash** transformation creates a nearly unique string representing a 64-bit number from the input values.

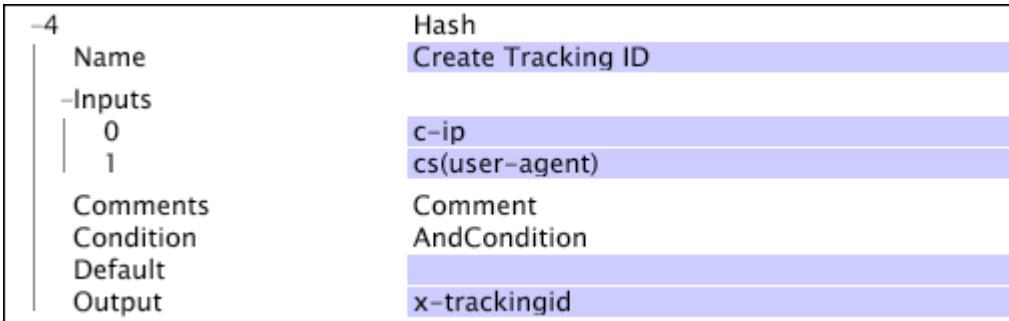
This transformation provides the same hash value when given the same inputs.

 **Note:** The resulting value is nearly unique because the transformation uses a 64-bit number as the space of possible hash values. For one million unique inputs to the **hash** transformation, there is a 1 in 38,000,000 chance of getting a duplicate hash value.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	

Parameter	Description	Default
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the input value is not available.	
Inputs	The set of inputs to use to create the hash value.	
Output	The name of the field for output.	

In this example, the values of the c-ip and cs(user-agent) fields are used to create a tracking ID, which is stored in the x-trackingid field.



-4	Hash
Name	Create Tracking ID
-Inputs	
0	c-ip
1	cs(user-agent)
Comments	Comment
Condition	AndCondition
Default	
Output	x-trackingid



Note: This example does not represent an ideal solution for creating unique tracking IDs. However, in situations in which archival log information is used, it may be the best method.

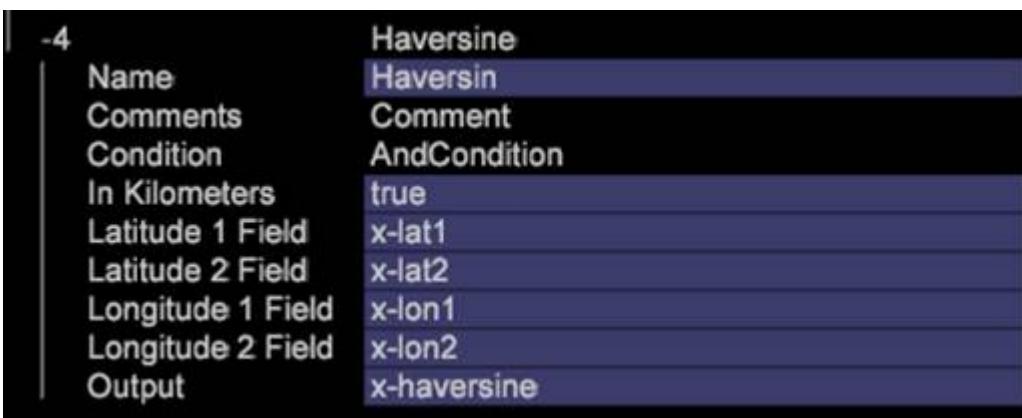
Haversine

Navigation title:Haversine

In mathematics, the haversine formula is an equation that gives circle distances between two points on a sphere identified from their longitudes and latitudes.

Like the formula, the **Haversine** transform requires two sets of **Latitude** and **Longitude** settings, using these four inputs to calculate the true distance across the Earth between two locations.

This distance can be represented as miles or kilometers by changing the "In Kilometers" flag from false to true.



-4	Haversine
Name	Haversin
Comments	Comment
Condition	AndCondition
In Kilometers	true
Latitude 1 Field	x-lat1
Latitude 2 Field	x-lat2
Longitude 1 Field	x-lon1
Longitude 2 Field	x-lon2
Output	x-haversine

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	

Parameter	Description	Default
Latitude 1 Field	The latitude of the point 1.	
Latitude 2 Field	The latitude of the point 2.	
Longitude 1 Field	The longitude of the point 1.	
Longitude 2 Field	The longitude of the point 2.	
Output	Once calculated, the Output field contains distances between the points designated as elements in a Dimension.	

As an example, if you code in a latitude and longitude of their store as Lat1, Lon1 and use an IP lookup lat and long for their customers, then distances to a store most customers buy from or come from can be determined.



Note: If you want to identify distances for other locations, then each individual location must have its own set of lat and lon fields.

IPLookup

Navigation title:IPLookup

The **IPLookup** transformation takes **IP geo-location** or **IP geo-intelligence** data (provided by any vendor of such data and converted into a proprietary format by Adobe) and transforms the data into geographical information that can be used in analysis.

Two **IPLookup** transformations are listed in the Add new > *Transformation type* menu:

- **IPLookup Quova** for **IP geo-location** data
- **IPLookup Digital Envoy** for **IP geo-intelligence** data

When defining an **IPLookup** transformation, choose the appropriate transformation for your **IP geo-location** or **IP geo-intelligence** data.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
File	Path and file name of the lookup file. Relative paths are with respect to the installation directory for the data workbench server. This file is typically located in the Lookups directory within the data workbench server installation directory.	
IP Address	The field from which to read the IP address.	c-ip
Outputs	The names of the output strings. The IPLookup Quova and IPLookup Digital Envoy transformations have different output parameters. Be sure to use the appropriate transformation for your IP lookup data.	

In this example, **IP geo-location** data (in the lookup file Quova .bin) is used to create the output fields listed. The outputs (AOL, ASN, Area Code, and so on) can be used to create dimensions for geographical analysis of visitor traffic.

-1	Name	IPLookup IP Geo-location (Quova)
	-Outputs	Quova
	AOL	x-q-asn
	ASN	x-q-area-code
	Area Code	x-q-carrier
	Carrier	x-q-city
	City	x-q-connection
	Connection	x-q-continent
	Continent	x-q-country
	Country	x-q-dma
	DMA	x-q-ip-routing
	IP Routing Method	
	IP Type	x-q-latitude
	Latitude	x-q-line-speed
	Line Speed	x-q-longitude
	Longitude	x-q-msa
	MSA	
	Message	
	Metro City	
	Metro Distance	
	Metro State	
	PMSA	x-q-pmsa
	Region	x-q-region
	Return Code	
	Second Level Domain	x-q-domain-2
	State	x-q-state
	Timezone	x-q-timezone
	Top Level Domain	x-q-domain-1
	Zip	x-q-zip
	Comments	Comment
	Condition	AndCondition
	File	Lookups\IP Geo-location\Quova.bin
	IP Address	c-ip

LookupRows

Navigation title:LookupRows

The **LookupRows** transformation looks at other log entries with the same tracking ID and sets the value of the output field to the value of a designated field in the input row.

Because the **LookupRows** transformation performs its lookup on log entries and not lookup files, it is very similar to the **CrossRows** transformation. See [CrossRows](#) on page 577.

To work, the **LookupRows** transformation requires that the data is ordered in time and grouped by the tracking ID in your source data. Therefore, **LookupRows** works only when defined in the `Transformation.cfg` file or in a **Transformation Dataset Include** file.

As you review the descriptions of the parameters in the following table, remember the following:

- The output row is the row of data that the transformation is working on at a given point in time.

-
- Input rows are all of the other rows of data (before, after, or including the output row) whose values of the input field serve as inputs to the transformation.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	Limits the output of the transformation to certain log entries. If the condition is not met for a particular log entry, the field in Output Row Value Output parameter is left unchanged. The input still may be used to affect other log entries.	
Input Condition	Accepts input for the transformation from only certain input rows. If the Input Condition is not met for a particular input row, the input field from that row is ignored and does not affect other output rows. However, the output field from that row is still modified per the specified Condition.	
Input Row Key Input	The name of the field to use as the key for the input rows.	
Input Row Value Input	The name of the field in the input row whose value is copied to the field in the Output Row Value Output parameter if all conditions are satisfied.	
Operation	An operation that, for each output row, is applied to all of the input rows satisfying all of the conditions defined by the Input Condition and Input Row Key Input parameters to produce an output: <ul style="list-style-type: none"> • FIRST outputs the value of the field in the Input Row Value Input parameter from the first matching input row in the data (not the first matching row after the output row). • LAST outputs the value of the field in the Input Row Value Input parameter from the last input row in the data (not the last matching row before the output row). 	
Output Row Key Input	The name of the field to use as the key for the output row.	
Output Row Value Output	The name of the field in the output row whose value is copied from the field in the Input Row Value Input parameter if all conditions are satisfied. All output rows with the same x-trackingid and Output Row Key Input values have the same Output Row Value Output value.	

The **Input Row Key Input**, **Input Row Value Input**, and **Input Condition** parameters together define the lookup file for each tracking ID, while the **Output Row Key Input**, **Output Row Value Input**, and **Condition** parameters control what is looked up in the file and what value is stored in the field specified by **Output Row Value Output**.

To better understand the operation of the transformation, consider the following outline:

- For each output row satisfying the Condition and having a nonempty **Output Row Key Input**:
 - Find the FIRST or LAST input row such that
 - the input row satisfies the **Input Condition**, and
 - the x-trackingid of the input row equals the x-trackingid of the output row, and
 - the **Input Row Key Input** of the input row equals the **Output Row Key Input** of the output row,
 - and set the **Output Row Value Output** of the output row to the **Input Row Value Input** of the input row.

Considerations for **LookupRows**

- Blank key values never match anything. Even if there are input rows with blank keys and nonblank values that match the **Input Condition**, an **Output Row Key Input** of "" will always produce an **Output Row Value Output** of "".
- If not forbidden by the **Input Condition**, a row may look up itself if its **Input Row Key Input** and **Output Row Key Input** values are the same.

If you have multiple key values, you can combine them using a **Format** transformation (see [Format](#) on page 580) before applying a **LookupRows** transformation.

Suppose that you have a website that has a pet registration page, where the name and the breed are entered, and a later "buy toy" page where only the name of the pet is used. You would like to be able to link the pet name with the pet breed entered on the registration page. To do so, you could create the following **LookupRows** transformation:

-1	LookupRows
Name	Look up pet breed
-Input Condition	AndCondition
-0	NonEmptyCondition
Input	cs-uri-query(petbreed)
Comments	Comment
Condition	AndCondition
Input Row Key Input	cs-uri-query(petname)
Input Row Value Input	cs-uri-query(petbreed)
Operation	LAST
Output Row Key Input	cs-uri-query(petname)
Output Row Value Output	x-pet-breed

Let's analyze this example using the previous outline:

- For each output row satisfying having a nonempty value of cs-uri-query(petname) :
 - Find the LAST input row such that
 - the input row contains a nonempty value of cs-uri-query(petbreed), and
 - the x-trackingid of the input row equals the x-trackingid of the output row, and
 - the value of cs-uri-query(petname) of the input row equals the value of cs-uri-query(petname) of the output row,
 - and set the value of x-pet-breed of the output row to the value of cs-uri-query(petbreed) of the input row.

The **LookupRows** transformation uses the pet name (the key) to make sure that the pet breed is linked to both the pet registration and buy toy pages so that you can analyze the toys bought for each breed of pet, even for visitors with multiple pets.

Math

Navigation title:Math

The **Math** transformation enables the use of arithmetic operations on fields within the log entries.

The operations can include decimal integers and floating point constants.

Field	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Expression	<p>An arithmetic expression that describes the computation to be performed.</p> <p>You can use any of the operations and functions listed below, and you can incorporate field names in the expression:</p> <p>Operations</p> <ul style="list-style-type: none"> • Addition (+) • Subtraction (-) • Multiplication (*) • Division (/) • Remainder (%) 	

Field	Description	Default
	<ul style="list-style-type: none"> • Exponentiation (^) <p>Functions</p> <ul style="list-style-type: none"> • sgn(x). Returns 1 if x is positive, 0 if x is zero, or -1 if x is negative. • abs(x). Returns the absolute value of x. • floor(x). Returns the greatest integer less than or equal to x. • round(x). Returns the nearest integer to x. • log(b,x). Returns the logarithm of x base b. • min(x,y,...). Returns the smallest of all its arguments. • max(x,y,...). Returns the largest of all its arguments. 	
Output	The name of the field containing the result of the arithmetic operation.	

In this example, which uses fields of data collected from website traffic, a new field named x-page-duration is calculated by subtracting x-last-pv-timestamp from x-timestamp, then adding 1. The output is calculated only if the user-defined field x-last-pv-timestamp (which represents the timestamp of a visitor's last page view), is populated, or "not empty."

-5	MathTransformation
Name	Calculate Page Duration
-Condition	AndCondition
-0	NotEmptyCondition
Comments	Comment
Input	x-last-pv-timestamp
Comments	Comment
Default	
Expression	x-timestamp - x-last-pv-timestamp + 1
Output	x-page-duration

For information about the **Not Empty** condition, see [Conditions](#) on page 634.

Merge

Navigation title:Merge

The **Merge** transformation takes the values from the input field (typically a vector of strings), combines them into a single string separated by the given delimiter, and places the resulting string in the given output field.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition is met and the input value is not available.	
Delimiter	String that is used to separate the individual elements of the input string vector in the single output string. If you hold down the Ctrl key and right-click within the Delimiter parameter, an Insert menu appears. This menu contains a list of special characters that often are used as delimiters.	
Input	A vector of string values that are combined to form the output string.	
Output	The name of the output string.	

In this example, an input vector of strings is assumed to contain a set of products that were selected for purchase. These products are placed into a single output string and are separated by ":" (two colons).

-6	Merge
Name	Products Selected for Purchase
-Comments	Comment
0	Filed of products purchased for display
Condition	AndCondition
Default	
Delimiter	::
Input	x-products
Output	x-show-products

So if the input field x-products contained the string values B57481, C46355, and Z97123, the resulting output string x-show-products would be B57481::C46355::Z97123.

PullNameValues

Navigation title:PullNameValues

The **PullNameValues** transformation is a special operation that takes the values in the cs-uri-query field and separates each of the name-value pairs into a separate string.

The entire collection of name-value pair strings is output in the specified output field as a vector of strings.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition is met and the input value is not available in the given log entry.	
Output	The name of the output string.	

The **PullNameValues** transformation is used in this example to capture visitors' use of the search form: which buttons were selected, what values were typed in the form, and so on. The example uses a **String Match** condition (see [Conditions](#) on page 634) to isolate the use of this transformation to only the page /search.php. The vector of name-value pairs is output into the field x-search-namevalues.

-7	PullNameValues
Name	Search Features and Values
-Comments	Comment
0	List the features and values selected in a search
-Condition	AndCondition
-0	StringCondition
-Matches	
0	/search.php
Case Sensitive	true
Comments	Comment
Input	cs-uri-stem
Default	
Output	x-search-namevalues

Using the transformation as defined above, if the cs-uri-stem field matched the page /search.php and cs-uri-query contained the following:

- Searchfor=Bob&State=Virginia&isMale=true

then x-search-namevalues would contain a vector containing the following three strings:

- Searchfor=Bob
- State=Virginia
- isMale=true

REMatch

Navigation title:REMatch

The **REMatch** transformation is a pattern-matching transformation that uses regular expressions to specify one or more patterns to look for and capture in the input.

The transformation constructs an output field for each capturing sub-pattern in the regular expression. If the regular expression does not match the input field, the outputs are blank, and if the output field already exists, the values are replaced by the blank values. For a brief guide to using regular expressions, see [Regular Expressions](#) on page 642.



Note: The **REMatch** transformation operates similarly to the **RETransform** transformation (see [RETransform](#) on page 590), which uses regular expressions to capture a string and stores that string in a single output field.

REMatch parses a string more efficiently than multiple **RETransform** transformations or a single **RETransform** transformation followed by a **Flatten** transformation. See [Flatten](#) on page 580.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Case Sensitive	True or false. Specifies whether the match is case-sensitive.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Expression	The regular expression used for matching.	
Input	The field against which the regular expression is evaluated.	
Outputs	<p>The name of the output string or vector. In the case of string vectors as input, the outputs are also string vectors.</p> <p>An output field must exist for each capturing sub-pattern in the expression.</p>	



Note: **REMatch** transformations can be very slow and may account for much of the data processing time.

In this example, a **REMatch** transformation parses a date of the format YYYY-MM-DD into the fields x-year, x-month, and x-day. For the date 2007-01-02, the values of x-year, x-month, and x-day would be 2007, 01, and 02, respectively.

-2	REMatch
Name	Parse yyyy-mm-dd date
-Outputs	
0	x-year
1	x-month
2	x-day
Case Sensitive	false
Comments	Comment
Condition	AndCondition
Expression	(\d+)-(\d+)-(\d+) .*
Input	x-date

RETransform

Navigation title:RETransform

The **RETransform** (regular expression) transformation is a pattern-matching transformation that uses regular expressions to specify a pattern to look for and capture in the input and stores the captured string in a designated output field.

Regular expressions are evaluated against the entire input string. If the input does not match the pattern specified in the regular expression, no data is captured. For a brief guide to using regular expressions, see [Regular Expressions](#) on page 642.



Note: The **RETransform** transformation operates similarly to the **REMatch** transformation (see [REMatch](#) on page 589), which constructs an output field for each capturing sub-pattern in the regular expression. You can think of **RETransform** as a combination of **REMatch** and **Format** transformations. If the Action parameter (see Action in the following table) is set to "RESULTS," then **RETransform** operates like a combination of **REMatch** and **Union** transformations.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition is met and the input value is either not available or the regular expression does not match the input value.	
Action	Specifies how the result is treated. The default setting of RESULTS simply takes the patterns matched and creates a vector of strings from the patterns being extracted. Alternatively, the action may be a formatting string to create a simple string output of a particular format. With this technique, you specify the number corresponding to the location of each matched pattern between % signs. For example, the 1st matched pattern would be %1%, and the 3rd matched pattern would be %3%. You would specify other characters in the formatting string literally.	RESULTS
Expression	The regular expression used for matching.	
Input	The field against which the regular expression is evaluated.	
Output	The name of the output string.	



Note: **RETransform** transformations can be very slow and may account for much of the data processing time.

This example isolates the version of the Windows operating system that a website visitor is using and creates a field x-windows-version from that value. The output value in this case would simply be the version number.

-8	RETransform
Name	Windows Users
Action	RESULTS
Case Sensitive	true
Comments	Comment
Condition	AndCondition
Default	
Expression	.+Windows ([^;]+);)
Input	cs(user-agent)
Output	x-windows-version

If you wanted to include the string "Version" in front of the version number for readability, you would change the Action parameter from "RESULTS" to "Version %1%." To include a literal percent sign (%) in your output, escape it with a second percent sign, as in "%%."

Sessionize

Navigation title:Sessionize

If you are working with data collected from website traffic, you can use the **Sessionize** transformation to determine how sessions are defined.

The transformation takes as its input a timestamp and a tracking ID and outputs a session number for each log entry. The session number is "1" for the first session with a given tracking ID, "2" for the second session with the same tracking ID, and so on. The output can be used directly as a session key because it has a unique value for each session.



Note: To work, the **Sessionize** transformation requires that the data is ordered in time and grouped by the tracking ID in your source data. Therefore, **Sessionize** works only when defined in the `Transformation.cfg` file or in a **Transformation Dataset Include** file.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Input Timestamp	The field containing the values of the timestamp to be used.	x-timestamp
Input Tracking ID	The field containing the values of the tracking ID to be used. The value must be a 64-bit (16 digit) or smaller hexadecimal number or a decimal integers of 16 digits or less.  Note: If you wish to use a field other than x-trackingid for the tracking ID, you need to hash the field first. See Hash on page 581.	x-trackingid
Maximum Session Duration	The longest length of session before a new session is started. (This keeps web pages that have auto content refreshing from creating sessions that are arbitrarily long.) If the Timeout Condition is satisfied and the referrer of a click is set to one of the entries in the Internal Domains parameter, Maximum Session Duration is used to define the end of a session. No session may be longer than the specified Maximum Session Duration regardless of how many clicks it contains. The recommended value is 48 hours. For more information about the Maximum Session Duration and Internal Domains parameters, see Configuration Settings for Web Data on page 649.	48 hours
Output Session Number	The field in which the session number is stored. This field has a unique value for each session for each visitor.	
Session Timeout	The amount of time that needs to pass between log entries of a given visitor to determine the end of one session and the start of a new session (that is, the typical timeout used to define a user session). The recommended value of this parameter is 30 minutes. If the Timeout Condition is not	30 minutes

Parameter	Description	Default
	<p>satisfied and the referrer of a click is not set to one of the referrers in the Internal Domains parameter, Session Timeout is used to define the session.</p> <p>If the Timeout Condition is satisfied and cs(referrer-domain) for a log entry is in the list of internal domains, then Maximum Session Duration determines whether the current log entry is part of an existing session or the start of a new session.</p> <p>For more information about the Session Timeout parameter, see Configuration Settings for Web Data on page 649.</p>	
Timeout Condition	The condition that must be satisfied for a log entry to be considered the start of a new session. Note that the amount of time that passes between the log entry and the previous log entry must be at least the value of the Session Timeout parameter.	

A new session begins when any one of the following situations occurs:

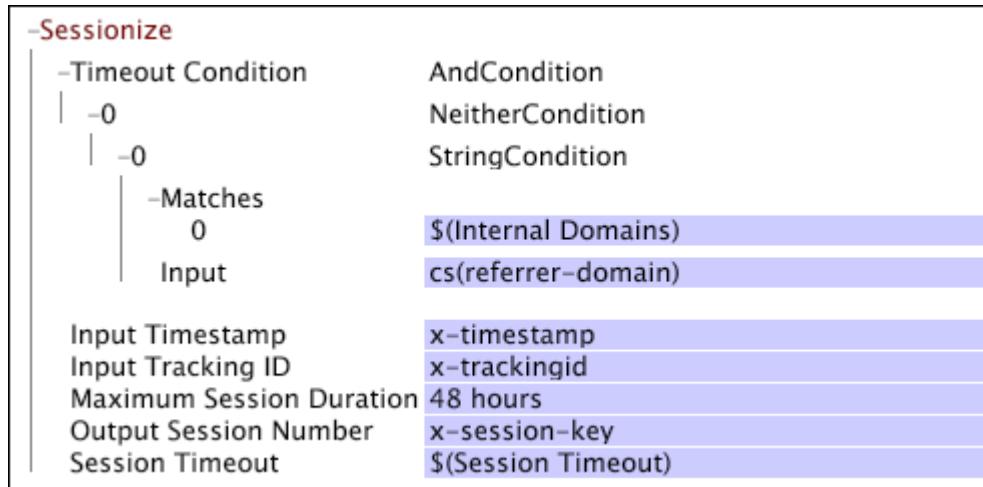
- The tracking ID changes.
- The time since the last log entry is at least equal to the value of the Session Timeout parameter and the Timeout Condition is satisfied.
- The time since the first log entry of the last session exceeds the value of the Maximum Session Duration parameter.



Note: If you have already defined Maximum Session Duration and Session Timeout as parameters in the `Session Parameters.cfg` file, do not enter values for them in the configuration. You can reference the parameters by typing `$(parameter name)` as shown in the following example. For more information about these parameters, see [Configuration Settings for Web Data](#) on page 649.

The **Sessionize** transformation in this example takes as its input the `x-timestamp` and `x-trackingid` fields and records the session number for each log entry in the `x-session-key` field. The transformation's **Timeout Condition** is based on a **Neither** condition: If the `cs(referrer-domain)` field for a log entry matches a member of the Internal Domains parameter, the condition evaluates to false. Note the references to the Internal Domains and Session Timeout parameters.

For information about the **NeitherCondition**, see [Conditions](#) on page 634. For information about the Internal Domains and Session Timeout parameters, see [Configuration Settings for Web Data](#) on page 649.



Split

Navigation title: Split

The **Split** transformation splits a string into a vector of substrings based upon a given delimiter character.

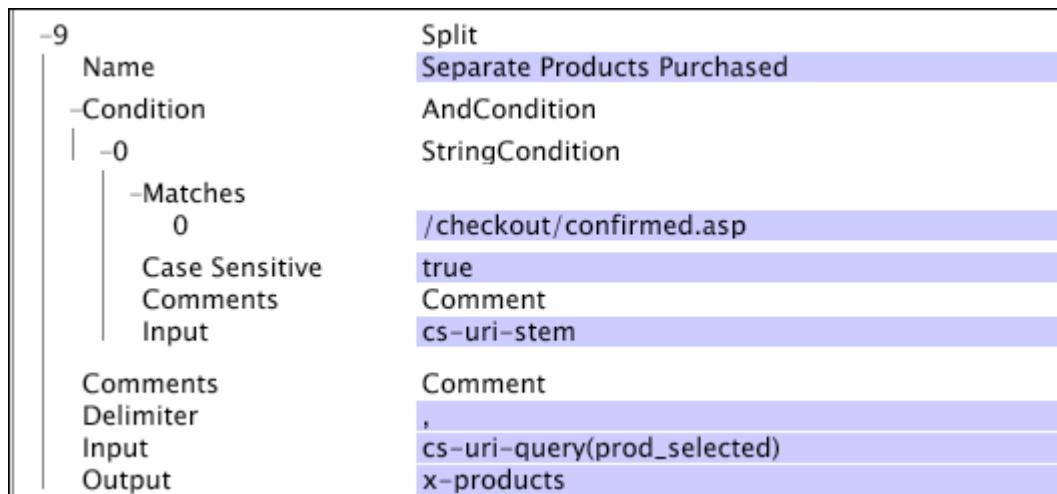
Split is particularly useful for extracting individual values from a collection of values associated with a single URI query name value.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Delimiter	String that is used to separate the input string into substrings. Must be a single character in length. If you hold down the Ctrl key and right-click within the Delimiter parameter, an Insert menu appears. This menu contains a list of special characters that often are used as delimiters.	
Input	The name of the field whose value is split to create the output string vector.	
Output	The name of the output field.	

Consider a website in which the products purchased by a customer are listed as part of the cs-uri-query value when the confirmation page associated with a successful purchase is accessed. The following is an example of such a string:

- /checkout/confirmed.asp?prod_selected=B57481,C46355,Z97123

The cs-uri-stem field is used to determine whether the page being requested by the log entry is the confirmation page. The codes for the products that the customer purchased are listed as the comma-separated values of the prod_selected name in the cs-uri-query. The **Split** transformation can be used to extract this information by splitting the product codes at the comma if the value of cs-uri-stem matches the value specified in the **String Match** condition. See [String Match](#) on page 638. The following transformation details the solution to this problem.



Here, the output field is x-products, which would be used to create the desired extended dimension that maps the products purchased to the sessions during which the purchase was made.

Tokenize

Navigation title: Tokenize

The **Tokenize** transformation iteratively applies a regular expression against the input string.

However, unlike **RETransform**, **Tokenize** does not have to match the entire string: the regular expression used for the **Tokenize** transformation can match a subset of the input. After a match is found, **Tokenize** applies the regular expression again, starting at the character after the end of the last match.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Case Sensitive	True or false. Specifies whether the match is case-sensitive.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition is met and the input value is either not available or the regular expression does not match the input value.	
Expression	The regular expression used for matching.	
Outputs	The names of the output strings. You can have multiple outputs for a given input string. The number of outputs must correspond to the number of capturing sub-patterns in the regular expression.	

In the following example, the **Tokenize** transformation uses a regular expression to capture the names of the query strings (in cs-uri-query) and output the captured sub-pattern (the query name) to x-pull-query-name.

-10	Tokenize
Name	Pull Query Names
-Outputs	x-pull-query-name
0	
Case Sensitive	true
Comments	Comment
Condition	AndCondition
Default	
Expression	[&]*([^\=]*)=[^\&]*
Input	cs-uri-query

For the query string "a=b&c=d," the output would be a vector containing "a" and "c."

For information about regular expressions, see [Regular Expressions](#) on page 642.

Union

Navigation title:Union

The **Union** transformation takes a set of inputs and creates a vector of strings as the output.

If one of the inputs is itself a vector, each element in the input vector is associated with one element in the output vector (that is, the transformation does not create a vector of vectors).

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition is met and the input value is not available.	
Inputs	One or more input values.	
Output	The name of the output field.	

This example uses fields of data from website traffic to create a list of the zip codes associated with the website's visitors (that is, within each log entry). The data provides two possible sources for this information: one in cs-uri-query and the other in a **zipcode** field of the cookie. If neither of these fields contains a zip code, the default value of 00000 is used.

-11	Union
Name	Available Zip Codes
-Inputs	
0	cs-uri-query(zipcode)
1	cs(cookie)(zipcode)
Comments	Comment
Condition	AndCondition
Default	00000
Output	x-zipcodes

While it is possible for both of these values to be available in a single log entry, you can select which value to use when you create a dimension based on the transformation's output. In a typical use case, you would create a simple dimension that takes either the first or the last of the encountered values. For information about creating simple dimensions, see [Extended Dimensions](#) on page 606.

URI Transformations

Navigation title:URI Transformations

The **URI** transformations modify the internal fields used to create the **URI** dimension provided with the Site application.

These transformations enable you to modify the **URI** elements to create values that are more useful and meaningful to your site analysts. In general, these transformations have the same basic operational principles as the standard transformations. See [Standard Transformations](#) on page 576.

AppendURI

Navigation title:AppendURI

The **AppendURI** transformation provides a way to add information to the default value that comes from the log entries used to build the dataset.

The transformation places a name-value pair at the end of the internal field used to create the URI dimension. The name-value pair is built using the Query String Key parameter as the name and the value of the identified Input parameter as the value of the pair. The **AppendURI** command adds any appropriate ? and & symbols necessary to separate the name-value pairs from the **URI** stem and from any previous **AppendURI** operations that may have been applied to the **URI**.

The **AppendURI** transformation works only when defined in the `Transformation.cfg` file or a **Transformation Dataset Include** file.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition is met and the input value is not available.	
Input	The name of the field whose value is appended to the URI .	
Query String Key	The name to use in the creation of the name-value pair being appended.	

Consider a website that was constructed using a traditional Model-View-Controller approach. In such systems, it is common to have a single web page be the point of access into the system. For such a site, visualizations of traffic patterns in the system would be very uninteresting and would provide no insights into visitor utilization and traffic flow. For example, consider a website that funnels all web requests through a **URI** of the following form:

- `http://www.examplesite.com/modelview.asp?id=login&name=bob`

The modelview ASP page receives all traffic and determines its actions based on the value of the id field in the query. By default, the URI dimension would contain a single entry:

- modelview.asp

This would result in a rather uninteresting mapping of the traffic through the site, as all traffic is being funneled through a single URI. To address this particular scenario and provide a more informative view into the underlying architecture of the website, **AppendURI** can be used to move some of the unique name-value pairs from the cs-uri-query field to the URI dimension used for visualizations. The transformation shown below gives the details of such a transformation:

-12	AppendURI
Name	Unique Pages
-Comments	Comment
0	Append ID to make pages unique
-Condition	AndCondition
-0	StringCondition
-Matches	
0	/modelview.asp
1	/xmlmodelview.asp
Case Sensitive	true
Comments	Comment
Input	cs-uri-stem
Default	
Input	cs-uri-query(id)
Query String Key	id

In this example, there are two pages used by the system to handle all requests: modelview.asp and xmlmodelview.asp. One page is used for browser traffic, and the other is used for system-to-system XML communications. The application server process uses the id name of the cs-uri-query to determine which action to take. Therefore, you can extract the value from the id field and append it to the URI. The result is a collection of URIs with a range of variation that reflects visitor traffic through the website. Here, a **String Match** condition determines the log entries to which the transformation is applied by searching the cs-uri-stem field for the two web pages of interest and ignoring all others. The input (the value of our name-value pair) is the result of cs-uri-query(id), which is "login." As specified by the Query String Key parameter, the name being appended is "id." Thus, for the incoming cs-uri value of our example, the resulting URI used by the **URI** dimension is /modelview.asp&id=login.

PrependURI

Navigation title:PrependURI

Similar to the **AppendURI** transformation, the **PrependURI** transformation affects the internal field used by the data workbench server to construct the URI dimension.

The **PrependURI** transformation works by adding the value in the identified input field to the front of the value currently in the URI.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition is met and the input value is not available.	
Input	The name of the field whose value is prepended to the URI.	

The following example simply prepends the s-dns field onto the URI, extending the representation of the URI dimension to include the domain requested by the client device.

-13	PrependURI
Name	Add Host to URI
Comments	Comment
Condition	AndCondition
Default	
Input	s-dns

In this example, prepending the s-dns field to the URI

- /modelview.asp?id=login

results in the following URL:

- www.adobe.com/modelview.asp?id=login

Now the URI is extended to include the domain requested.

ReplaceURI

Navigation title:ReplaceURI

The **ReplaceURI** transformation changes the value in the internal URI dimension to a new value.

If **URI Prefix** is specified, the resulting value is simply the URI prefix concatenated with the provided input value.

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition is met and the input value is not available.	
Input	The value to replace the URI.	
URI Prefix	The value (string) to be prepended to the value in the Input field.	



Note: Before applying **ReplaceURI** transformations, you should create a new simple dimension with a parent of **Page View** from a copy of cs-uri-stem or cs-uri. For assistance with this, contact Adobe.

This example demonstrates the use of **ReplaceURI** to replace the "page=*pageid*" query strings with "homepage.html" whenever *pageid* indicates that the website's homepage was viewed. The end result is a more user-friendly view of the URI.



For the transformation shown, the page

- www.examplesite.com/info.html?page=1550

would be changed to

- www.examplesite.com/homepage.html

UnescapeURI

Navigation title:UnescapeURI

The **Unescape URI** transformation unescapes any characters in a string that have been escaped.



Note: Escaped characters replace the unsafe characters in a URI string. They are represented by a triplet consisting of a percent sign followed by two hexadecimal digits (for example, %20).

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition is met and the input value is not available.	
Input	The URI string to be unescaped.	
Output	The name of the field in which the unescaped string is to be stored.	

The following transformation unescapes the docname value in a HTTP header field and stores the output in the field x-docname-unescaped:

-0	URIUnescape
Input	cs(docname)
Output	x-docname-unescaped

If the docname value were

• mysite.net/lending%20and%20leasing%20forms/document%20library/credit%20application.doc

then the value of x-docname-unescape would be

• mysite.net/lending and leasing forms/document library/credit application.doc

DeviceAtlas Distribution

The DeviceAtlas JSON file will now be distributed in a .bundle file (a renamed .tar.gz) along with DeviceAtlas.dll and DeviceAtlas64.dll files.

When the administrator upgrades the Insight Server to version 6.0, the DeviceAtlas.bundle file is included with the upgrade package in the Software and Docs profile (softdocs profile) located at:

Server Packages > v6.00 > Server_6.00.zip

The DeviceAtlas.bundle file is extracted to Server\Lookups\DeviceAtlas.

The DeviceAtlas.bundle file should be placed in a directory that is synchronized to the DPUs, and the DeviceAtlas.cfg file corresponding to the new DeviceAtlasComponent should be placed in the "Components for Processing Servers" directory on the synchronization master. When the DeviceAtlas.bundle file is changed, the very next DeviceAtlas lookup call will get results based on the updated API and/or JSON file.

Modify the Transformation.cfg file

The DeviceAtlas Transformations will no longer need to specify the path to the JSON file. Any previous DeviceAtlasTransformation that is defined in the transformation.cfg file should no longer include the File parameter that points to the obfuscated JSON file.

This example Transformation.cfg file shows the File argument that should be deleted to avoid confusion. (Leaving it there will not cause harm, but only potential confusion because it will be ignored.)

```
6 = DeviceAtlasTransformation:  
  Comments = Comment: 0 items  
  Condition = AndCondition: 0 items  
  
File = string: Lookups\\DeviceAtlas\\20110106_private.json.obfuscated
```

```

^^ DELETE THE ABOVE LINE FROM ALL PREVIOUS TRANSFORMATIONS ^^

Name = string: DeviceAtlas Lookup
Outputs = vector: 4 items
0 = Column:
    Column Name = string: vendor
    Field Name = string: x-vendor
1 = Column:
    Column Name = string: model
    Field Name = string: x-model
2 = Column:
    Column Name = string: isBrowser
    Field Name = string: x-isbrowser
3 = Column:
    Column Name = string: usableDisplayHeight
    Field Name = string: x-usable-display-height
User Agent = string: x-ua

```

Modify the DeviceAtlas.cfg file

This is an example of the component argument required in the DeviceAtlas.cfg file.

```

component = DeviceAtlasComponent:
    DeviceAtlas Bundle File = string:Lookups\\DeviceAtlas\\DeviceAtlas.bundle
    Unsynchronized Bundle Extraction Path = string: Temp\\DeviceAtlas\\

```

This DeviceAtlas.bundle file will be treated just like a configuration file from the perspective of the Profile Synchronization feature. In addition, the JSON data and DLL will be used at the Component level rather than at the individual Transformation level.

A new DeviceAtlasComponent, upon startup, finds the .bundle conglomeration, de-obfuscates the JSON file into memory, extracts the files into a temporary directory, and loads the appropriate DLL for the running platform. This component also monitors changes to the bundle file, and reloads the DLL and .cfg file automatically if it changes.

Running DeviceAtlas

Proper configuration makes a big difference in the time required for transformation. The transformation can be configured to run only once per visitor per session to allow DeviceAtlas to speed up the process.

If deployed using Log Processing.cfg:

Run the transformations twice.

1. Look up only the mobile_id field, then
2. Create conditions to ignore the mobile_id and then look up the rest of the fields.

If deployed using Transformation.cfg:

Deploy as in Step 1 in Log Processing above, or use cross-rows to support a conditional setting.

- Cross-Rows—Grab the previous session key. Then identify if the current session key is different from the one found with cross-rows. If so, then the DeviceAtlas transformation will only run on one record per session.

Integrating Lookup Data

Navigation title:Integrating Lookup Data

Data workbench provides a set of transformations that enables the data workbench server to incorporate lookup data into the dataset.

Lookup data is external data from corporate databases or lookup files that you can combine with event data to create the dataset. In general, you use lookup data to augment the event data from your log sources. Conceptually, you can think of using lookup data to populate event data records with additional columns of information.

When you use lookup data, you load the data into a memory-resident lookup table. A column in the table must contain a common key that also exists in the event data records. The data in the lookup table itself can be loaded from a flat file or from an ODBC data source. Lookup data can be incorporated into the dataset during the log processing or transformation phase of the dataset construction process.

To incorporate lookup data, you must first generate a lookup file or have the information needed to access an SQL database, then define one or more of the following transformations in the dataset configuration files for log processing and transformation.

To integrate lookup data into the dataset

1. Generate your lookup file. See [Populating the Lookup Table](#) on page 600.
2. Define one of the following types of transformations in the Transformations parameter in the appropriate dataset configuration file:
 - **Categorize**
 - **FlatFileLookup**
 - **ODBCLookup**



Note: Note that the **ODBCLookup** transformation works only when defined in the Transformation.cfg file or in a **Transformation Dataset Include** file.

Populating the Lookup Table

Navigation title:Populating the Lookup Table

If you use the **Categorize** or **FlatFileLookup** transformations, the lookup table is loaded in memory and populated from a flat file whose location you specify when you define the transformation.

The flat file you specify must meet the following requirements:

- Each line in the file must represent a row in the lookup table.
- Columns in the file must be separated by an ASCII delimiter. You may use any character that is not a line-ending character and does not appear anywhere within the event data itself. When you define the transformation, you specify which character has been used to delimit the columns in the flat file.

If you use an **ODBCLookup** transformation, the lookup table is loaded into memory and populated from a table or view in an **ODBC** database that you specify. When you define the transformation, you must also specify the data source, user name, and password that the data workbench server must use to establish a connection to the database.



Note: Lookup tables are loaded when the the data workbench server initially begins constructing the dataset. Once established, lookup files are not meant to be changed. If you change the flat file or **ODBC** table that is used for the transformation phase, you are required to retransform the entire dataset. If you change a flat file that is used during the log processing phase, the new lookup data is applied to all new records that enter the dataset, but the changes are not applied retroactively.

Defining Lookup Transformations

Navigation title:Defining Lookup Transformations

Information about the transformations that you can use to incorporate lookup data into the dataset.

Note that not all types can be used during both phases of the dataset construction process.

- [Categorize](#) on page 601
- [FlatFileLookup](#) on page 603

-
- [ODBCLookup](#) on page 605

Categorize

The **Categorize** transformation uses a two-column lookup table composed of pattern-string/value pairs. During this transformation, the data workbench server reads each event data record in turn and compares the contents of a designated field in the record to each of the pattern strings listed in the first column of the lookup table. If the designated field matches one of the pattern strings, the data workbench server writes the value (found in the second column) that is associated with that pattern string to a designated output field in the record.

The strings in the first column of the lookup table optionally can start with the ^ character and/or end in the \$ character to force matching at the beginning and/or end. This transformation does not accept regular expressions for defining match conditions in the first column. If the input value is a vector of strings, each string is run through the transformation and the result(s) are appended to an output string vector.

A **Categorize** transformation is generally easier and faster than using a **Regular Expression** transformation to accomplish the same thing.



Note: The substring test used in **Categorize** is case-sensitive unless otherwise specified using the **Case Sensitive** parameter.

Categorize

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Case Sensitive	True or false. Specifies whether the substring test is case-sensitive.	true
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition test passes and no entry in the categorization file matches the input, or the input field is not defined in the given log entry.	
Delimiter	String that is used to separate the columns in the lookup file. Must be a single character in length. If you hold down the Ctrl key and right-click within the Delimiter parameter, an Insert menu appears. This menu contains a list of special characters that often are used as delimiters.	
Multiple Values	True or false. If true, when multiple rows in the file match the input, each match results in a value being appended to the output vector of strings. If false, only the first matching row in the file is used in the output. In the latter case, if the input is a vector, the output is also a vector of equivalent length. If the input is a simple string, the output is also a simple string.	false
File	Path and file name of the categorization file. Relative paths are with respect to the installation directory for the data workbench server. This file is typically located in the Lookups directory within the data workbench server installation directory.	
Input	The categorization file matches its substrings against the value in this field to identify the matching row in the file.	
Output	The name of the field associated with the result.	



Note: Considerations for Categorize

- Changes to lookup files in **Categorize** transformations defined in the Transformation.cfg file or in a **Transformation Dataset Include** file require retransformation of the dataset. Lookup files for **Categorize**

transformations defined in the `Log Processing.cfg` file or a **Log Processing Dataset Include** file are not subject to this limitation. For information about reprocessing your data, see [Reprocessing and Retransformation](#) on page 627.

- **Categorize** transformations defined in the `Log Processing.cfg` file or a **Log Processing Dataset Include** file reload their lookup files whenever the lookup files change. Changes are not applied retroactively, but they apply to all log data read after the change takes place.

This example illustrates the use of the **Categorize** transformation to integrate lookup data with event data collected from website traffic. Suppose that a particular website has business sections, and there is a requirement to be able to look at and make comparisons based on traffic flow and value generated by the different sections. You can create a lookup file that lists the substrings used to identify these different sections.

The lookup file `Lookups\custommap.txt` contains the following table:

/products/	Products
^/sports/	Sports
^/news/	News
...	...

This categorization file maps anything containing the string "/products/" to the value "Products," anything starting with "/sports/" to the value "Sports," and anything starting with "/news/" to the value "News." The following categorization transformation uses the value in the `cs-uri-stem` field as the string within which we are looking for a matching substring. The result of the transformation is placed into the `x-custommap` field.

-14	Categorize
Name	Custom Site Mapping
Case Sensitive	true
Comments	Comment
Condition	AndCondition
Default	,
Delimiter	
File	lookups\custommap.txt
Input	cs-uri-stem
Multiple Values	false
Output	x-custommap

Assuming that the Multiple Values parameter is set to false, the example would produce the following values for `x-custommap` given the listed values for `cs-uri-stem`.

cs-uri-stem	x-custommap
/sports/news/today.php	Sports
/sports/products/buy.php	Products
/news/headlines.php	News
/news/products/subscribe.php	Products

The output is based on the order of the substrings in the lookup file. For example, the `cs-uri-stem` `/sports/products/buy.php` returns "Products." Although the URI stem starts with "/sports/", the string "/products/" is listed before "/sports/" in the lookup file. If the Multiple Values parameter were set to true, there would be an additional value for `x-custommap`, as the last example would match two rows in the lookup table: Products and News.

FlatFileLookup

The **FlatFileLookup** transformation uses a lookup table composed of any number of columns and rows (although, recall that it resides in memory). During this type of transformation, the data workbench server reads each event data record in turn and compares the contents of a designated field in the record to each of the values in a designated column of the lookup table. If there is a match, the data workbench server writes one or more values from the matching row in the lookup table to one or more designated output fields in the event data record.

The lookup table used during this transformation is populated from a flat file whose location you specify when you define the transformation.

FlatFileLookup

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Default	The default value to use if the condition is met and if no entry in the lookup file matches the input.	
Delimiter	String that is used to separate the columns in the lookup file. Must be a single character in length. If you hold down the Ctrl key and right-click within the Delimiter parameter, an Insert menu appears. This menu contains a list of special characters that often are used as delimiters.	
File	Path and file name of the lookup file. Relative paths are with respect to the installation directory for the data workbench server. This file is typically located in the Lookups directory within the data workbench Server installation directory.	
Header Row	True or false. Indicates that the first row in the table is a header row to be ignored in processing.	
Input	Column Name is the name of the column used for matching the input to the row(s) in the file. If Header Row is true, this can be the name of a column in the lookup file. Otherwise, this must be the zero-based column number to match against. Field Name is the name of the field used to locate the row in the lookup file.	
Multiple Values	True or false. Determines whether a single value (a matching row) or multiple values should be returned (one for each matching row).  Note: If Multiple Values is set to false, you must ensure that there are not multiple matches. When multiple matches occur, there is no guarantee which match will be returned.	
Outputs	A vector of column objects (results) in which each object is defined by column and field names. Column Name is the column from which the output value is obtained. If Header Row is true, this can be the name of a column in the lookup file. Otherwise, this must be the zero-based column number to match against. Field Name is the name of the field used to capture the output. Note that this can be a vector of results, one for each row identified in the case where the Multiple Values parameter is true.	



Note: Considerations for FlatFileLookup

- Matching the input field to the lookup file is always case-sensitive.
- Changes to lookup files in **FlatFileLookup** transformations defined in the `Transformation.cfg` file or `Transformation Dataset Include` files require retransformation of the dataset. Lookup files for **FlatFileLookup**

transformations defined in the `Log Processing.cfg` file or **Log Processing Dataset Include** files are not subject to this limitation. For information about reprocessing your data, see [Reprocessing and Retransformation](#) on page 627.

- **FlatFileLookup** transformations in the `Log Processing.cfg` file or **Log Processing Dataset Include** files reload their lookup files whenever the lookup files change. Changes are not applied retroactively, but they apply to all log data read after the change takes place.

This example illustrates the use of the **FlatFileLookup** transformation to integrate lookup data with event data collected from website traffic. Suppose that you want to isolate website partners that are routing traffic to the website and transform their partner IDs into more user-friendly names. You then can use the user-friendly names to create extended dimensions and visualizations that map more clearly to the business relationship than the site-to-site relationship used for routing traffic.

The example transformation searches the `cs(referrer-query)` field for the `PartnerID` name-value pair, and, if located, the lookup file `Lookups\partners.txt` is used to compare the `PartnerID` value against the values in the **Partner** column of the table. If a row is located, the output field `x-partner-name` is given the name from the **PrintName** column of the identified row.

-15	FlatFileLookup
Name	Renamed Referrer Partner
Comments	Comment
0	Gets nice names for referrer partners
Input	Column
Column Name	Partner
Field Name	<code>cs(referrer-query)(PartnerID)</code>
Outputs	Column
-0	PrintName
Column Name	<code>x-partner-name</code>
Field Name	
Condition	AndCondition
Default	No Partner
Delimiter	[TAB]
File	<code>Lookups\partners.txt</code>
Header Row	true
Multiple Values	false

If the lookup table contained the following information:

ID	Partner	Started	PrintName
1	P154	Aug 21, 1999	Yahoo
2	P232	July 10, 2000	Microsoft
3	P945	Jan 12, 2001	Amazon

The following examples would transform as follows:

- If `cs(referrer)(PartnerID)` returned P232, the field `x-partner-name` would be given the value "Microsoft."
- If `cs(referrer)(PartnerID)` returned P100, the field `x-partner-name` would be given the value "No Partner."
- If `cs(referrer)(PartnerID)` returned nothing, the field `x-partner-name` would be given the value "No Partner" as specified by the `Default` parameter.

ODBCLookup

The **ODBCLookup** transformation operates like a **FlatFileLookup** transformation. The only difference is that the lookup table used during this transformation is populated from an ODBC database and not a flat file.



Note: **ODBCLookup** transformations can be executed only during the transformation phase of the dataset construction process. When possible, Adobe recommends that you use the **FlatFileLookup** transformation instead of the **ODBCLookup** transformation. **FlatFileLookup** transformations are inherently more reliable because they do not depend on the availability of an outside system. Additionally, there is less risk that the lookup table is modified if it resides in a flat file that you control locally.

ODBCLookup

Parameter	Description	Default
Name	Descriptive name of the transformation. You can enter any name here.	
Comments	Optional. Notes about the transformation.	
Condition	The conditions under which this transformation is applied.	
Data Source Name	A DSN, as provided by an administrator of the data workbench server machine on which the dataset is processed, that refers to the database from which data is to be loaded.	
Database Password	The password to be used when connecting to the database. If a password has been configured for the DSN in the Data Source Administrator , this may be left blank. Any password supplied here overrides the password configured for the DSN in the Data Source Administrator .	
Database User ID	The user ID to be used when connecting to the database. If a user ID has been configured for the DSN in the Data Source Administrator , this may be left blank. Any user ID supplied here overrides the user ID configured for the DSN in the Data Source Administrator .	
Default	The default value to use if the condition is met and no entry in the lookup file matches the input.	
Input Column	Column Name is the column name or SQL expression for the data that is matched against the input. Field Name is the name of the field containing the data to be looked up.	
Multiple Values	True or false. Determines whether a single value (a matching row) or multiple values should be returned (one for each matching row).  Note: If Multiple Values is set to false, you must ensure that there are not multiple matches. When multiple matches occur, there is no guarantee which match will be returned.	
Output Columns	A vector of column objects (results) where each object is defined by column and field names. Column Name is the name of or SQL expression for the column from which the output value is obtained. Field Name is the name of the field used to capture the output.	
Table Identifier	An SQL expression that names the table or view from which data is to be loaded. A typical table identifier is of the form SCHEMA.TABLE.	

- The Data Source Name, **Database User ID**, **Database Password**, and Table Identifier parameters are the same as the parameters of the same names that are described for ODBC data sources. See [ODBC Data Sources](#) on page 547.
- Unlike ODBC data sources, **ODBCLookup** transformations do not require an increasing ID column. See [ODBC Data Sources](#) on page 547. That is because the contents of the lookup table must not change in any way while the dataset is active. Changes in a lookup table or view can not be detected until retransformation occurs. For information about reprocessing your data, see [Reprocessing and Retransformation](#) on page 627.

Suppose that you want to convert outdated DNS records to the updated records. Both sets of records are stored in an SQL database. To perform this task, you would reference a lookup table that is generated from the database and replace the outdated DNS records.

Our example transformation searches the log entries for the s-dns field, and, if located, the lookup table VISUAL.LOOKUP is used to compare the s-dns entry against the entries in the **OLDDNS** column of the table. If a row is located in the table, the output field s-dns is given the updated DNS record entry from the **NEWDNS** column of the identified row.

-16	ODBCLookup
Name	Domain Name
-Input Column	Column
Column Name	OLDDNS
Field Name	s-dns
-Output Columns	
-0	Column
Column Name	NEWDNS
Field Name	s-dns
Comments	Comment
Condition	AndCondition
Data Source Name	VSTest0
Database Password	visual
Database User ID	visual
Default	
Multiple Values	false
Table Identifier	VISUAL.LOOKUP

Extended Dimensions

Navigation title:Extended Dimensions

Instructions to create extended dimensions and describe the types of extended dimensions that you can define for creation during the transformation phase of data set construction.

Derived dimensions make up another category of dimensions used by the Insight Server. As the name suggests, derived dimensions are created from existing extended dimensions or metrics. You do not define derived dimensions within a **Transformation Dataset Configuration** file as you do extended dimensions. Instead, you define them as individual .dim files within an inherited profile or a data set profile.

See [Extended Dimensions](#) for steps to create a derived dimension.

About Extended Dimensions

Navigation title:About Extended Dimensions

The Insight Server (InsightServer64.exe) enables you to define custom dimensions from event data or lookup data.

Any custom dimensions that you define are referred to as extended dimensions. You can use them to create visualizations, build extended metrics, or perform analysis to understand the operations and issues associated with your business channel. You can define several types of extended dimensions in the Transformation.cfg file or in **Transformation Dataset Include** files.

An extended dimension represents a relationship between log field values and a parent dimension. A parent dimension can be any user-defined countable dimension. See [Countable Dimensions](#) on page 608. You specify the parent when defining the dimension in the Transformation.cfg file or a **Transformation Dataset**

Include file. A dimension's parent is the same as its level. For example, if you define a dimension with a parent of Session, then that dimension is a session-level dimension.



Note: The log field values can come from the inherent values available in the log (.vsl) files or other event data sources or from extended log fields created through the use of transformations.

To add an extended dimension to a visualization, you access it from the Extended list within the **Select Dimension** menu. For example, to add an extended dimension to a graph visualization, you would right-click within the workspace and click **Add Visualization > Graph > Extended > <dimension name>**. If you would like to organize the list of your extended dimensions within the data workbench interface, you can move the extended dimensions into subfolders that you create. See the Administrative Interfaces chapter of the *Data Workbench User Guide*. If you do this, the subfolders' names also appear in the menu, as in **Add Visualization > Graph > Extended > <subfolder name> > <dimension name>**.

To see all the dimensions that have been defined for your dataset profile and the buffer size for each, open the **Detailed Status** interface in data workbench and click **Performance**, then **Dimensions** to expand the nodes. The buffer size, which controls query times, is expressed in MB. For more information about the **Detailed Status** interface, see the Server Administration and Installation guide.

Defining Extended Dimensions

Navigation title:Defining Extended Dimensions

Steps to define extended dimensions.

1. While working in your dataset profile, open the **Profile Manager** and click **Dataset** to show its contents.
2. Open the **Transformation.cfg** file or the **Transformation Dataset Include** file in which you want to define the extended dimension.
 - (Recommended) To open a dataset include file, see *Dataset Include Files* on page 557.



Note: Adobe recommends defining extended dimensions in one or more new **Transformation Dataset Include** files. For more information, see *Creating New Dataset Include Files* on page 558.

- To open the **Transformation.cfg** file, see *Editing the Transformation Configuration File* on page 551.
3. Right-click **Transformations** and click **Add new > <Extended dimension type>**.
 4. Input the appropriate information for your extended dimension. For descriptions of the transformation types and information about their parameters, see the following sections:
 - *Countable Dimensions* on page 608
 - *Simple Dimensions* on page 610
 - *Many-to-Many Dimensions* on page 611
 - *Numeric Dimensions* on page 612
 - *Denormal Dimensions* on page 614
 - *Time Dimensions* on page 615

For any extended dimension that you define, you can add one or more comment lines to the **Comments** parameter to further describe the dimension or add notes about its use. To add a comment, right-click the **Comments** label and click **Add new > Comment Line**.

5. After you have defined your extended dimension(s) in the configuration file, save the file locally and save it to your dataset profile on the data workbench server.

Types of Extended Dimensions

Navigation title:Types of Extended Dimensions

The Insight Server enables you to define countable, simple, many-to-many, numeric, denormal, and time dimensions for inclusion in your data set.

Each dimension type has a set of parameters whose values you edit to provide specific instructions for the Insight Server to create the dimensions during the transformation phase of data set construction.

While some of the parameters differ among the dimension types, all require the specification of a parent dimension (the Parent parameter). The parent dimension determines which log entries from the log sources are provided as input to the new dimension. In other words, the log entries that are associated with the elements of the parent dimension are the ones that are associated with the new dimension before any filtering is applied. The parent dimension also determines the new dimension's position within the data set's hierarchy, referred to as the data set schema. For information about the interface that shows the data set schema, see [Dataset Configuration Tools](#) on page 628.

After the Insight Server uses the parent dimension to determine which log entries should be considered in the creation of the dimension, it applies the specified condition(s) (the Condition parameter) to blank the log entries that do not satisfy the condition. The server then identifies the value of the specified input field (the Input parameter) for each log entry and applies the specified operation (the Operation parameter), if applicable.



Note: If a log entry does not satisfy an extended dimension's Condition, the Insight Server substitutes blank values for all of the fields in the log entry. The actual log entry still exists, and the specified Operation determines whether the blank value of the **Input** field is used.

Countable Dimensions

Navigation title:Countable Dimensions

A countable dimension's elements can be counted by the system.

Countable dimensions are typically used to create sum metrics, which return the count, or sum, of all the elements of the dimension. You can define countable dimensions to count instances such as reservation bookings or product orders. For example, you could define the countable dimension Orders whose elements (log entries corresponding to orders from your online store) could be counted. If you want to show a count of orders within a visualization, you would define the Orders sum metric, which can be evaluated over a dimension or have filters applied to it.

Countable dimensions can be parents of other dimensions or children of other countable dimensions.



Note: If you need a dimension that only provides a count of something, you should use a numeric dimension with an operation of COUNT. See [Numeric Dimensions](#) on page 612.

Countable dimensions are defined by the following parameters:

Parameter	Description	Default
Name	Descriptive name of the dimension as appears to the user in data workbench. The dimension name cannot include a hyphen (-).	
Comments	Optional. Notes about the extended dimension.	
Condition	The conditions under which the input field contributes to the creation of the countable dimension. If specified, a condition restricts the set of log entries visible to the dimension and all of its children in the dataset schema.	
Hidden	Determines whether the dimension appears in the data workbench interface. By default, this parameter is set to false. If, for example, the dimension is to be used only as the basis of a metric, you can set this parameter to true to hide the dimension from the data workbench display.	false

Parameter	Description	Default
Key	<p>Optional. The name of the field to use as the key. If you define this parameter, an element of the countable dimension exists for every combination of an element of the countable dimension's parent and a distinct value of the field specified as the key.</p> <p>Each element of the countable dimension is required to relate to a contiguous set of log entries. Therefore, if the log entries are not ordered by the key, an element of the countable dimension is created each time the key field changes. To prevent this situation, Adobe recommends that you use a unique key which is contiguous in time order.</p>	
Parent	<p>The name of the parent dimension. Any countable dimension can be a parent dimension. To make a dimension the top-level dimension in the dataset's schema, set the parameter to "root." The defined dimension becomes the root countable dimension for the dataset. For example, if you are working with Site, the Visitor dimension is the root countable dimension for your dataset.</p> <p> Note: Although your root countable dimension does not have to be associated with the tracking IDs in the data, Adobe recommends that you configure your dataset's root countable dimension to use the tracking ID field (x-trackingid) as its Key. As a result, each element of the root countable is associated with a unique value of x-trackingid, and all of the data about each element is grouped together. If you would like to configure your dataset differently, contact Adobe.</p>	

This example illustrates the definition of a countable dimension using event data collected from website traffic. The countable dimension counts the web campaign events within a given session. The assumption is that all email campaign resources are requested from the web server with "email=" as part of cs-uri-query. In the example, the number of times that the visitor responds to an email campaign during a given session is of interest, not the actual value of the cs-uri-query(email) field.

-0	Countable
Name	Campaign Event
-Condition	AndCondition
-0	RECondition
-Matches	
0	.+
Case Sensitive	true
Comments	Comment
Input	cs-uri-query(email)
Comments	Comment
Hidden	false
Key	
Parent	Session

This example also illustrates the definition of a countable dimension using event data collected from website traffic, but it has a defined Key parameter. The Session countable dimension uses the x-session-key field as its key. (The x-session-key field is the output of the **Sessionize** transformation and has a unique value for each session.) Every unique combination of an element of the Visitor dimension (the parent) and the x-session-key field is an element of the Session dimension.

-0	Countable
Name	Session
Comments	Comment
Condition	AndCondition
Hidden	false
Key	x-session-key
Parent	Visitor

Simple Dimensions

Navigation title:Simple Dimensions

A simple dimension has a one-to-many relationship with its parent countable dimension.

A simple dimension is always a child of a countable dimension. You can think of a simple dimension as a representation of a property of the elements in its parent dimension. For example, if you are working with web data, you could define the Visitor Referrer dimension, which is a simple dimension with a parent dimension of Visitor. It represents the first HTTP referrer for each visitor in the Visitor dimension. Each visitor in the Visitor dimension has only one visitor referrer, but many visitors can have the same visitor referrer. Therefore, the Visitor Referrer dimension has a one-to-many relationship with the Visitor dimension.

Simple dimensions are defined by the following parameters:

Parameter	Description	Default
Name	Descriptive name of the dimension as it appears in data workbench. The dimension name cannot include a hyphen (-).	
Comments	Optional. Notes about the extended dimension.	
Condition	The conditions under which the relationship between the Parent and the input field's value should be created.	
Hidden	Determines whether the dimension appears in the data workbench interface. By default, this parameter is set to false. If, for example, the dimension is to be used only as the basis of a metric, you can set this parameter to true to hide the dimension from the data workbench display.	false
Input	The field of values that is related to the parent dimension (Parent).	
Load File	Optional. A file of available values for the relationship. You use a load file when either of the following applies: <ul style="list-style-type: none"> The values have a specific sort order that you want to preserve in the data workbench display. For example, you might want to create a Quarter dimension whose elements (the quarters of the year) always display in chronological order. You want to create place holders for values that may not be found in the data but need to appear in the data workbench display. If a value is encountered that is not present in the file, it is added to the end of the values when viewed in data workbench.	
Operation	Available operations are as follows: <ul style="list-style-type: none"> FIRST NONBLANK: The first non-blank input value is used, regardless of whether it comes from the first log entry. If Input is a vector field, the first row in the vector for the relevant log entry is used. FIRST ROW: The value for the first log entry related to the parent dimension element is used, even if the input is blank. If Input is a vector field, the first row in the vector for the relevant log entry is used. If this value is blank or not a number, or if the relevant log entry does not meet the dimension's Condition, no value is used. LAST NONBLANK: The last non-blank input value is used, regardless of whether it comes from the last log entry. If Input is a vector field, the first row in the vector for the relevant log entry is used. LAST ROW: The value for the last log entry related to the parent dimension element is used, even if the input is blank. If Input is a vector field, the first row in the vector for the relevant log entry is used. If this value is blank or not a number, or if the relevant log entry does not meet the dimension's Condition, no value is used. 	

Parameter	Description	Default
	 Note: If Operation yields no value or a blank value for a particular log entry, the corresponding element of the parent dimension will relate to the "None" element of the simple dimension. You should specify an operation to ensure that the dimension is defined as intended.	
Parent	The name of the parent dimension. Any countable dimension can be a parent dimension.	

This example illustrates the definition of a simple dimension using event data collected from website traffic and a load file.

Consider the example of a poll of site visitors' favorite Girl Scout cookies. A web page captures this vote and returns it to the web server in the name-value pair favoritecookie. Only one vote per visitor is counted, but visitors can change their minds and vote again if desired. This is a one-to-many relationship: one visitor can have many votes, but each vote is associated with only one visitor. Therefore, the parent of the dimension is visitors (only one vote per visitor) and the operation is LAST ROW (so that they can change their minds and vote again).

Placeholders must exist for all types of cookies so that cookie types receiving no votes appears in the data workbench display. For these reasons, a load file has been defined that contains the list of cookie types that may be selected. This file's contents, saved in a file named cookietypes.txt, looks something like the following:

Animal Treasures

Caramel Delights

Lemon Pastry Creams

Peanut Butter Patties

Shortbreads

Thin Mints

The final dimension is defined as shown here:

-2	Simple
Name	Favorite Cookie
Case Sensitive	true
Comments	Comment
Condition	AndCondition
Hidden	false
Input	cs-uri-query(favoritecookie)
Load File	Lookups\cookietypes.txt
Maximum Elements	32768
Operation	LAST ROW
Parent	Visitor

Many-to-Many Dimensions

Navigation title:Many-to-Many Dimensions

A many-to-many dimension has a many-to-many relationship with its parent countable dimension.

You can think of a many-to-many dimension as a representation of a set of values for each element in its parent dimension. For example, the many-to-many dimension Search Phrase is a Session-level dimension (it has a parent of Session). It represents the set of search phrases associated with each session in the Session dimension. A single search phrase can be used in any number of sessions, and a single session can include zero or more search phrases. Therefore, the Search Phrase dimension has a many-to-many relationship with the Session dimension.

Many-to-many dimensions are defined by the following parameters:

Parameter	Description	Default
Name	Descriptive name of the dimension as it appears to the user in data workbench. The dimension name cannot include a hyphen (-).	
Comments	Optional. Notes about the extended dimension.	
Condition	The conditions under which the relationship between the parent and the input field's value should be created.	
Hidden	Determines whether the dimension appears in the data workbench interface. By default, this parameter is set to false. If, for example, the dimension is to be used only as the basis of a metric, you can set this parameter to true to hide the dimension from the data workbench display.	false
Input	<p>The value that is related to the parent dimension (Parent). If this field is a vector of strings, then each element of the vector has its own relationship with the parent.</p> <p> Note: If the input value for every log entry for an element of the parent dimension is empty, no element of the many-to-many dimension will relate to that element of the parent dimension.</p>	
Parent	The name of the parent dimension. Any countable dimension can be a parent dimension.	

This example illustrates the definition of a many-to-many dimension using event data collected from website traffic. This many-to-many dimension, named "Selected Product," relates sessions to the products purchased by the visitor during that session. The x-products field contains a vector of values, each of which is associated with a page view, which, in turn, is associated with a session.

-3	ManyToMany
Name	Selected Product
Case Sensitive	true
Comments	Comment
Condition	AndCondition
Hidden	false
Input	x-products
Maximum Elements	32768
Parent	Session

By creating such a transformation, you can create a visualization in data workbench that depicts the relationship between the selected product dimension and the number of sessions that involve each of the products.

Numeric Dimensions

Navigation title:Numeric Dimensions

A numeric dimension consists of ordered, numerical elements and has a one-to-many relationship with its parent countable dimension.

You can think of a numeric dimension as a representation of the numeric properties of the parent dimension's elements. For example, if you are working with web data, you could define the numeric dimension Session Revenue, which defines an amount of revenue, in dollars, for each session in the Session dimension. Each session has a single amount of associated revenue, but several sessions can have the same amount of associated revenue. Therefore, the Session Revenue dimension has a one-to-many relationship with the Session dimension.

Numeric dimensions are often used to define metrics that sum values, count occurrences of a condition, or locate a minimum or maximum value. For example, a metric named "Revenue" might be defined using the Session Revenue dimension: $\text{sum}(\text{Session_Revenue}, \text{Session})$. Defined this way, the Revenue metric would give the total amount of revenue for the selected sessions.

Numeric dimensions cannot be parents of other dimensions.

Numeric dimensions are defined by the following parameters:

Parameter	Description	Default
Name	Descriptive name of the dimension as it appears in data workbench. The dimension name cannot include a hyphen (-).	
Clip Values	True or false. Specifies whether the input value (after Operation) is to be clipped to be between the values of Min and Max. If Clip Values is true, the value is clipped to that range. If Clip Values is false, no value is returned for the element of the parent dimension.	
Comments	Optional. Notes about the extended dimension.	
Condition	The conditions under which the input field contributes to the creation of the numeric dimension.	
Fixed Size	True or false. Controls the number of elements in a dimension (cardinality). If true, all elements from Min to Max are included in the dimension. If false, the dimension's size grows as values are added.	false
Hidden	Determines whether the dimension appears in the data workbench interface. By default, this parameter is set to false. If, for example, the dimension is to be used only as the basis of a metric, you can set this parameter to true to hide the dimension from the data workbench display.	false
Input	The value to use with the specified Operation or the input value for which you want to count occurrences. If this field is a vector of strings, the evaluation occurs for each element in the vector. So for example, a vector with length 3 and an Operation of COUNT adds 3 to the count.	
Min	Lower limit on the final dimension result.	0
Max	Upper limit on the final dimension result.	1e6
Offset	See Scale in this table.	0
Operation	<p>Available operations are as follows:</p> <ul style="list-style-type: none"> COUNT: The total number of nonblank values in the Input field across all log entries that meet the dimension's Condition is used. If the Input field is a vector field, the total number of nonblank values in each log entry is counted. FIRST NONBLANK: The first non-blank input value is used, regardless of whether it comes from the first log entry. If Input is a vector field, the first row in the vector for the relevant log entry is used. If the value is not a number, no value is used. FIRST ROW: The value for the first log entry related to the parent dimension element is used, even if the input is blank. If Input is a vector field, the first row in the vector for the relevant log entry is used. If this value is blank or not a number, or if the relevant log entry does not meet the dimension's Condition, no value is used. LAST NONBLANK: The last non-blank input value is used, regardless of whether it comes from the last log entry. If Input is a vector field, the first row in the vector for the relevant log entry is used. If the value is not a number, no value is used. LAST ROW: The value for the last log entry related to the parent dimension element is used, even if the input is blank. If Input is a vector field, the first row in the vector for the relevant log entry is used. If this value is blank or not a number, or if the relevant log entry does not meet the dimension's Condition, no value is used. SUM: The total of all the numeric values in the Input field across all log entries that meet the dimension's Condition is used. If there are no such log entries or no numeric values found, the numeric value 0 is used. MIN or MAX: The minimum or maximum numeric value found in the Input field across all log entries that meet the dimension's Condition is used. If there are no such log entries or no numeric values, no value is used. <p> Note: You should specify an operation to ensure that the dimension is defined as intended.</p>	
Parent	The name of the parent dimension. Any countable dimension can be a parent dimension.	
Scale	To yield the ordinal value of the dimension, the result of Operation is transformed as follows: $(\text{scale} * \text{input}) + \text{offset}$	1.0



Note: If **Operation** yields no value, or **Clip Values** is false and the value is not between **Min** and **Max**, no element of the numeric dimension is related to the element of the parent dimension.

This example illustrates the definition of a numeric dimension using event data collected from website traffic. This numeric dimension, named "Ad View Counter," counts the number of times that the visitor sees an advertisement during a given session. The assumption is that all advertisement resources are requested from the web server with ad= as part of the cs-uri-query. In the example, the number of times (COUNT) that the visitor is presented with an advertisement is the value of interest, not the actual value in the field.

-4	Numeric
Name	Ad View Counter
Clip Values	false
Comments	Comment
Condition	AndCondition
Fixed Size	false
Hidden	false
Input	cs-uri-query(ad)
Max	1e6
Min	0.0
Offset	0.0
Operation	COUNT
Parent	Session
Scale	1.0

Denormal Dimensions

Navigation title: Denormal Dimensions

A denormal dimension has a one-to-one relationship with its parent countable dimension.

You would define a denormal dimension whenever the desired dimension contains a unique element for each element of its parent. For example, **EMail Address** is a denormal dimension with a parent of **Visitor**. Each **Visitor** has an email address, and each element in the **EMail Address** dimension is the email address of a single visitor. Even if two visitors have the same email address, the addresses are distinct elements of the **EMail Address** dimension.

You can use denormal dimensions in any table visualization, in detail tables, or to create filters. In addition, you can use denormal dimensions with the data workbench server's segment export functionality to export values of fields (such as **Tracking ID** or **EMail Address**) that have lots of values. Because any segment data that you want to export must be defined as a dimension within the profile, you must create a denormal dimension that stores the raw strings of the field's data.



Note: When using a denormal dimension in a table or other visualization that expects a normal dimension, a derived denormal dimension is created automatically. The derived denormal dimension has a one-to-many relationship with the parent dimension.

For information about the detail table visualization and filters, see the Analysis Visualizations chapter in the *Data Workbench User Guide*. For information about segment export, see the Configuring Interface and Analysis Features chapter in the *Data Workbench User Guide*.



Note: Denormal dimensions can be very expensive in query time and disk space. A denormal dimension with parent **Page View** and a 50-byte average input string could add 25 GB of data to the buffers in a typical, large dataset, equivalent to about 13 simple or numeric page view dimensions, or about 125 session level dimensions. Never add a denormal dimension to a dataset without a careful evaluation of the performance impact.

Denormal dimensions are defined by the following parameters:

Parameter	Description	Default
Name	Descriptive name of the dimension as it appears in data workbench. The dimension name cannot include a hyphen (-).	
Comments	Optional. Notes about the extended dimension.	
Condition	The conditions under which the relationship between the Parent and the input field's value should be created.	
Hidden	Determines whether the dimension appears in the data workbench interface. By default, this parameter is set to false. If, for example, the dimension is to be used only as the basis of a metric, you can set this parameter to true to hide the dimension from the data workbench display.	true
Input	The value that is related to the parent dimension (Parent).	
Normalized Elements	A performance tuning parameter that specifies the number of dimension elements whose names are to be stored in system memory. Setting this parameter to a higher value causes a denormal dimension to use more RAM but results in faster queries. The default value is 16383.	
Operation	<p>Available operations are as follows:</p> <ul style="list-style-type: none"> FIRST NONBLANK: The first non-blank input value is used, regardless of whether it comes from the first log entry. If Input is a vector field, the first row in the vector for the relevant log entry is used. FIRST ROW: The value for the first log entry related to the parent dimension element is used, even if the input is blank. If Input is a vector field, the first row in the vector for the relevant log entry is used. If this value is blank or not a number, or if the relevant log entry does not meet the dimension's Condition, no value is used. LAST NONBLANK: The last non-blank input value is used, regardless of whether it comes from the last log entry. If Input is a vector field, the first row in the vector for the relevant log entry is used. LAST ROW: The value for the last log entry related to the parent dimension element is used, even if the input is blank. If Input is a vector field, the first row in the vector for the relevant log entry is used. If this value is blank or not a number, or if the relevant log entry does not meet the dimension's Condition, no value is used. <p> Note: If Operation yields no value, a blank value ("") is used.</p> <p>You should specify an operation to ensure that the dimension is defined as intended.</p>	
Parent	The name of the parent dimension. Any countable dimension can be a parent dimension.	

The denormal dimension shown in this example takes all of the data in the field x-trackingid as input and includes it in a dimension named Visitor ID. For a segment of visitors that you have created, you can export the data in the Visitor ID dimension (as well as any other defined dimension).

-5	Denormal
Name	Visitor ID
Comments	Comment
Condition	AndCondition
Hidden	false
Input	x-trackingid
Normalized Elements	16383
Operation	
Parent	Visitor

Time Dimensions

Navigation title: Time Dimensions

A time dimension enables you to create a set of periodic or absolute local time dimensions (such as Day, Day of Week, Hour of Day, Reservation Time, and so on) based on any timestamp field that you specify for the Input Time (1970 epoch) parameter.

When defining time dimensions, you also can choose a day other than Monday to be used as the start of a week by specifying the Week Start Day parameter. You can define more than one set of time dimensions in your dataset as long as the dimensions have different names.

Time dimensions are defined by the following parameters:

Parameter	Description	Default
Name	Descriptive name of the dimension as it appears in data workbench. The dimension name cannot include a hyphen (-).	
Comments	Optional. Notes about the extended dimension.	
Dimensions	<p>You can specify dimension names for any of the following periods:</p> <ul style="list-style-type: none">• Day• Day of Week• Hour• Hour of Day• Month• Week <p>The names that you enter here are the names that appear in dimension menus and in visualizations in data workbench. If you leave the name of a time dimension blank, the dimension is not created in the dataset.</p>	
Hidden	Determines whether the dimension appears in the data workbench interface. By default, this parameter is set to false. If, for example, the dimension is to be used only as the basis of a metric, you can set this parameter to true to hide the dimension from the data workbench display.	true
Input Time (1970 epoch)	<p>The name of the timestamp field to use as input.</p> <p> Note: The field's values must represent the number of seconds since January 1st, 1970, at 00:00:01. If the input time is not a valid time (1970 to 2037), the transformation process will fail, and the data workbench server will generate an error.</p>	
Parent	The name of the parent dimension. Any countable dimension can be a parent dimension. For web data, the parent is Session.	
Week Start Day	<p>The day to use as the first day of a week.</p> <p>This parameter affects the Week dimension, the Day of Week dimension, and any reporting time dimensions defined in terms of weeks.</p>	Mon

This example creates a set of time dimensions based on the user-defined input field x-time-1970. The set of time dimensions is named "Session Time." Because the parent of each dimension is the Session dimension, each element of the time dimensions corresponds to the time at which a session began. The Week Start Day parameter specifies that each week of the Week dimension starts on Monday.

-Extended Dimensions	
-0	Name
-Dimensions	TimeDimensions
Day	Session Time
Day of Week	map
Hour	Day
Hour of Day	Day of Week
Month	Hour
Week	Hour of Day
Comments	Month
Hidden	Week
Input Time (1970 epoch)	Comment
Parent	false
Week Start Day	x-time-1970
	Session
	Mon

Transform Functionality

Navigation title:Transform Functionality

Information about transformation functionality and the steps to configure the server to export data to external targets.

About Transformation Functionality

Navigation title:About Transformation Functionality

Transform functionality (**Transform**) runs on a data workbench server machine to enable the export of log source data for use by other applications.

Transform can read .vs1 files, log files, XML files, and ODBC data and export the data as .vs1 files, text files, or delimited text files that can be used by DataWarehouse loading routines, auditing agencies, or other targets. The data extraction and transformation can be performed on a continuous or other scheduled basis.



Note: Typically, **Transform** is configured on a data workbench server FSU. However, your implementation may require configuration on an data workbench server DPU. For more information, contact Adobe.

You can view memory usage information for **Transform** in the Detailed Status interface. For more information, see the Administrative Interfaces chapter of the *Data Workbench User Guide*.

The segment export feature provides another means of exporting data from an Adobe application. **Transform** enables you to export unprocessed data to an external target, but the segment export feature enables you to output processed data from the dataset and requires that the exported data be defined as a dimension in the dataset. For more information about segment export, see the *Data Workbench User Guide*.

Configuration Files for Transform

Navigation title:Configuration Files for Transform

Transform functionality includes three configuration files that contain the required parameters for reading and exporting data.

- [The Transform.cfg File](#) on page 618
- [The Transform Mode.cfg File](#) on page 625
- [The Required Field Definitions.cfg File](#) on page 626



Note: Before you begin to configure your implementation of transformation functionality, please confirm that you have properly installed the data workbenchTransform.cfg and data workbench Transform

Mode .cfg files and all of the Adobe software products and applications needed for your implementation. For installation instructions, see the *Server Products Installation and Administration Guide*.

The Transform.cfg File

Navigation title:

The data workbench **Transform.cfg** file contains the parameters that define the log sources, data transformations, and exporters.

The transformations that you define manipulate raw data collected by Sensors (.vs1 files) or contained in text files, XML files, or ODBC-compliant databases and output them either into existing fields, overwriting the current data, or into newly defined fields.

To configure transformation functionality, you edit the data workbench **Transform.cfg** file within the **Dataset** folder for the profile for which you want to export event data. Typically, this profile is dedicated to transformation functionality (that is, you perform no other data processing than what is defined in the data workbench **Transform.cfg** file). It is important to note that any processing instructions specified in the **Log Processing Dataset Include** files for any inherited profiles are applied in addition to those specified in the data workbench **Transform.cfg** file.

For information about dataset include files, see [Dataset Include Files](#) on page 557.

If the data that you want to export is processed by a data workbench server cluster, each of the processing servers (DPUs) in the cluster processes the data, but only the first DPU (processing server #0 in the **profile.cfg** file) will write the output data to its local file system.

To edit the data workbench **Transform.cfg** file

1. While working in the profile for which you want to export data, open the **Profile Manager** and click **Dataset** to show the contents of the directory.
2. Right-click the check mark next to data workbench **Transform.cfg**, then click **Make Local**. A check mark for this file appears in the **User** column.
3. Right-click the newly created check mark and click **Open > from the workbench**. The data workbench **Transform.cfg** window appears.
4. Edit the parameters in the configuration file using the table below as a guide:

Data Workbench Transform.cfg

Parameter	Description
End Time	<p>Optional. Filter data to include log entries with timestamps up to, but not including, this time. Adobe recommends using one of the following formats for the time:</p> <ul style="list-style-type: none">• January 1 2013HH:MM:SS EDT• Jan 1 2013 HH:MM:SS GMT <p>For example, specifying July 29 2013 00:00:00 EDT as the End Time includes data through July 28, 2013, at 11:59:59 PM EDT.</p> <p>You must specify a time zone. The time zone does not default to GMT if not specified. For a list of time zone abbreviations supported by the data workbench server, see Time Zone Codes on page 658.</p> <p>The Use Start/End Times parameter for Sensor and log file sources is related to this parameter.</p>
Exporters	<p>The subfields of an exporter specify how the output data is processed and/or formatted. You can define multiple exporters for a set of log sources. Each exporter type creates output independently.</p> <p>Three types of exporters exist:</p> <ul style="list-style-type: none">• ExportTextFile• ExportDelimitedTextFile• ExportVSLFile <p>For more information about exporter types, see Defining Exporters on page 620.</p>

Parameter	Description
Hash Threshold	Optional. A sampling factor for random sub-sampling of rows. If set to a number n, then only one out of each n tracking IDs are selected for exporting, reducing the total number of rows exported by a factor of n. To export all rows, you would set Hash Threshold to 1.
Log Entry Condition	Optional. Defines the rules by which log entries are considered for export. For more information about the Log Entry Condition , see Log Processing Configuration File on page 528.
Log Sources	The sources of data. Log sources can be .vs1 files, log files, or XML files or data from ODBC-compliant databases. For information about log sources , see Log Processing Configuration File on page 528. Transform expects all source data to be in chronological order within lexicographically ordered input files. If this requirement is not satisfied, As Of calculations are incorrect, and additional input data may be processed after the output files are closed.
Offline Mode	Optional. True or false. If true, Transform assumes that all of the input files are present when it starts processing the data. When all of the input data has been read, Transform closes all of the output files without waiting for additional data to be received. The default value is false.  Note: If Offline Mode is set to true, Transform expects all source data to be present before processing starts. A warning message is generated in the <code>VisualServer.log</code> file if additional data is received after the output files are closed.
Reprocess	Optional. Any character or combination of characters can be entered here. Changing this parameter and saving the file to the Transform machine initiates data reprocessing. For information about reprocessing your data, see Reprocessing and Retransformation on page 627.
Stages	Optional. The names of the processing stages that can be used in Log Processing Dataset Include files that are executed in addition to the data workbenchTransform.cfg file. Processing stages provide a way to order the transformations that are defined in Log Processing Dataset Include files. This parameter is very helpful if you have defined one or more transformations within multiple Log Processing Dataset Include files and you want specific transformations to be performed at specific points during the export process. The order in which you list the stages here determines the order in which the transformations in the Log Processing Dataset Include files are executed during data export. Preprocessing and Postprocessing are built-in stages; Preprocessing is always the first stage, and Postprocessing is always the last stage. By default, there is one named stage called Default . To add a new processing stage <ul style="list-style-type: none">• In the data workbenchTransform.cfg window, right-click Stages, then click Add New > Stage.• Enter a name for the new stage. To delete an existing processing stage <ul style="list-style-type: none">• Right-click the number corresponding to the stage that you want to delete and click Remove<#stage_number>.  Note: When you specify a Stage in a Log Processing Dataset Include file the name of the stage must match exactly the name that you enter here. For more information about dataset include files, see Dataset Include Files on page 557.
Start Time	Optional. Filter data to include log entries with timestamps at or after this time. Adobe recommends using one of the following formats for the time: <ul style="list-style-type: none">• January 1 2013 HH:MM:SS EDT• Jan 1 2013 HH:MM:SS GMT For example, specifying July 29 2013 00:00:00 EDT as the Start Time includes data starting from July 29, 2013, at 12:00:00 AM EDT. You must specify a time zone. The time zone does not default to GMT if not specified. For a list of time zone abbreviations supported by the data workbench server, see Time Zone Codes on page 658.

Parameter	Description
	 Note: The Use Start/End Times parameter for Sensor and log file sources is related to this parameter.
Transformations	<p>Optional. Defines the transformations that are to be applied to the data. For information about the available transformation types, see Data Transformations on page 573.</p>  Note: The following transformation types do not work when defined in the data workbenchTransform.cfg file: <ul style="list-style-type: none"> • AppendURI • CrossRows • ODBCLookup • Sessionize



Note:

If additional data is received after the output files are closed (see **Log Sources** and **Offline Mode** in the preceding table), **Transform** creates new output files with the additional data. The names of the new output files are generated from the original output file's name with the addition of a parenthesized version number just before the extension. For example, if the original output file is 20070701-ABC.vs1, subsequent versions of this file will be named 20070701-ABC(1).vs1, 20070701-ABC(2).vs1, and so on. Note that using the versioned files as input to the data workbench server may result in processing errors.

Adobe recommends avoiding the creation of versioned output files by making sure that all source data is in chronological order within lexicographically ordered input files and, if **Offline Mode** is set to true, that all source data is present before processing starts. For more information, see the **Log Sources** and **Offline Mode** entries in the preceding table.

5. Add transformations by right-clicking **Transformations** and clicking **Add new > <Transformation type>**. Complete the transformation fields.

See [Data Transformations](#) on page 573 for descriptions and examples of the transformations that you can use with transformation functionality.

6. Right-click (**modified**) at the top of the window, then click **Save**.
7. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for data workbench Transform.cfg in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the profile for which you are exporting data. Reprocessing of the data begins after synchronization of the profile.



Note: For information about reprocessing your data for export, see [Reprocessing and Retransformation](#) on page 627.

Defining Exporters

Navigation title:Defining Exporters

Exporters provide the instructions for outputting the **event data**.

Transformation functionality provides three types of exporters for exporting .vs1 files, log files, XML files, and ODBC data as .vs1 files, text files, or delimited text files that can be used by DataWarehouse loading routines, auditing agencies, or other targets.



Note: For an exporter to work properly, the log source must meet the appropriate requirements discussed in the [Log Sources](#) on page 532 section of [Log Processing Configuration File](#) on page 528.

To define an exporter

1. Open `Transform.cfg` in data workbench. See [The Transform.cfg File](#) on page 618.
2. Right-click **Exporters**, then click **Add New**.
3. Select one of the following options:
 - **ExportTextFile**
 - **ExportDelimitedTextFile**
 - **ExportVSLFile**



Note: For the **ExportVSLFile** option, all of the extended fields in the input file and all user-defined fields of the form `cs(header)` are always written to the VSL output file. If you overwrite an existing extended field, the new value is written to the output file, even if the field is blank.

4. Edit the Exporters parameters in the configuration file using the following table as a guide:

Exporter Parameters

Parameter	Description
Data Format	For ExportTextFile only. The format of each output line, consisting of field name escapes (expressed as <code>%fieldname%</code>) and any other desired fixed text. The format should include a line separator, typically [CR] [LF]. A literal percent sign (%) can be embedded in the format string by escaping the character as shown here: <code>%%</code> An example of an entry for the Data Format parameter is <code>%x-timestring% %x-trackingid% [CR] [LF]</code> .
Fields	For ExportDelimitedTextFile only. Names of the fields to be output.
Delimiter	Optional. For ExportDelimitedTextFile only. Character that is used to separate the fields in the output file. The software can not escape delimiters that are included in the data's values. As a result, Adobe does not recommend using commas as delimiters. If you hold down the Ctrl key and right-click within the Delimiter parameter, an Insert menu appears. This menu contains a list of special characters that often are used as delimiters.
Line Separator	Optional. For ExportDelimitedTextFile only. The character(s) used to separate lines in the output files. The default value is [CR] [LF].
Name	Optional. Identifier for the exporter. This name appears in the Detailed Status interface. For information about the Detailed Status interface, see the Data Workbench User Guide .
Comments	Optional. Notes about the exporter.
Output Path	Path where output files are to be stored. The path is relative to the data workbench server installation folder.  Note: The data workbench server that stores output data is processing server #0 in the <code>profile.cfg</code> file.
File Rotation Period	Optional. The frequency at which data is exported to the output file. Each output file contains data related to a specific time period called the rotation period. All time calculations are in GMT: One day starts at midnight GMT and ends the following day at midnight GMT, even if the data written to the file includes a field which has been transformed to local time. Available values are as follows:

Parameter	Description
	<ul style="list-style-type: none"> YEAR. Each file contains data for one calendar year. MONTH. Each file contains data for one calendar month. Months are numbered 1 (January) through 12 (December). WEEK. Each file contains data for one week. A week starts on Monday. The week starting on one of the first seven days of the year is week 1, and the prior (partial) week, if any, is week 0. DAY. Each file contains data for one calendar day. HOUR. Each file contains data for one hour. NONE. No rotation is performed. All of the data is written to the same file (or a set of files as determined by other parameter settings). See the File Name Format parameter in this table. <p>The default file rotation period is DAY.</p> <ul style="list-style-type: none"> Set the file rotation to NONE only when working in Offline Mode. See the <i>The Transform.cfg File</i> on page 618 parameter description.
File Name Format	<p>Optional. The format of the output file name.</p> <p>Each log entry can be stored in a file whose name is derived from the start time of the rotation period, and optionally, from values of fields in the rows that it contains. The fields to use in the file name are embedded as field name escapes (expressed as %<i>fieldname</i>%).</p> <p>The file name components related to the rotation period are embedded in the format string using the following escape sequences:</p> <ul style="list-style-type: none"> %yyyy% (four-digit year) %yy% (two-digit year) %mm% (two-digit month, 01 - 12) %ww% (two-digit week, 01 - 52) %dd% (two-digit day, 01 - 31) %HH% (two-digit hour, 00 - 23) <p>The default file name format is %yyyy%mm%dd%-%x-mask%.txt</p> <ul style="list-style-type: none"> The escape sequences are case-sensitive. When File Rotation Period is set to NONE, an empty string is substituted for each of the escape sequences, if present. An error is generated if File Name Format does not result in a unique file name for each rotation period (see the File Rotation Period parameter in this table). For example, when using the DAY rotation period, the %dd%, %mm%, and %yy% or %yyyy% escape sequences must be present in the pattern to avoid data loss. If you are using field name escape sequences within the pattern and the given field has many distinct values, many output files are written for each rotation period. Note that this scenario may result in poor performance, so you should use this feature with caution. All times calculations are in GMT.
Execute at Rollover	<p>Optional. When each file is finalized, an external (Windows) command can be executed. The command line is derived from the final file name by substituting the following escape sequences into this parameter:</p> <ul style="list-style-type: none"> %dir%. The directory part of the final file name, including the trailing backslash. %file%. The file name (excluding the directory and extension) of the final file. %ext%. The extension (including the leading ".") of the final file name. %path%. The full path name of the file, equivalent to %dir%%file%%ext%. <p>By default, this parameter is empty (no command is executed).</p>
Memory Limit	<p>Optional. The amount of memory in bytes used for buffering the exporter's output. The default value is 10,000,000 bytes.</p> <p> Note: If you have many output files that are open at the same time, you may want to increase this value, but you may decrease the amount of memory available for use by other components of the system. Decreasing this value, however, may slow down the export process. For assistance, contact Adobe.</p>

Parameter	Description
Open Files Limit	Optional. The maximum number of files that may be open at the same time for output from the exporter. If this number is exceeded, an error is recorded in the event log and the data workbench server stops running. The default value is 1000.

5. After you have defined your exporter (and made changes to other parameters) in the `Transform.cfg` file, save the file locally and save it to the appropriate profile on the data workbench server machine.

Sample Data Workbench Transform.cfg Files

Navigation title: Sample Insight Transform.cfg Files

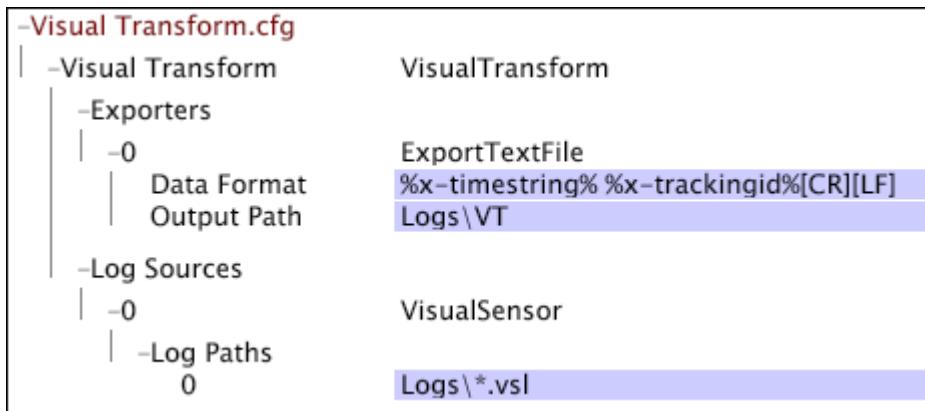
Information about how to specify parameters in the `Transform.cfg` file based on the several scenarios.

- [A Simple Data Workbench Transform.cfg File](#) on page 623
- [Output with Comma-Separated Values](#) on page 623
- [Sample Log Files](#) on page 624
- [Splitting Log Files by Web Site Section](#) on page 625

In each sample, the file is displayed as a `Transform.cfg` window in data workbench.

A Simple Data Workbench Transform.cfg File

The following `Transform.cfg` window provides instructions to read `.vs1` files from the `Logs` directory and export the `x-timestamp` and `x-trackingid` fields to a text file stored in the `Logs\VT` directory. Because no file rotation period or output file name format is specified, each file contains data for one calendar day and has a name in the default format `%yyyy% %mm% %dd% - %x-mask%.txt`.



Output with Comma-Separated Values

The following `Transform.cfg` window provides instructions to read `.vs1` files from the `Logs` directory and export fields 0 through 13 to a comma-delimited (`.csv`) file stored in the `Logs\VT\CSV` directory. Because no file rotation period is specified, each file contains data for one calendar day. The output files are `.csv` files named in the format `%yyyy% %mm% %dd% - %x-mask%.csv`.

-Visual Transform	VisualTransform
-Exporters	
-0	ExportDelimitedTextFile
-Fields	
0	x-timestring
1	x-mask
2	x-trackingid
3	c-ip
4	sc-bytes
5	sc-status
6	cs-method
7	cs-uri-stem
8	cs-uri-query
9	s-dns
10	cs(referrer)
11	cs(cookie)
12	cs(user-agent)
13	sc(content-type)
Delimiter	,
File Name Format	%yyyy%-%mm%-%dd%-%x-mask%.csv
Output Path	Logs\VT\CSV
-Log Sources	
-0	VisualSensor
-Log Paths	
0	Logs*.vsl

Sample Log Files

You can configure transformation functionality to create and maintain an up-to-date, compact version of your full log files. Doing so enables you to test your dataset configurations quickly, with reprocessing times of seconds or minutes instead of hours needed to reprocess the entire dataset. The following example provides an example of how to configure transformation functionality to do this.

The following `Transform.cfg` window provides instructions to read `.vsl` files from the `Logs` directory and export the `x-timestring` and `x-trackingid` fields to a text file stored in the `Logs\VT` directory. The specified Hash Threshold filters certain tracking IDs from the dataset, thereby creating a dataset that is sampled by a factor of 100. Because no file rotation period is specified, each file contains data for one calendar day. The names of the output files are in the default format `%yyyy%-%mm%-%dd%-%x-mask%.txt`.

-Visual Transform.cfg	
-Visual Transform	VisualTransform
-Exporters	
-0	ExportTextFile
Data Format	%x-timestring% %x-trackingid%[CR][LF]
Output Path	Logs\VT
-Log Sources	
-0	VisualSensor
-Log Paths	
0	Logs*.vsl
Hash Threshold	100

Splitting Log Files by Web Site Section

The following `Transform.cfg` window provides instructions to read `.vsl` files from the `Logs` directory and export the `x-timestring` and `x-trackingid` fields to a text file stored in the `Logs\VT` directory. The regular expression transformation (**RETransform**) takes as its input the `cs-uri-stem` field and creates a new field (`x-site`) that defines a section of the site. The `x-site` field is included in the name of output text files, each of which contain data for one calendar day.

-Visual Transform.cfg	
-Visual Transform	VisualTransform
-Exporters	
-0	ExportTextFile
Data Format	%x-timestring% %x-trackingid%[CR][LF]
File Name Format	%x-site%-%yyyy%-%mm%%dd%.txt
Output Path	Logs\VT
-Log Sources	
-0	VisualSensor
-Log Paths	
0	Logs*.vsl
-Transformations	
-1	RETransform
Action	RESULTS
Default	MAIN
Expression	/?([^/]*)/*
Input	cs-uri-stem
Output	x-site

The `Transform Mode.cfg` File

Navigation title:

The configuration file `Transform Mode.cfg` enables you to pause processing of data into a dataset, specify offline sources, or specify the frequency at which the data workbench server running transformation functionality saves its state files.

Making changes to the `Transform Mode.cfg` file, including adding or removing sources, does not cause reprocessing of the data.

To edit the Transform Mode.cfg file for a dataset profile

1. While working in the profile whose data you want to export, open the **Profile Manager** and click **Dataset** to show the contents of the directory.
2. Right-click the check mark next to `Transform Mode.cfg` and click **Make Local**. A check mark for this file appears in the **User** column.
3. Right-click the newly created check mark and click **Open > from the workbench**. The `Transform Mode.cfg` window appears.
4. Edit the parameters in the configuration file using the following table as a guide:

Transform Mode.cfg

Parameter	Description
Offline Sources	<p>Mask of the offline log source.</p> <p>To specify an offline source:</p> <ul style="list-style-type: none">• Right-click Offline Sources and click Add new > Source.• In the parameter for the new source, enter the mask of the log sequence. For Sensor log sources with file names of the format <code>YYYYMMDD-SENSORID.vsl</code>, the mask is <code>SENSORID.SENSORID</code> is case-sensitive. For log file log sources, the mask is the string extracted by the Mask Pattern (see Log Files on page 534). <p>Adding or removing sources from Offline Sources does not cause reprocessing of the dataset.</p> <p>As Of time measurements are maintained for the processing of the profile's online sources. When the offline source is again online, the processing of incoming log files for that source resumes.</p> <p> Note: Whenever a source comes back online, you should remove it from Offline Sources. If you do not do so, the data workbench server treats the source as an online source and updates the As Of time as long as the source is sending data. If the source goes offline again, the As Of time measurements stop.</p>
Paused	True or false. If true, new data is not processed into the dataset. The default value is false.
Save Interval (sec)	Frequency at which the data workbench server on which transformation functionality is running saves its state files. The default value is 3600. <p> Note: You should not change this value without consulting Adobe.</p>

When editing the `Transform Mode.cfg` file within a data workbench window, you can use shortcut keys for basic editing features, including cut (Ctrl+x), copy (Ctrl+c), paste (Ctrl+v), undo (Ctrl+z), redo (Ctrl+Shift+z), select section (click+drag), and select all (Ctrl+a). In addition, you can use the shortcuts to copy and paste text from one configuration file (.cfg) to another.

5. Right-click (**modified**) at the top of the window and click **Save**.
6. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the data workbench `Transform Mode.cfg` in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the profile for which you are exporting data. Reprocessing of the data begins after synchronization of the profile.

For information about reprocessing your data for export, see [Reprocessing and Retransformation](#) on page 627.

The Required Field Definitions.cfg File

Navigation title: The Required Field Definitions.cfg File

The Required Field Definitions.cfg file is a **Log Processing Dataset Include** file that defines the required fields x-trackingid and x-timestamp.

The file is located in the **Transform** profile's **Dataset\Log Processing** folder. For more information about defining fields, see [Log Processing Dataset Include Files](#) on page 559.

Reprocessing and Retransformation

Navigation title: Reprocessing and Retransformation

Information about dataset reprocessing and retransformation and how to prepare for either.

Understanding Reprocessing and Retransformation

Navigation title: Understanding Reprocessing and Retransformation

During reprocessing, the data workbench server reconstructs your dataset as you have specified in the **Log Processing** and **Transformation Dataset Configuration** files.

To do so, the data workbench server (InsightServer64.exe) must complete both the log processing phase and the transformation phase of dataset construction. When log processing finishes, it triggers transformation to occur automatically, but transformation also can occur independently of log processing.

During the log processing phase, data workbench users do not have access to the data in the dataset. During the transformation phase, data workbench users do have access to up-to-date data, but the data is sampled instead of complete. Data analysis can continue during transformation, but queries will complete only as quickly as the transformation is occurring.

For more information about the log processing and transformation phases of dataset construction, see [Dataset Configuration](#) on page 524.

This section discusses the following topics:

- [Reprocessing](#) on page 627
- [Retransformation](#) on page 627

Reprocessing

Each time you complete one of the following tasks, log processing, and therefore transformation, occurs automatically to reconstruct your dataset as you have specified in the dataset configuration files:

- Add a new data source.
- Add a new data workbench server to your cluster in the `Profile.cfg` file.
- Change the `Cluster.cfg` file.
- Change the `Log Processing.cfg` file or a **Log Processing Dataset Include** file, including but not limited to the following:
 - Add a new parameter
 - Change a transformation
 - Change the Start Time or End Time parameters
- Upgrade your `Insight Server.exe` file.

You also can initiate reprocessing at any time by entering any character or combination of characters in the `Reprocess` parameter of the `Log Processing.cfg` file and saving the file.



Note: For reprocessing to occur, the `Pause` parameter in the `Log Processing Mode.cfg` file must be set to false. This parameter's default value is false, so changing the parameter may not be required. For more information about `Log Processing Mode.cfg`, see .

Retransformation

Each time you change any information in the `Transformation.cfg` file or in a **Transformation Dataset Include** file, such as change a transformation or define a new dimension, transformation occurs automatically.

Each time you change lookup files that are referenced in the Transformation.cfg file or in a **Transformation Dataset Include** file (including lookup files for **Categorize**, **FlatFileLookup**, and **ODBCLookup** transformations), you must initiate transformation by entering any character or combination of characters in the Reprocess parameter of the Transformation.cfg file and saving the file.

Preparing for Reprocessing or Retransformation

Navigation title:Preparing for Reprocessing or Retransformation

Steps to ensure that reprocessing or retransformation goes smoothly and finishes in time for data workbench users to get back to work

1. Determine the elapsed time of previous processing or transformation by checking the dataset profile's **Processing Mode History** in the **Detailed Status** interface.
 1. While working in your dataset profile, open the **Detailed Status** interface.
 2. Click **Processing Status** > <*dataset profile name*> > **Processing Mode History** to view the elapsed times of previous processing or transformation.
 - Fast Input is the total time needed for log processing.
 - Fast Merge is the total time needed for transformation.
 - The sum of the two times (Fast Input + Fast Merge) is the total time needed for processing the dataset.

The example below indicates that log processing took approximately 43 seconds, while transformation took less than 2 minutes.

-Processing Mode History	
0	Real Time for 00:00:01 (hh:mm:ss)
1	Fast Input for 00:00:43 (hh:mm:ss)
2	Fast Merge for 00:01:40 (hh:mm:ss)

For more information about the **Detailed Status** interface, see the *Data Workbench User Guide*.

2. Schedule and plan the reprocessing. Because data workbench users do not have access to the data during the log processing phase, ensure that you schedule reprocessing to occur during an appropriate time, such as over the weekend.
3. Monitor the progress of the reprocessing and retransformation using the fields in the **Processing Legend**. For more information about the **Processing Legend**, see the *Data Workbench User Guide*.

Dataset Configuration Tools

Navigation title:Dataset Configuration Tools

Information about the data workbench interfaces that enable you to monitor or manage your dataset and the options for hiding dataset components.

Working With Dataset Configuration Interfaces

Navigation title:Working With Dataset Configuration Interfaces

The dataset configuration interfaces show the relationships among the components (log sources, fields, transformations, and dimensions) of your dataset.

Some of the interfaces provide options for viewing information about, searching within, or editing dataset components.

To open the dataset configuration interfaces

You can open the interfaces for dataset configuration using one of the following methods:

- **Admin Menu:** Right-click within a workspace, click **Admin**, then click the appropriate menu option.

-OR-

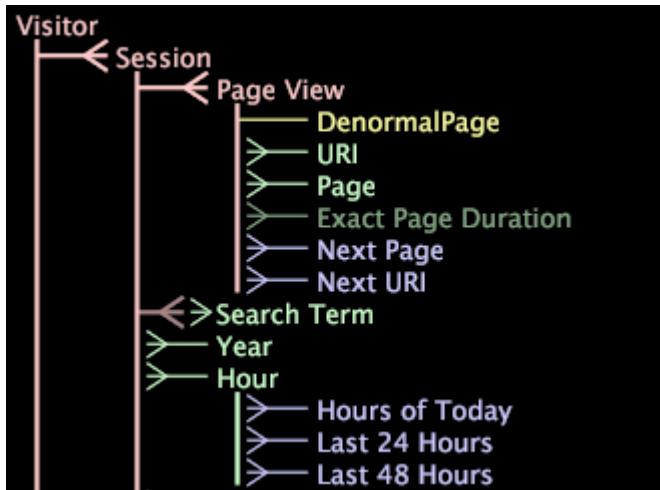
- **Admin Tab:** Click to open the appropriate workspace thumbnail on the **Admin > Dataset and Profile tab**.

Dataset Schema

Navigation title:Dataset Schema

The **Dataset Schema** interface displays the extended dimensions (countable, simple, many-to-many, numeric, denormal, and time dimensions) defined in any **Transformation Dataset Configuration** file and the relationships between those dimensions.

In addition, the **Dataset Schema** interface shows any derived dimensions that you have defined, as well as any extended dimensions that are configured to be hidden.



This section discusses the following topics:

- [To Interpret Dimension Type Using the Dataset Schema Interface](#) on page 629
- [To Display the Default Visualization for a Dimension](#) on page 630
- [To Display a Specific Visualization for a Dimension](#) on page 630

To Interpret Dimension Type Using the Dataset Schema Interface

The following table lists the dimension types and the colors in which their names appear in the **Dataset Schema** interface. Parents for the sample dimensions (from the example above) are noted as well.

Dataset Schema Interface

Dimension Type	Color	Sample Dimension and Parent
Countable	Pink	Visitor - In this schema, Visitor is a root countable dimension. Session - parent is Visitor.
Denormal	Yellow	DenormalPage - parent is Page View.
Derived	Blue	Next Page - parent is Page View.
Many-to-Many	Pink and Green (The stem from the parent is pink, while the dimension name is green.)	Search Term - parent is Session.
Numeric	Green	Exact Page Duration - parent is Page View In this example, Exact Page Duration is a hidden numeric dimension. See the Hidden dimension type in this table.

Dimension Type	Color	Sample Dimension and Parent
Simple	Green	Page - parent is Page View .
Time	Green	Hour - parent is Session.
Hidden	Hidden dimensions are a darker version of the appropriate dimension type color. For example, a hidden numeric dimension is a darker, less bright green.	Exact Page Duration - parent is Page View.

To Display the Default Visualization for a Dimension

- In the **Dataset Schema** interface, click the desired dimension. The default visualization displays. For example, if the default visualization is a table displaying Sessions and the selected dimension, and you click the URI dimension, data workbench displays a table with URI by Sessions.



Note: If you want to change the default visualization that displays, see the Configuring Interface and Analysis Features chapter in the *Data Workbench User Guide*.

To Display a Specific Visualization for a Dimension

- In the **Dataset Schema** interface, right-click the desired dimension and click **Add Visualization > <visualization type>**.

Dependency Maps

Navigation title:

Dependency maps enable you to visualize and manage the configuration of the components of your profile.

- Dataset components:** Log sources, filters, fields, transformations, and extended dimensions defined in your dataset's `Log Processing.cfg`, `Transformation.cfg`, and **dataset include** files.
- Query model components:** Metrics, dimensions, and filters defined in the Dimensions, Metrics, and Filters folders.
- Workspaces and visualizations:** Workspaces, reports, menu options, and globe layers.

For more information about working with query model components, workspaces, and visualizations in dependency maps, see the *Data Workbench User Guide*.

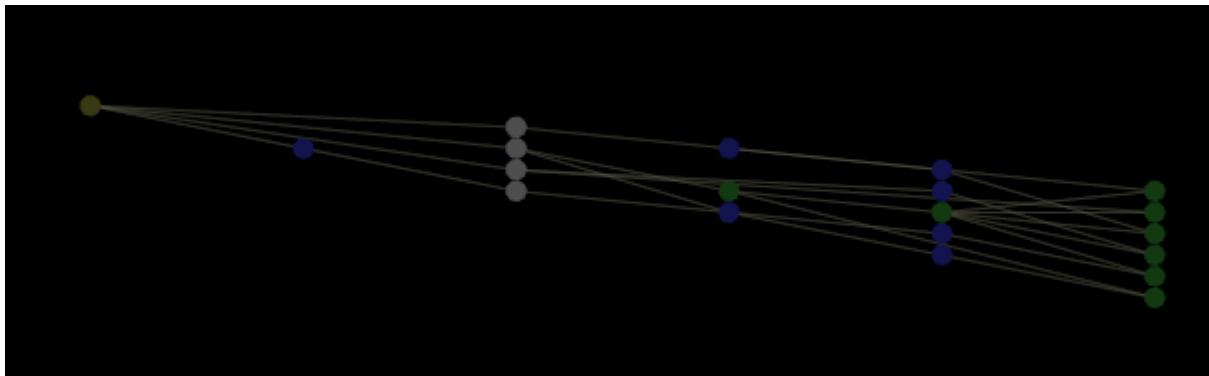
Profile components are represented by colored dots (nodes) in the map. The lines connecting the nodes depict dependencies, that is, how the components relate to one another. A line between two nodes means that an output of the node on the left is an input of the node on the right, that is, the right node depends on the left node.

Displaying Dataset Components

- Right-click within the dependency map and click **Display**.
- Choose **Dataset**. An X appears to the left of **Dataset**.

For more information about the other display options see the *Data Workbench User Guide*.

The following figure shows a dependency map whose nodes represent a dataset's log sources, fields, transformations, and extended dimensions.

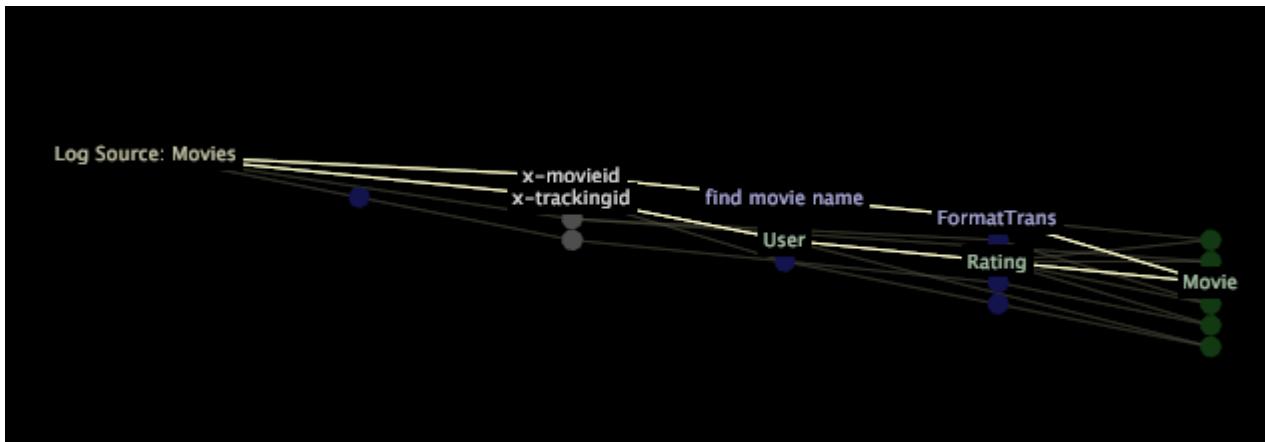


- A yellow-green node represents one or more log sources or a filter defined in the dataset. A node for a log source always appears furthest to the left in the map.
- A gray node represents a field that is listed in the Fields parameter in a `Log Processing.cfg` or **Log Processing Include** file.
- A blue node represents a transformation.
- A green node represents an extended dimension.

 **Note:** If your dataset has a single log source, the map displays Log Source: *log source name*. If your dataset has multiple log sources, the map displays *number* Log Sources, where number is the count of log sources. For example, if you have three log sources in your dataset, your map displays 3 Log Sources.

If you cannot see all of the nodes on the map, you can move the map or zoom in or zoom out to display the entire map or to focus on a particular section. For more information about zooming, see the Working with Visualizations chapter of the *Data Workbench User Guide*.

When you click a node, all of the nodes that depend on that node and all of the nodes on which that node depends are highlighted and their names display.



 **Note:** A highlighted path in a dependency map does not constitute a selection.

When you right-click a node, you can see identifying information about each component shown on the map and choose menu options that enable you to view more detail about the component or to edit the component. In addition, you can perform text searches and display performance information for transformations and extended dimensions.

For information about these functions for dependency maps, see the Administrative Interfaces chapter of the *Data Workbench User Guide*.

Field Viewers

Navigation title:Field Viewers

A field viewer is a table containing the values of one or more data fields.

See [Field viewer](#) on page 474 in *Data Workbench Client*.

Hiding Dataset Components

Navigation title:Hiding Dataset Components

As you configure your dataset, you might want to hide certain extended dimensions so they do not show on the dimension menu in data workbench.

You might want to override certain dataset configuration files in a profile that you cannot or do not want to modify. The following sections provide instructions to perform these tasks.

Hiding Extended Dimensions

Navigation title:Hiding Extended Dimensions

You can use either the Hidden parameter or the Show parameter to hide extended dimensions so they do not show on the dimension menu in data workbench.

When you enter the appropriate setting for either parameter, the dimension name is not listed in the menu in data workbench, but it is still in the profile and available to be used. Any data workbench user can temporarily unhide hidden dimensions by setting the Unhide All parameter in the `Insight.cfg` file to true.

For more information about the Unhide All parameter, see the appendix on data workbench configuration parameters in the *Data Workbench User Guide*.

The following sections describe how to use the Hidden and Show parameters to hide extended dimensions.

- [Hiding Extended Dimensions Using the Hidden Parameter](#) on page 632
- [Hiding Extended Dimensions Using the Show Parameter](#) on page 632

Hiding Extended Dimensions Using the Hidden Parameter

The Hidden parameter is an optional parameter that you can use when defining extended dimensions in **Transformation Dataset Configuration** files.

1. Open **Transformation Dataset Configuration** files in which the extended dimension that you want to hide is defined. See [Editing Existing Dataset Include Files](#) on page 557.
2. Locate the Hidden parameter for the desired dimension in the configuration window and type *true*.
3. Save the file locally, then save it to the appropriate profile on the server. See [Editing Existing Dataset Include Files](#) on page 557.

The dataset retransforms, after which the extended dimension does not appear on the dimension menu in data workbench. For more information about the Hidden parameter, see [Extended Dimensions](#) on page 606.

If you change the Hidden parameter's setting, you must retransform the dataset for the change to take effect.

Hiding Extended Dimensions Using the Show Parameter

The Show parameter is not one of the parameters available for defining extended dimensions in **Transformation Dataset Configuration** files. Instead, the parameter is defined in the `.dim` files for any derived dimensions that you create. Therefore, to use the Show parameter to hide an extended dimension, you first must create a derived dimension that is based on the extended dimension as described in the following procedure:

1. Use a text editor such as Notepad to create an empty file called `<dimension name>.dim` The file name should match the name of the dimension that you want to hide. For example, to hide the Next Page dimension, you would create a `Next Page.dim` file.
2. Add the following text to the file:

- entity = ref: wdata/model/dim/Parent/+name
- show = bool: false

3. Save the file to the profile's Dimensions directory. You can save the file to a subdirectory if desired.

When you use the Show parameter to hide an extended dimension, you do not have to retransform your dataset to make the change take effect. You can choose to hide or show the dimension in your local version of the profile (that is, you can save the .dim file to your User folder), or you can save the .dim file to the server for use by other users of the profile.

You also can use the Show parameter to hide metrics and filters. For information, see the Configuring Interface and Analysis Features chapter in the *Data Workbench User Guide*.

Hiding Dataset Configuration Files

Navigation title:Hiding Dataset Configuration Files

If you do not want to inherit a configuration file from an internal or other inherited profile (that is, you want the instructions in the file to be ignored during dataset construction), but you do not want to modify the file, you can create an empty (zero-byte) file with the same name and store the file in another profile.

To zero-byte a dataset configuration file

1. In the **Profile Manager**, open the necessary folders and subfolders to locate the file that you want to zero-byte.
2. Right-click the check mark next to the name of the file and click **Make Local**.
3. Open the local file in a text editor such as Notepad and delete its contents.
4. Save and close the file.
5. In the **Profile Manager**, save the zero-byte file to a profile to the right of the profile in which the original file resides. (You want the zero-byte file to take precedence over the original file.)

In the **Profile Manager**, a hyphen (-), instead of a check mark, in a column identifies the zero-byte file as shown in the example below.

File	Base	Movie Data	User
profile.cfg		✓	
Context\	✓		
Dataset\			
Client.cfg	✓		
Cluster.cfg	✓		
Log Processing.cfg		✓	
Log Processing Mode.cfg	✓		
Server.cfg	✓		
Transformation.cfg		✓	✓
Log Processing\		✓	✓
TimeZone\	✓		
Transformation\			
Movie Ratings.cfg	✓	-	
Movie Transformation.cfg		✓	
Dimension Legend\	✓		
Dimensions\	✓		
FileEditor\	✓		
Filters\		✓	
Images\	✓		
Menu\	✓	✓	
Metrics\	✓	✓	✓

When you reprocess your dataset, the dataset does not contain the dataset components that the original file defines.



Note: If you zero-byte a configuration file that defines an extended dimension that is used in a visualization or a metric definition, data workbench produces an error for that visualization or metric, respectively.

You can also use zero-byte files to move a metric, dimension, or filter to another location in the profile or to hide menu items. For information, see the *Data Workbench User Guide*.

Conditions

Navigation title:Conditions

Information about the types of condition operations that you can use when configuring your dataset.

About Conditions

Navigation title:About Conditions

Transformations and dimensions use conditions to determine when certain instructions or actions apply to log fields.

The Log Entry Condition parameter uses conditions to determine which log entries should be included in the dataset construction process. This appendix describes the different types of conditions available within data workbench server.

A condition falls into one of two categories:

- **Test Operations:** Compare, Not Empty, Range, Regular Expression, and String Match conditions are used to test for different states in the available log fields.
- **Boolean Operations:** The And, Or, and Neither conditions are used to combine the test operations described above. For example, if you have a Range condition and a String Match condition that must both be false to take the appropriate action, you would make these two operations children of the Neither condition. Note that the And condition is used as the root of all condition testing in the system.

Working With Conditions

Navigation title:Working With Conditions

Information about adding, removing, or copying a condition.

- [To Add a Condition to a Dataset Configuration File](#) on page 634
- [To Remove a Condition from a Dataset Configuration File](#) on page 635
- [To Copy a Condition](#) on page 635

To Add a Condition to a Dataset Configuration File

1. While working in your dataset profile, use the **Profile Manager** to open the dataset configuration file that you want to edit.
 - To open the Log Processing.cfg file, see [Editing the Log Processing Configuration File](#) on page 529.
 - To open the Transformation.cfg file, see [Editing the Transformation Configuration File](#) on page 551.
 - To open a **Dataset Include** file, see [Dataset Include Files](#) on page 557.
2. Within the dataset configuration file, locate the Log Entry Condition or Condition parameter that you would like to define.
3. Right-click the parameter and click **Add new**. Choose one of the following condition types:
 - **Not Empty**
 - **String Match**
 - **Regular Expression**
 - **Range**

- And
- Or
- Neither
- Compare

4. Edit the parameters of the condition as desired. For descriptions of the parameters of each condition, see the appropriate section of this appendix.
5. To save your changes, right-click (**modified**) at the top of the window and click **Save**.
6. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the file belongs.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

To Remove a Condition from a Dataset Configuration File

1. Right-click the name of the condition or the number corresponding to the condition that you want to remove.
2. Click **Remove <#number>**, where number is the number corresponding to the condition that you want to remove.

To Copy a Condition

You can copy a condition from one location to another location in the same file, or you can copy a condition from one dataset configuration file to another.

1. Right-click the name of the condition or the number corresponding to the condition that you want to copy and click **Copy**.
2. Right-click the name or number of the condition below which you want to place the copied condition and click **Paste**.

Test Operations

Navigation title:Test Operations

Test operations take input from the available log fields, perform a test based upon their values, and return true if the test succeeded or false if the test failed.

These tests are always the child of one of the **Boolean** operations described in the next section. See [Boolean Operations](#) on page 640.



Note: If the input is a vector of strings, only the first value (string) in the vector is used for the test. This rule applies to all of the test operations.

Test Operation Conditions

Navigation title:Test Operation Conditions

Information about test operation conditions including compare, not empty, range, regular expression, and string match.

- [Compare](#) on page 636
- [Not Empty](#) on page 636
- [Range](#) on page 637
- [Regular Expression](#) on page 638
- [String Match](#) on page 638

Compare

The **Compare** condition compares string or numeric values. For comparisons of string values, you can specify whether case should be considered.

The parameters of the **Compare** condition are described in the following table:

Parameter	Description	Default
Case Sensitive	True or false. Used only if the Type is LEXICAL . If set to false, upper and lower case letters are considered equal.	true
Comments	Optional. Notes about the condition.	Comments
Input A	The first of the two values to compare. This value represents the left operand in the condition.	
Input B	The second of the two values to compare. This value represents the right operand in the condition.	
Operation	The comparison operation. The available operations (and their meanings) are as follows: <ul style="list-style-type: none">• = or == (Input A equals Input B)• <> or != (Input A is not equal to Input B)• < (Input A is less than Input B)• <= (Input A is less than or equal to Input B)• > (Input A is greater than Input B)• >= (Input A is greater than or equal to Input B)	=
Type	The type of comparison to be made. Available types are LEXICAL , NUMERIC , and DATETIME . For descriptions of the types, see Test Types for Test Operations on page 639.	LEXICAL

This example uses a **Compare** condition to define the **Log Entry Condition**. As the data workbench server reads each event data record, it compares the numeric values x-age and 55. If for a given log entry, x-age is less than or equal to 55, the log entry is included in the dataset construction process.

-Log Entry Condition	AndCondition
-0	CompareCondition
Case Sensitive	true
Comments	Comment
Input A	x-age
Input B	"55"
Operation	<=
Type	NUMERIC

Not Empty

The **Not Empty** condition checks a field to see if it contains a value or is empty. The condition is satisfied for any log entry whose value for the **Input** field is not empty.

The parameters of the **Not Empty** condition are described in the following table:

Parameter	Description	Default
Comments	Optional. Notes about the condition.	Comments
Input	The name of the field from the log entry to check for content.	

This example takes as its input x-some-field and tests whether the field is not empty. The condition is satisfied if the field is populated.

	-1	Comments Input	NotEmptyCondition Comment x-some-field
--	----	-------------------	--

Range

The **Range** condition takes an input field and determines whether the value of that field falls, inclusively, within the given minimum (Min) and maximum (Max) parameter values.

The parameters of the **Range** condition are described in the following table:

Parameter	Description	Default
Case Sensitive	True or false. Used only if the Type is LEXICAL . If set to false, upper and lower case letters are considered equal.	true
Comments	Optional. Notes about the condition.	Comments
Input	The name of the field from the log entry to use as input.	
Min	<p>Lower bound of range.</p> <p>This parameter's value must be a literal value or a string - not a field name. If you use a date for this field, you must specify a time zone. For a list of supported time zone abbreviations, see Time Zone Codes on page 658.</p>	
Max	<p>Upper bound of range.</p> <p> Note: This parameter's value must be a literal value or a string - not a field name. If you use a date for this field, you must specify a time zone. For a list of supported time zone abbreviations, see Time Zone Codes on page 658.</p>	
Type	<p>The type of comparison to be made. Available types are LEXICAL, NUMERIC, and DATETIME.</p> <p>For descriptions of the types, see Test Types for Test Operations on page 639.</p>	

This example uses a **Range** condition to define the **Log Entry Condition**. As the data workbench server reads each **event data** record, it compares the numeric values x-age and 55. If for a given log entry, x-age is at least 55, the log entry is included in the dataset construction process. This example performs the same function as the **Compare** condition example. See [Compare](#) on page 636.



Note: If the Min or Max parameter is left blank, the data workbench server substitutes the minimum or maximum integer values available. The minimum value is zero (0), and the maximum value is infinity.

	-Log Entry Condition	AndCondition
	-0	RangeCondition
	Case Sensitive	true
	Comments	Comment
	Input	x-age
	Max	
	Min	55
	Type	NUMERIC

Regular Expression

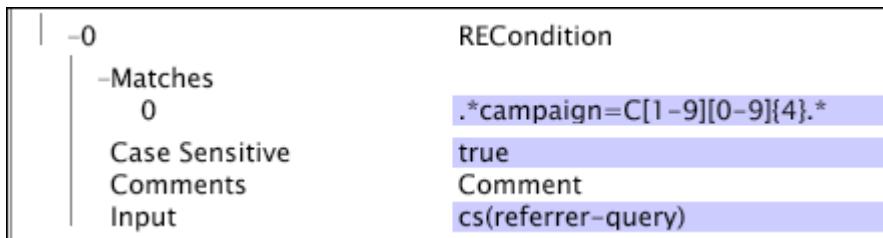
The **Regular Expression** condition test uses regular expressions pattern matching (see [Regular Expressions](#) on page 642) to determine whether the value of the specified input field contains a string that matches one of the patterns specified in the Matches parameter.

If the input is a vector of strings, only the first value in the vector is used for the test. The **Regular Expression** condition performs full string comparisons. If you want to identify substrings, you must prepend and append ".*" to the string.

The parameters of the **Regular Expression** condition are described in the following table:

Parameter	Description	Default
Case Sensitive	True or false. If set to false, upper and lower case letters are considered equal.	true
Comments	Optional. Notes about the condition.	Comments
Input	The name of the field from the log entry to use as input.	
Matches	<p>The regular expression pattern(s) to match against the value of the input field.</p> <p>To add a regular expression pattern</p> <ol style="list-style-type: none">1. Right-click Matches.2. Click Add new > Regular Expression.3. Enter the desired regular expression in the text box.	

This example illustrates the use of the **Regular Expression** condition to match a field of data collected from website traffic. The condition returns true only if the cs(referrer-query) field contains a string matching the regular expression campaign=C[1-9][0-9]{4}. This regular expression matches any string containing "campaign=C12345." However, the pattern would not match the string "campaign=C0123&" because the first character after the "C" is not in the range 1-9.



String Match

The **String Match** condition tests for string equality. It takes a specified field as input and tests the value of that field in each log entry against the strings specified in the operation's Matches parameter. If any one of these case-sensitive match strings is the same as the value in the provided input field, the operation returns true. In the event that the **StringCondition** contains no match strings, the condition returns false. If the input is a vector of strings, only the first value (string) in the vector is used for the test.

Parameter	Description	Default
Case Sensitive	True or false. If set to false, upper and lower case letters are considered equal.	true
Comments	Optional. Notes about the condition.	Comments
Input	The name of the field from the log entry to use as input.	

Parameter	Description	Default
Matches	<p>The string(s) to match against the value of the input field.</p> <p>To add a string</p> <ol style="list-style-type: none"> 1. Right-click Matches. 2. Click Add new > String. 3. Enter the desired string in the text box. 	

This example uses data collected from website traffic to illustrate the use of the **String Match** condition. The condition tests whether the input field (cs-uri-stem) matches either of the two strings specified in the Matches parameter, and it succeeds if the field cs-uri-stem is either the exact string /navigation/footer.asp or the exact string /navigation/header.asp.

StringCondition	
-0	
-Matches	
0	/navigation/footer.asp
1	/navigation/header.asp
Case Sensitive	true
Comments	Comment
Input	cs-uri-stem

Test Types for Test Operations

Navigation title: Test Types for Test Operations

The **Compare** condition and **Range** condition require that you specify the type of comparison to be made for the condition.

The following table describes the available types (**LEXICAL**, **NUMERIC**, and **DATETIME**).

Test Types for Test Operations

Test Type	Description	Notes
INTEGER	First turns the input field into an integer. If this is not possible, a value of zero is used. The test returns true only if the resulting integer input value is greater than or equal to the specified minimum value and less than or equal to the specified maximum value.	If either of the min or max fields is left blank, the system uses the appropriate min or max value available to 64-bit signed integers. If the min or max value specified in the condition does not successfully parse to an integer value, the system substitutes zero and does not stop processing the dataset.
DATETIME	First turns the input field into a date. If the input field cannot be turned into a valid date, the condition test returns false. If the field can be turned into a date, the test returns true only if the input date falls on or after the specified minimum date and on or before the specified maximum date.	If the min and max dates are not valid, the dataset are not constructed. If the min or max dates are not supplied, the system substitutes appropriately either the min date (Jan 1, 1600) or the max date (sometime in the 24th century). Adobe recommends using one of the following formats for DATETIME : <ul style="list-style-type: none"> • January 1 2013 HH:MM:SS EDT • Jan 1 2013 HH:MM:SS GMT The time zone defaults to GMT if not specified.
LEXICAL	Returns true only if the input field is lexically greater than or equal to the string specified as the minimum and less than or equal to the string specified in the maximum value.	Lexical comparison uses the ASCII value of characters in the strings moving from left to right comparing the characters. For the first character that does not match, the one with the larger ASCII value is considered to be the greater of the two. In the event that one string is shorter than the other, but up until that point all of the characters have

Test Type	Description	Notes
		been the same, the longer string is considered the greater of the two. If the strings are character for character equivalent and the exact same length, they are considered lexically equivalent.

Boolean Operations

Navigation title: Boolean Operations

The **Boolean** operations combine the results of the test operations, which function as children of the **boolean** operations.

For information about the test operations, see [Test Operations](#) on page 635. When you define a **boolean** operation, you can define zero or more children for the operation.

To add a child condition to a Boolean operation

1. Right-click the name or the number corresponding to the **Boolean** operation.
2. Click **Add new child** and choose one of the available condition types to add.
3. Repeat steps 1 and 2 until you have added all of the desired child conditions for the **Boolean** operation.



Note: When you right-click the name or the number corresponding to a **Boolean** operation, you see the **Add new sibling** menu option. A sibling is another condition at the same relative position in the condition hierarchy as the **Boolean** operation that you right-clicked. Adding a new sibling for a **Boolean** operation is the same as adding a new condition by right-clicking the **Condition** or **Log Entry Condition** parameter.

To remove a child condition from a Boolean operation:

1. Right-click the name of the child condition or the number corresponding to the child condition that you want to remove from the **Boolean** operation.
2. Click **Remove <#number>**, where number is the number corresponding to the child condition that you want to remove.

This section discusses the following conditions:

- [And](#) on page 640
- [Neither](#) on page 641
- [Or](#) on page 642

And

The **And** condition can have zero or more child conditions and returns true when none of its children nodes return false.

The **And** condition forms the root operation of all condition testing within the data workbench server. If the **And** condition contains no children, the condition evaluates to true and the associated operation proceeds. This is why actions that have only the **And** condition as the condition test always execute and why it is used as the root for all condition tests.

This example shows how an **And** condition is used to make sure that the **Copy** transformation occurs when only both the date of the log entry occurred in the year 2006 and that the page requested was `/products/purchase.asp`.

-Transformations	
-0	Copy
Name	Sample Copy
-Condition	AndCondition
-0	RangeCondition
Case Sensitive	true
Comments	Comment
Input	x-timestamp
Max	December 31, 2006
Min	January 1, 2006
Type	DATETIME
-1	StringCondition
-Matches	
0	/products/purchase.asp
Case Sensitive	true
Comments	Comment
Input	cs-uri-stem

Neither

The **Neither** condition can have zero or more child conditions and returns false if any of its child conditions evaluate to true. If the **Neither** condition contains no children, none of its children can return true. As a result, the **Neither** condition evaluates to true.

The following example shows a **Neither** condition with two **Range** conditions as its children. As defined, the **Neither** condition excludes log entries that occurred between January 1, 2007 and January 10, 2007 or during the period January 12, 2007, to January 14, 2007. Such a condition might be used as the **Log Entry Condition** to eliminate transactions from a dataset during periods in which there was a known problem with the data collected.

-Condition	
-0	AndCondition
-0	NeitherCondition
Case Sensitive	RangeCondition
Comments	true
Input	Comment
Max	x-timestamp
Min	January 10, 2007
Type	January 1, 2007
-1	RangeCondition
Case Sensitive	true
Comments	Comment
Input	x-timestamp
Max	January 14, 2007
Min	January 12, 2007
Type	DATETIME

Or

The **Or** condition can have zero or more child conditions and returns true if at least one of its child conditions evaluate to true. If the **Or** condition contains no children, none of its children can return true. As a result, the **Or** condition evaluates to false.

This example shows the **Or** condition with a **String Match** condition and a **Range** condition as its children. The **Or** condition is satisfied only if the log entry has the **x-hasproblem** value set to yes or the log entry occurred during the time range January 1, 2007, to January 10, 2007.

Condition	AndCondition
-0	OrCondition
-0	StringCondition
-Matches	yes
0	true
Case Sensitive	Comment
Comments	x-hasproblem
Input	Comment
Comments	RangeCondition
-1	true
Case Sensitive	Comment
Comments	x-timestamp
Input	January 10, 2007
Max	January 1, 2007
Min	DATETIME
Type	

Regular Expressions

Navigation title:Regular Expressions

Regular expressions are used across all data workbench search fields including the query entity panels.

- [About Regular Expressions](#) on page 642
- [Regular Expression Terminology](#) on page 643
- [About Literal Matching](#) on page 643
- [Using Metacharacters](#) on page 643
- [Pattern Extraction](#) on page 645

About Regular Expressions

A regular expression is a text pattern, consisting of a combination of alphanumeric characters and special characters known as metacharacters, that locates patterns and extract substrings from text. Regular expressions are widely used in computer programming and are an integral part of languages such as Perl.

To identify and extract complex string patterns, the data workbench server uses regular expressions in some of the transformations and conditions. What follows is a brief guide to regular expressions.

This appendix is not a comprehensive introduction to regular expressions. A particularly good reference is the O'Reilly publication *Mastering Regular Expressions, 2nd Edition* by Jeffrey E. F. Friedl.

Regular Expression Terminology

Term	Definition
Literal	A literal is a character we use in a regular expression to locate a specific sequence of characters. For example, to find product in <code>shop/products.html</code> , the string product is a literal, or what we are literally looking for in the string.
Metacharacter	A metacharacter is a special character that has a unique interpretation in the context of regular expressions. For example, the period (.) is a metacharacter that is used to match any character.
Escape Sequence	An escape sequence is simply a way to tell the regular expression engine that we would like to use one of the metacharacters as a literal. Escape sequences always start with the backslash character (\). By placing the backslash (which is also a metacharacter) in front of a metacharacter, the regular expression engine interprets the escaped metacharacter as a literal. For example, if you want to match the metacharacter period (.), you need to use an escape sequence. However, to match one of the periods in the string 168.196.0.11, you could use the regular expression consisting of a backslash and a period (\.).
Pattern	This is shorthand terminology for the regular expression. In essence, a regular expression is a pattern you are trying to match against the target string.
Target String	This term refers to the string we are searching within to locate the desired pattern.

About Literal Matching

Literal matching takes a literal string without any escape characters and looks in the target string to see if it is a substring of the target string.

In this example, you see how literal matching works. Consider a situation in which data is collected from website traffic, and the cs(referrer) field contains the following value:

`http://www.abc.com/adventurenews/today.html?ad=123AZ45`

To determine whether the referrer represents someone who clicked on one of the advertisements, you need to see whether the referrer contains the string ad. You could simply use the literal string ad to search the target string and determine if an advertisement was used to route the traffic to the site. While this would match the target string, it would match in two locations and is thus ambiguous and can lead to false positives.

The following URL contains the string ad in two different places:

`http://www.abc.com/ad vertnews/today.html?ad =123AZ45`

Thus, if you are trying to determine which sessions started as a result of a particular advertisement campaign, simply using the literal ad as the regular expression is clearly not sufficient. Changing the literal to "ad=" would eliminate this ambiguity and result in the expression making only a single match. However, even this may not be sufficient to ensure that the referrer was part of the advertisement campaign. Consider the following referrer:

`http://www.xyz.com/hello.html?pad=something`

You do not have any control over the URLs that others may be using to create links to the site. Literal matching is too simple a mechanism to locate sessions that started as a result of the advertisement campaign. The following section discusses how you can use metacharacters for more flexible and powerful matching.

Using Metacharacters

A metacharacter is a special character in a program or data field that provides information about other characters.

metacharacter	description
. (dot)	Matches a single character, for example: <code>re:x . z</code> matches "xyz" or "xxz".

metacharacter	description
*	(star) Matches one or more characters, for example: <code>x* : Z*</code> matches "ZZZ".
?	(wildcard) Matches 0 or 1 of previous expression to force minimal matching, for example: <code>xy?z</code> matches "xy" and "xyz".

Additional common regular expressions can also be used to create more complex search strings.

Lists, Ranges, and OR

Literal matching lets you look for a single string, but brackets, dashes, and pipes let you define a list of things to look for in the target string.

For this metacharacter...	The regular expression processor will...
Square Brackets ([])	Match any of the characters inside of the bracket with a single character position. For example, [AB] is an instruction to match either the letter A or the letter B and [0123456789] says match to any character in the range 0 to 9.
Dash (-)	Match a range of characters. Thus, instead of writing [0123456789] we could simply write [0-9]. This can be extended to ranges of characters and multiple ranges within one set of brackets. For example, [0-9A-C] matches the characters 0 through 9 and A to C.  Note: To test for a dash (-) as a literal inside the brackets, it must come first or last. For example, [-0-9] tests for - and 0 to 9.
Pipe ()	Match one of two choices to a given target string. For example, <code>b nat</code> matches either bat or nat.

Consider the following examples:

Pattern	String	Match
Win9[58]	OS=Win95	Win95
Win95 8	OS=Win98	Win98
[0-9]	Mozilla/3.0	3
Lesson[A-Z]	Lesson a	No match because lower-cased a is not in the range of upper-cased A through Z.

Negation

Negation is a way to say that you would like to match anything except the given characters. The negation metacharacter, the circumflex or caret (^), is used as the first character inside brackets to say that you would like the match to be anything but the remaining characters in the brackets. For example, to match any character but a semicolon (;), you would write

`[^;]`

This would match any character except the semicolon.

Positioning

To force a match to the beginning or end of a target string, one of two metacharacters are used.

For this metacharacter...	The regular expression processor will...
Circumflex or Caret (^)	Match against the beginning of the string. For example, <code>^The</code> would match the target string "The Beginning" but would not match "This is the beginning."

For this metacharacter...	The regular expression processor will...
Dollar sign (\$)	Match against the end of the string. For example, [Ee]nd\$ would match "This is the end" but would not match "The end is a special time."



Note: When the regular expression contains ^ at the beginning and \$ at the end, the entire target string must match the regular expression.

Matching Anything

The period (.) is a special metacharacter that matches any character in the target string. For example, the regular expression ^...\$ matches any target string that is exactly three characters long. The regular expression "...." matches any target string that contains at least three characters.

Repeated Patterns

Iteration metacharacters let you match a pattern more than once.

For this metacharacter...	The regular expression processor will...
Question Mark (?)	Match no instances or one instance of the character immediately preceding the metacharacter (?). For example, the pattern rea?d matches red and read.
Asterisk (*)	Match zero or more occurrences of the character immediately preceding the metacharacter (*). For example, the pattern [0-9]* matches any number of the characters 0 through 9 (any integer).
Plus (+)	Match one or more occurrences of the preceding character or range. For example, the pattern thre+ would match three but not through.
{n}	Match the proceeding character or range exactly n times. The following pattern matches United States phone numbers: [0-9]{3}-[0-9]{3}-[0-9]{4}. While not an optimal pattern, it determines whether the target string is in the proper format.
{n,m}	Match the preceding character at least n times and at most m times. For example, fo{1,2}d would match fod and food but not food.

Pattern Extraction

Pattern matching is only part of the power of regular expressions. Regular expressions also provide a mechanism for extracting key portions of a target string. This is done through the use of the left and right parentheses. These extractions are typically used as input into another process and are accessed through the use of %position%, where position is an integer referring to the count of which set of parentheses was matched.

Consider the following examples of pattern extraction:

Pattern	String	Match	Extraction
Win(9[58])	OS=Win95	Win95	%1% = 95
(Win)(95 8)	OS=Win98	Win98	%1% = Win %2% = 98
Mozilla/([0-9]).([0-9])	Mozilla/3.0	Mozilla/3.03	%1% = 3 %2% = 0

Pattern	String	Match	Extraction
Lesson([A-Z])	Lesson a	No match because lower-cased a is not in the range of upper-cased A through Z	

Event Data Record Fields

Navigation title: Event Data Record Fields

Information about the fields of data that the data workbench server can process to construct a dataset.

- [About Event Data](#) on page 646
- [Baseline Event Data Record Fields](#) on page 646
- [Derived Fields](#) on page 648

About Event Data

The event data used to build a dataset resides in files referred to as log sources. The data available in the log sources is called event data because each data record represents a transaction record or a single instance of an event with an associated timestamp.

A log source's event data is collected in real-time by **Sensors**. Event data collected by **Sensors** from HTTP and application servers is transmitted to data workbench servers, which convert the data into compressed log (.vs1) files. Event data that resides in a flat file, XML file, or an ODBC data source is read by the data workbench server, which provides decoders that you define to extract a common set of data fields from these different formats.

The following sections provide information about the data fields (referred to as event data record fields or log entry fields) that are collected by **Sensors** or read and made available to the data workbench server.



Note: The names of the fields generally follow the naming convention for the W3C extended log file format. Many of the fields have prefixes that indicate the source of the information contained in the field:

- cs indicates communication from the client to the server.
- sc indicates communication from the server to the client.
- s indicates information from the server.
- c indicates information from the client.
- x indicates information that is created by an Adobe software product.

Baseline Event Data Record Fields

Log (.vs1) files contain the fields of event data that are collected from servers by **Sensors** and used by the data workbench server in the dataset construction process. The following table lists the fields in a typical event data record as recorded by **Sensor**:

Field	Description
c-ip	The IP address of the client as included in the request made to the server. Example: 207.68.146.68
cs(cookie)	The cookies sent by the client with the request. Example: v1st=42FDF66DE610CF36; ASPSESSIONIDQCATDAQC=GPIBKEIBFBFIPOJMKCAAEPM;
cs(referrer)	The HTTP referrer string sent by the client to the server with the request.

Field	Description
	Example: <code>http://www.mysite.net/cgi-bin/websearch?qry</code>
cs(user-agent)	The string sent by the client with its request to the server that indicates what type of user agent the client is. Example: Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.7) Gecko/20040707 Firefox/0.9.2
cs-method	The method type of the HTTP request. Example: GET Reference: http://www.w3.org/TR/2000/NOTE-shoplogfileformat-20001115/#field_method
cs-uri-query	The query string portion of URI (stem + query string = URI). This is preceded by a question mark (?) and may contain one or more name-value pairs separated by ampersands (&). Example: page=homepage
cs-uri-stem	The stem portion of URI (stem + query string = URI). The stem is the actual or logical path to the requested resource on the server. Example: /index.asp
sc(content-type)	The content type of the resource being requested by the client as reported by the server. Examples: text/html, image/png, image/gif, video/mpeg
sc-bytes	The number of bytes of data sent from the server to the client in response to the request Example: 4996
sc-status	The status code returned to the client by the server. Example: 200 Reference: http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html
s-dns	The fully qualified domain name or IP address of the host of the requested resource. Example: www.adobe.com
x-experiment	The list of all the controlled experiment names and groups that the client is a member of at the time of the request. Example: VSHome_Exp.Group_1,VSRegistration_Exp.Group_2
x-timestamp	The date and time (GMT) at which the request was received by the server. The time is expressed as the number of 100 nanoseconds since January 1, 1600. Example: 12771098932000000 would be the x-timestamp value for 11:28:52.0000000 on Tuesday, September 13, 2005.
x-trackingid	The 64-bit, hexadecimal value of the unique browser identifier found in a persistent cookie as set by a Sensor and provided by the client with a request to a server. Example: 42FDF66DE610CF36

Derived Fields

The table below lists examples of fields that are derived by the data workbench server from the baseline event data record fields:

Field	Description
cs(cookie)(name)	The value of a given name-value pair within a cookie.
cs(referrer-domain)	The domain name or IP address of the HTTP referring URI.  Note: This field is read-only.
cs(referrer-host)	The entire hostname of the referrer. Example: If cs(referrer) is <code>http://my.domain.com/my/page</code> , cs(referrer-host) is <code>my.domain.com</code> .
cs(referrer-query)(name)	The value of a referrer query string.  Note: You cannot access a referrer query string value using the cs(referrer)(name) field.
cs-uri	The complete URI (stem + query string = entire URI). Example: <code>/shopping/checkout.html?product1=8Track&product2=casette&product3=cd</code>
cs-uri-query(name)	The value associated with the given name. If multiple values exist for the given name, this field returns the last of those values. Examples: <ul style="list-style-type: none">• For the URI <code>/shopping/checkout.html?product1=8Track&product2=casette&product3=cd</code>, cs-uri-query(product3) would return cd.• For the URI <code>/shopping/checkout.html?product1=8Track&product1=casette</code>, cs-uri-query(product1) would return cassette.
ctime	x-timestamp expressed as seconds since January 1, 1970. This field is also called x-unixtime.
date	x-timestamp in the format YYYY-MM-DD.
time	x-timestamp in the format HH:MM:SS.
x-local-timestring	x-timestamp converted to the local timezone that is specified in the <code>Transformation.cfg</code> file for the dataset. The format is YYYY-MM-DD HH:MM:SS.mmm.  Note: You also can define time conversions such as x-local-timestring in the <code>Log Processing.cfg</code> file. For information, see Log Processing Configuration File on page 528.
x-log-source-id	The identifier corresponding to the log source for a particular log entry. For the identifier to be recorded, you must specify it in the Log Source ID field of the <code>Log Processing.cfg</code> file when defining Sensor , log file, or ODBC data sources. For more information, see Log Processing Configuration File on page 528.

Field	Description
	Example: from VSensor01.
x-mask	The mask pattern of the Sensor data sources (derived from the .vsl file names). For a file whose name is of the format YYYYMMDD-SENSORID.vsl, x-mask is SENSORID.
x-timestamp	x-timestamp in the format YYYY-MM-DD HH:MM:SS.mmm.
x-unixtime	The decimal UNIX time derived from x-timestamp.

Sensor, when used on a server, can collect fields of event data from any valid HTTP request or response header or variable available to it through the server's API. To collect such fields of data, you must specify the desired header fields or variables in the txlogd.conf configuration file for **Sensor**. For more information, see the *Data Workbench Sensor Guide*.

Configuration Settings for Web Data

Navigation title: Configuration Settings for Web Data

Many web-specific configuration settings for **Site** are defined in **Log Processing** and **Transformation Dataset Include** files.

Web-Specific Settings for Log Processing

Navigation title: Web-Specific Settings for Log Processing

Information about web-specific settings that are defined in **Log Processing Dataset Include** files that are delivered with Adobe profiles for **Site**.

The filtering defined by these settings occurs after the log entries leave the decoders and the transformations are applied but before evaluation by the **Log Entry Condition**.

- [HTTP Status Filtering](#) on page 649
- [Robot Filtering](#) on page 650

HTTP Status Filtering

You can configure your implementation of **Site** to remove log entries with sc-status codes of 400 or above from the dataset. Successful requests have status codes that are less than 400. Your default implementation includes a **Log Processing Dataset Include** file in which HTTP status filtering is configured.

To edit the configuration settings for HTTP status filtering

1. Open the **Profile Manager** within your dataset profile and open the Dataset\Log Processing\Traffic\HTTP Status Filter.cfg file.



Note: If you have customized your implementation of **Site**, the file in which these configuration settings exist may differ from the location described.

2. Review or edit the values of the parameters of the file as desired. Use the following example as a guide.

-HTTP Status Filter.cfg	
-Log Processing	Include LogProcessingInclude
-Log Entry Condition	AndCondition
-0	RangeCondition
Case Sensitive	true
Comments	Comment
Input	sc-status
Max	399
Min	0
Type	NUMERIC
Decoder Groups	
Fields	
Parameters	
Stage	Default
Transformations	

For information about the **Range** condition, see [Conditions](#) on page 634.

3. Save the `HTTP Status Filter.cfg` file by right-clicking **(modified)** at the top of the window and clicking **Save**.
4. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

Robot Filtering

You can configure your implementation of **Site** to use lookup files to remove log entries generated by known robots, test scripts, and IP addresses for internal users from your dataset. Your default implementation includes a **Log Processing Dataset Include** file in which robot filtering is configured.

To edit the configuration settings for robot filtering

1. Open the **Profile Manager** within your dataset profile and open the `Dataset\Log Processing\Traffic\Robot Filter.cfg` file.



Note: If you have customized your implementation of **Site**, the file in which these configuration settings exist may differ from the location described.

2. Review or edit the parameters of the file using the following example and information as guides:

-Robot Filter.cfg	
-Log Processing Include	LogProcessingInclude
-Log Entry Condition	AndCondition
-0	NotRobotCondition
Case Insensitive	true
Robot Lookup File, Baseline	\$\\$(Traffic Lookups)\Robots.txt
Robot Lookup File, Extended	
Decoder Groups	
Fields	
Parameters	
Stage	Default
Transformations	

The file includes a **NotRobotCondition** that is defined by the following three parameters:

- **Case Insensitive Robot Filtering:** True or false. If true, letter case (upper/lower) is not considered in robot filtering.
- **Robot Lookup File, Baseline:** The path and filename of the text file that contains a list of browser user agents that are known robots and are to be filtered out of the dataset. Adobe provides the baseline robot lookup file. If you do not specify a path, the data workbench server looks for this file in the Lookups directory within the data workbench server installation directory.
- **Robot Lookup File, Extended:** The path and filename of an optional text file that contains a list of browser user agents or IP addresses that define robots specific to your implementation. This list can include internal monitoring robots, test scripts, and IP addresses for internal users that should be filtered out of the dataset. If you do not specify a path, the data workbench server looks for this file in the Lookups directory within the data workbench server installation directory.

If a log entry's browser user agent is not listed in either lookup file, the log entry is considered to be generated by a real visitor and is not filtered from the dataset.



Note: Matching in the robot lookup files uses substrings to compare against the c-ip and the cs(user-agent) log fields. If the search string starts with "\$" it must match the front of the string being tested, and if it ends with "\$" the search string must match the end of the string being tested. If the search string both begins with and ends with "\$" the strings must match exactly for the log entry to be filtered out. For example, to test for all IP addresses in a class C block, you would use a string such as \$231.78.123. to force a match at the front of the string. This would match addresses 231.78.123.0 through 231.78.123.255.

3. Save the file by right-clicking (**modified**) at the top of the window and clicking **Save**.
4. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs.

Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.



Note: If it is critical that the underlying log entries used to construct a dataset do not change (even if the transformations used to construct and update the dataset and its dimensions change), the Robot Lookup File, Baseline, and the Robot Lookup File, Extended, should be version controlled. Placing a version number on these files ensures that updates to the default robot lookup files do not unintentionally change previously constructed reporting datasets by adding or deleting entries in these files.

Web-Specific Settings for Transformation

Navigation title:Web-Specific Settings for Transformation

Information about web-specific settings that are defined in **Transformation Dataset Include** files that are delivered with Adobe profiles for **Site**.

The conditions, dimensions, and parameters that are defined by these settings are created during the transformation phase of dataset construction.

- [Page View Condition](#) on page 652
- [URI Dimension](#) on page 653
- [Referrer Dimension](#) on page 655
- [Session Parameters](#) on page 656

Page View Condition

The **Page View Condition** is a condition operation that determines whether a particular log entry (that is, a page request) should be included in the data gathered about a visitor's page view history. When the log entry satisfies the **Page View Condition**, it becomes an element of the Page View countable dimension. If a log entry does not satisfy the **Page View Condition**, its data fields still are accessible by other dimensions. In addition to the Page View dimension, the following dimensions can be affected by the results of the **Page View Condition**:

- **URI and Page:** These dimensions are directly affected by the **Page View Condition**. If the given page does not pass the **Page View Condition**, it is not be included in the URI or Page dimensions.
- **Visitor Page Views and Session Page Views:** The Visitor Page Views and Session Page Views dimensions are a count of the number of pages viewed by a visitor to or in a given session, respectively. Pages filtered out by the **Page View Condition** are not part of this count.
- **Session Number:** The **Page View Condition** has an indirect effect on the Session Number dimension. The Session Number dimension is created prior to the **Page View Condition**; therefore, when considering **Session Number** in relation to the **Page Views**, it is possible to have sessions with no page views.

Your default implementation of **Site** includes a **Transformation Dataset Include** file in which the Page View countable dimension and the related **Page View Condition** are defined.

For information about countable dimensions, see [Extended Dimensions](#) on page 606.

To edit the configuration settings for the Page View Condition

1. Open the **Profile Manager** within your dataset profile and open the `Dataset\Transformation\Traffic\Page_View.cfg` file.



Note: If you have customized your implementation of **Site**, the file in which these configuration settings exist may differ from the location described.

2. Review or edit the values of the parameters of the **Page View Condition** as needed. Use the following example as a guide. In this file, the **Page View Condition** is defined by a **Copy** transformation. Note that this file also contains the definition of the Page View countable dimension.

-Page View.cfg	
-Transformation Include	TransformationInclude
-Extended Dimensions	
-0	Countable
Name	Page View
-Condition	AndCondition
-0	NotEmptyCondition
Comments	Comment
Input	x-is-page-view
Comments	Comment
Hidden	false
Key	
Parent	Session
-Transformations	
-0	Copy
Name	Page View Condition
Comments	Comment
Condition	AndCondition
Default	
Input	"1"
Output	x-is-page-view



Note: For information about countable dimensions, see [Extended Dimensions](#) on page 606. For information about the **Copy** transformation, see [Data Transformations](#) on page 573.

3. Save the file by right-clicking (**modified**) at the top of the window, then click **Save**.
4. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

URI Dimension

If you are working with **Site**, you need to define the URI dimension whose elements are the URI stems of the website pages viewed. Your default implementation includes a **Transformation Dataset Include** file in which the URI simple dimension is defined.

For information about simple dimensions, see [Extended Dimensions](#) on page 606.

To edit the configuration settings for the URI dimension

1. Open the **Profile Manager** within your dataset profile and open the **Dataset\Transformation\Traffic\URI.cfg** file.



Note: If you have customized your implementation of **Site**, the file in which these configuration settings exist may differ from the location described.

2. Review or edit the values of the parameters of the file as desired. Use the following example and information as guides.

URI.cfg	
- Transformation Include	TransformationInclude
- Extended Dimensions	
- 0	Simple
Name	URI
Case Sensitive	true
Comments	Comment
Condition	AndCondition
Hidden	false
Input	x-page-dim
Load File	
Maximum Elements	32768
Operation	
Parent	Page View
- Transformations	
- 0	Copy
Name	URI is cs-uri-stem by default
Comments	Comment
Condition	AndCondition
Default	
Input	cs-uri-stem
Output	x-page-dim
- 1	Format
Name	AppendURI transformations' output appended to URI
- Condition	AndCondition
- 0	NotEmptyCondition
Comments	Comment
Input	x-page-dim-qs
- Inputs	
0	x-page-dim
1	x-page-dim-qs
Comments	Comment
Format	%1%?%2%
Output	x-page-dim

The configuration settings for the URI dimension include the following two parameters:

- **Case Sensitive:** True or false. If true, letter case (upper/lower) is considered in identifying unique pages. The default value is true.
- **Maximum Elements:** The maximum number of elements (that is, URIs) for the URI dimension. The default value is 32768.



Note: Changing this value can cause serious performance issues. Do not change this value without consulting Adobe.

- Save the **URI.cfg** file by right-clicking (**modified**) at the top of the window, then click **Save**.

-
- To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

Referrer Dimension

If you are working with **Site**, you need to define the Referrer dimension whose elements consist of the second level domains of the referrers of the first log entries in all sessions. Your default implementation includes a **Transformation Dataset Include** file in which the Referrer simple dimension is defined.

For information about simple dimensions, see [Extended Dimensions](#) on page 606.

To edit the configuration settings for the Referrer dimension

- Open the **Profile Manager** within your dataset profile and open the `Dataset\Transformation\Traffic\Referrer.cfg` file.



Note: If you have customized your implementation of **Site**, the file in which these configuration settings exist may differ from the location described.

- Review or edit the values of the parameters of the file as desired. Use the following example and information as guides.

-Referrer.cfg	
-Transformation Include	TransformationInclude
-Extended Dimensions	
-0	Simple
Name	Referrer
-Condition	AndCondition
-0	NeitherCondition
-0	StringCondition
-Matches	
0	\$(Internal Domains)
Case Sensitive	false
Comments	Comment
Input	cs(referrer-domain)
Comments	Comment
Case Sensitive	false
Comments	Comment
Hidden	false
Input	cs(referrer-domain)
Load File	
Maximum Elements	0
Operation	FIRST ROW
Parent	Session
Transformations	

The configuration settings for the Referrer dimension include the Maximum Elements parameter, which specifies the maximum number of elements (that is, referrers) for the Referrer dimension. The default value is 32768.



Note: In the example above, the **Maximum Elements** parameter is set to 0. When this parameter is set to 0, the data workbench server uses its internal default value of 32768.

3. Save the `Referrer.cfg` file by right-clicking (**modified**) at the top of the window, then click **Save**.
4. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

Session Parameters

If you are working with **Site**, you can specify parameters that define the boundaries of a visitor's session on a website. These parameters are valid only when defined in a **Transformation Dataset Include** file within your **Site** implementation.

The following parameters are unique in that they can be members of **Transformation Dataset Include** file's **Parameters** vector, or they can be listed as individual parameters in the `Transformation.cfg` file. A parameter can be defined exactly once, so these parameters are defined either in the `Transformation.cfg` file or in the **Parameters** vector of the dataset include file - not in both files.

Maximum Session Duration and Session Timeout

Maximum Session Duration and Session Timeout are string parameters that define the length of a visitor's session. These parameters work with the Internal Domains parameter to determine session length.

Maximum Session Duration specifies the longest length of session before a new session is started. This keeps web pages that have auto content refreshing from creating sessions that are arbitrarily long. If the referrer of a click is set to one of the entries in the Internal Domains parameter, this timeout is used to define the end of a session. No session may be longer than the specified Maximum Session Duration regardless of how many clicks it contains. The recommended value is 48 hours.

Session Timeout specifies the amount of time that needs to pass between log entries of a given visitor to determine the end of one session and the start of a new session (that is, the typical timeout used to define a user session). The recommended value of this parameter is 30 minutes. If the referrer of a click is not set to one of the referrers in the Internal Domains parameter, this timeout is used to define the session. If `cs(referrer-domain)` for a log entry is in the list of internal domains, then Maximum Session Duration determines whether the current log entry is part of an existing session or the start of a new session.

Consider a situation in which a visitor is called away from his computer for a period of time longer than the Session Timeout while in the middle of browsing the site. Upon returning, he continues browsing where he left off. Because the visitor never leaves the site or closes his browser, the `cs(referrer-domain)` of his next click is the same as the internal domain, and his original session remains active as long as the Maximum Session Duration setting is not reached. If the domain of the site is listed as an internal domain, and the maximum timeout is not reached, the visitor's interaction appears as a single session and not two separate sessions. However, if the visitor returns to his computer and his next click has an external (or blank) referrer, a new session begins.



Note: The **Sessionize** transformation's **Timeout Condition** also plays a role in determining the length of a visitor's session. If Session Timeout and Maximum Session Duration do not apply, the **Timeout**

Condition is checked to determine whether a log entry should be considered the start of a new session. For more information, see [Data Transformations](#) on page 573.

To edit the Maximum Session Duration and Session Timeout parameters

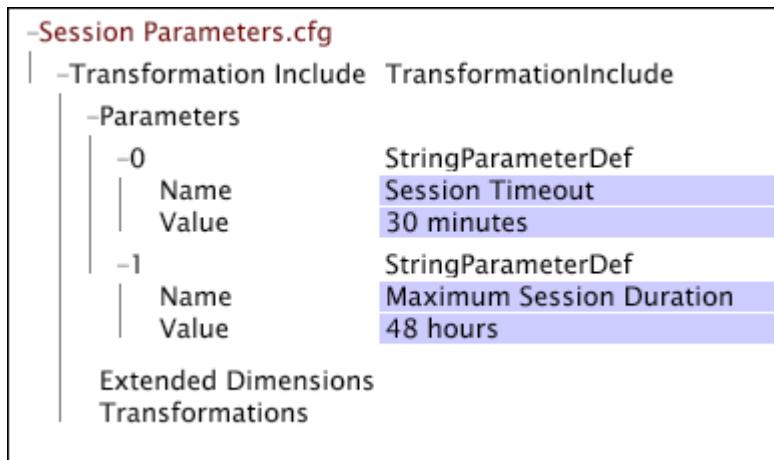
If you are working with **Site**, your default implementation likely includes a **Transformation Dataset Include** file in which the names and recommended values of these parameters are specified.

1. Open the **Profile Manager** within your dataset profile and go to `Dataset\Transformation\Traffic\Session Parameters.cfg`.



Note: If you have customized your implementation of **Site**, the file in which these parameters are defined may differ from the location described.

2. Edit the values of the parameters as desired. Be sure to specify the desired units (minutes, hours, and so on).



3. Save the `Session Parameters.cfg` file by right-clicking (**modified**) at the top of the window and clicking **Save**.
4. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

Internal Domains

Internal Domains is a vector parameter that lists domain level hosts (internal referrers) that should be treated as part of a particular website. These hosts are removed from the referrer dimension (which is a list of the external referrer information). When `cs(referrer-domain)` matches any of the strings listed in the set of internal domains, Session Timeout is ignored and Maximum Session Duration is used to determine session length.

The Internal Domains parameter also can be used to prevent the start of a new session when visitors move among a company's multiple domains associated in a way that exceeds session timeout. For example, consider a company that has parts of its site split across two domains: one is logged (`xyz.com`), and the other is not logged (`xyz-unlogged.com`). If these sites are integrated in a way that facilitates the seamless movement of traffic across the two domains, it is not desirable to generate a different session each time the visitor moves from `xyz-unlogged.com` domain back to the `xyz.com` domain. Listing `xyz-unlogged.com` as an internal domain keeps sessions from being split into multiple sessions as a result of traffic across these two domains as long as the Maximum Session Duration setting is not reached.

To add an internal domain

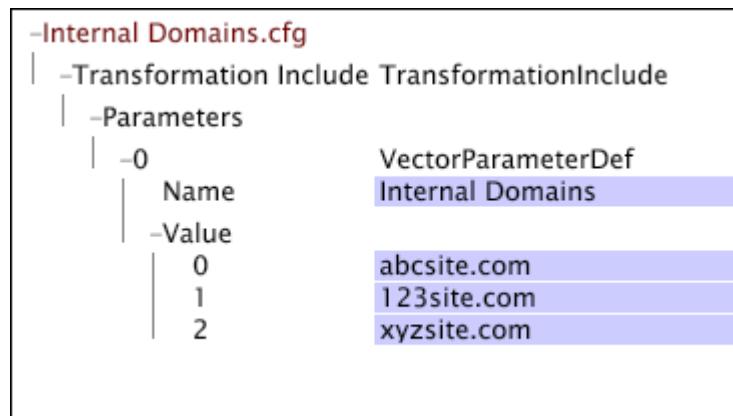
If you are working with **Site**, your default implementation includes a **Transformation Dataset Include** file for defining the Internal Domains parameter. In this file, the parameter is named; you just enter the internal domains that you want to include and save the updated file.

1. Open the **Profile Manager** within your dataset profile and go to `Dataset\Transformation\Traffic\Internal Domains.cfg`.



Note: If you have customized your implementation of **Site**, the file in which the Internal Domains parameter is defined may differ from the location described.

2. Right-click **Value** for the Internal Domains vector parameter and click **Add new > Value**.
3. Edit the values as desired.



4. Save the `Internal Domains.cfg` file by right-clicking (**modified**) at the top of the window and clicking **Save**.
5. To make the locally made changes take effect, in the **Profile Manager**, right-click the check mark for the file in the **User** column, then click **Save to > <profile name>**, where profile name is the name of the dataset profile or the inherited profile to which the dataset include file belongs.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe, as your changes are overwritten when you install updates to these profiles.

Time Zone Codes

Navigation title: Time Zone Codes

Information about the Time Zone codes and formats.

Most time-based parameters in the data workbench server are specified in the following format:

- Month DD , YYYY HH :MM :SS TZone
- Example: August 13, 2013 22:30:00 EST

Time zones are expressed in a system-independent time zone format (Coordinated Universal Time) of the following format:

- UTC +hhmm dstrules

The sign (+) can be either a plus (+) or a minus (-) sign, and hhmm is the offset from UTC in hours and minutes. The optional variable dstrules specifies a set of rules to implement Daylight Saving Time or a similar clock-shifting policy.

If you specify `dstrules`, a tab-delimited file named `dstrules.dst` must be present within the `Dataset\TimeZone` directory of either the **Base** profile (for configuration files that are not associated with a particular dataset) or the dataset profile (for configuration files that are dataset-specific). The file specifies a time-zone independent set of rules for Daylight Saving Time. You can have different sets of rules for different years. The `DST.dst` file provided by Adobe in the **Base** profile specifies the standard U.S. rules established by the Energy Policy Act of 2005 (in effect starting 2007) and the U.S. rules for prior years.

Sample time zone entries are listed below:

- U.S. Eastern Daylight Time: Time Zone = string: UTC -0500 DST
- UTC time with no offset and no `dstrules` (corresponding to GMT): Time Zone = string: UTC -0000

When this format is used, the system time zone of data workbench server, data workbench, and Report machines need not be the same as the specified time zone. In addition, all active dataset profiles on a data workbench server machine need not have the same time zone setting.

The following table contains the list of codes you can use to specify time zones in time-based parameters.

Time Zone Code Table

If you are implementing Daylight Saving Time or a similar clock-shifting policy, you must save the `.dst` file containing the appropriate rules in the profile name `|Dataset|Timezone` directory on the data workbench server machine.

Code	Time Zone	Offset from GMT
gmt	Greenwich Mean	0
est	Eastern Standard	5
edt	Eastern Daylight	5
cst	Central Standard	6
cdt	Central Daylight	6
mst	Mountain Standard	7
mdt	Mountain Daylight	7
pst	Pacific Standard	8
pdt	Pacific Daylight	8

Data Workbench Report Server

Navigation title:

Using the Report Server for Data Workbench, you can schedule and automatically generate report sets and distribute to others.

For example, you can create a status report containing the particular visualizations that members of upper management are interested in seeing and then distribute the resulting report to the appropriate executives.

You can also create reports that alert you when a metric has reached a defined threshold within a specified time frame. For example, you can create a report that alerts you when the number of orders for the previous day was less than 10 or greater than 100.

Understanding Report Sets

Navigation title:Understanding Report Sets

A report set is a collection of workspaces that Report Server generates based on the values specified in a Report.cfg configuration file.

In your Insight installation folder, each subfolder within the <working profile name>\Reports folder represents a report set that has been created. Each report set has its own Report.cfg configuration file within that subfolder.



Note: In the **Profile Manager** in Data Workbench, report sets appear as subfolders within the Reports folder. For more information about the **Profile Manager**, see the [Data Workbench User Guide](#).

By defining specific configuration settings for a report set in its Report.cfg file, you can schedule the creation and distribution of the reports, including who receives which reports and in what formats.

Generating Reports

Navigation title:Generating Reports

Generate reports by processing workspaces and specifying them as reports.

Report generates your reports at the interval set in the Every parameter in the Report.cfg file (such as "day," which processes the report on a daily basis), and based on the other Report.cfg file settings.

While generating reports, the percent complete displays on the **Reports** tab under the thumbnail for that particular report. If Report encounters a problem during report generation, the most recent error message displays on the **Reports** tab in the report set's folder. If Report does encounter an error for a particular report, it continues to process the other reports in the set.

You can generate the reports in a report set in any or all of the following formats using the Report Types parameter in the Report.cfg file:

- Microsoft Excel file (.xls or .xlsx)
- Portable network graphic file (.png)
- Thumbnail (.jpg)

Along with the types of output specified, Report creates an .xml file named the same as your report. This <report name>.xml file contains the description of the report that displays in Data Workbench on the **Reports** tab below the report's thumbnail. This makes the description available to use when distributing your reports through a reporting portal. For information about the **Report Portal**, see [Using the Report Portal](#) on page 669.



Note: If you redefine an internal metric, the system behaves unexpectedly because of the wrong value. Your reports will not generate unless a metric reads 100%. It is recommended that you do not change metric definitions.

Distributing Reports

Navigation title:Distributing Reports

After the reports have been generated, Report distributes the reports in the set based on the settings in its Report.cfg file.

Report enables you to distribute the reports in a set using the following methods:

- **Email:** To distribute reports as Excel files, .png files, or thumbnails via email (in-line or as attachments), specify the Mail Report parameters in the report set's Report.cfg file. All reports in that set are emailed in one message to the specified recipients.

-
- **Shared Directory:** To distribute reports as Excel files, .png files, or thumbnails to a shared directory, specify the directory in the Output Root parameter in the report set's Report.cfg file.
 - **Reporting Portal:** A reporting portal enables you to view reports through your web browser. Adobe's **Report Portal** displays reports generated as Excel or .png files, but not those generated as thumbnails (.jpg). To distribute reports to a portal, specify the document root of the web server used for the portal in the Output Root parameter in the report set's Report.cfg file. For more information about installing and using **Report Portal**, see [Using the Report Portal](#) on page 669.



Note: To read the output that is currently supported by Report, you must have an application capable of displaying the reports in the desired format(s). For example, to view .xlsx files, you must have Microsoft Excel 2007 or later.

You can also use a combination of these generation and distribution options. For example, if you set the Report Types parameter to generate a report set as Excel and .png files, and then set the Mail Report and Output Root parameters, all of the reports in that set are distributed to the specified output directory (perhaps to be viewed in a portal) and emailed in one email to recipients.

For steps to configure your report sets, see [Working with Report Sets](#) on page 661. For more information about the specific Report.cfg parameters, see [Report.cfg Parameters](#) on page 704.

Report Set Requirements

Navigation title:Report Set Requirements

Conceptual information about requirements for your report sets.

All report sets must:

- Be stored in a subfolder within the <working profile name>\Reports folder. They must appear on the **Reports** tab of the **Worktop**.
- Contain a Report.cfg file in the report set folder.
- Contain (or point to) at least one workspace.



Note: To be viewed using **Report Portal**, reports must follow specific naming conventions. Additionally, the directory into which reports are saved must follow a prescribed structure. For a description of these requirements, see [Ensuring that Your Report Sets are Compatible with Report Portal](#) on page 698.

Working with Report Sets

Navigation title:Working with Report Sets

Information about creating your report set, reviewing the status of your report set, and editing the existing Report.cfg files.

Creating a Report Set

Navigation title:Creating a Report Set

Steps to create a new report set.

You must complete the following steps in order:

1. [Create a New Folder for the Report Set](#) on page 662
2. [Create the Report Server Workspaces](#) on page 662
3. [Configure the Report Set](#) on page 662
4. [Publish the Report Set to the Data Workbench Server](#) on page 666

You also can copy and paste an existing report set folder and modify its reports and Report.cfg file.

Create a New Folder for the Report Set

Steps to create a new report set folder.

1. In data workbench, open the **Profile Manager** by right-clicking within a workspace and clicking **Admin > Profile > Profile Manager**.
2. Click **Reports** to open the **Reports** folder.

File	Base	Traffic	Value	Marketing	Operations
Traffic Summary.vw report.cfg		✓			

3. In the **User** column for the folder in which you want to create the new report set, right-click and select **Create > Folder**. A folder named **New Folder** appears in the **File** column for **Reports**.
4. Right-click in the **User** column for the new folder, click in the **Dir** parameter, then type the new name of the folder.

Create the Report Server Workspaces

Navigation title:Create the Report Workspaces

You can create multiple workspaces for a single report set directly in the **Worktop**.

For information about data workbench and using the interface, see the [Data Workbench Client Guide](#).

To create a report workspace using the Worktop

1. In data workbench, on the **Reports** tab, select the report set folder (tab or drop-down subdirectory) in which you want to add the new report(s).
2. Add a report workspace using one of the following methods:
 - Right-click within the desired report set subdirectory and click **New > Blank Workspace** or **New > Report Template**.
 - or-
 - Copy a workspace from another location and paste it into the appropriate report set folder on the **Reports** tab.
3. Click the thumbnail to open the workspace and add the desired visualizations, legends, and annotations to the report.
4. To name the report, click the text in the title bar of the report to enter edit mode and type the new title.
5. To save the report, right-click the title bar of the report and click **Save**.

Configure the Report Set

Navigation title:Configure the Report Set

After you have created and saved the workspaces within your report set folder, you must create a new **Report .cfg** file.

You must specify in the **Report .cfg** file for the report set when and how the reports are to be generated and distributed.

To create a new Report.cfg

1. In data workbench, open the **Profile Manager** by right-clicking within a workspace and clicking **Admin > Profile > Profile Manager**.
2. Click **Reports** to open the **Reports** folder.

-
3. Click the folder for your report set.
 4. In the **User** column for your report set folder, right-click and select **Create > Report**. A new **Report.cfg** file appears in the **File** column.
 5. In the **User** column for the new **Report.cfg** file, right-click the check mark for the **Report.cfg** file, then click **Open > from the workbench**.

Traffic	Marketing	Metrics	Operations	Admin	Reports
Sent every to default output path					
-report.cfg					
+Attachments					
+Mail Report					MailReport
+Report Types					
Alert Threshold			0		
Allow Report Regeneration					
Color Set			1		
Command To Execute					
Default Excel Template					
Dimension Name					
Email Only If Perfect			false		
End Date					
Every					
Filter Formula					
Gamma Correction			1.6		
Hide Logos			false		
Lookup File					
Output Root					
Start Date					
Thumbnail X			240		
Thumbnail Y			180		
Top N Metric					
Top N Value			0		
Use Local Sample Only			false		
Workspace Path					
XSL Output File					
XSL Template					

6. Edit the configuration parameters as desired. For information about these parameters, see [Report.cfg Parameters](#) on page 704.



Note: The sample **Report.cfg** shown in this example contains only the parameters included in the **Report.cfg** file by default. If you need to add additional parameters to a **Report.cfg** file, you must do so using a text editor. For steps to do so, see [Editing Existing Report.cfg Files](#) on page 667.

7. Save the file by right-clicking **Report.cfg (modified)** at the top of the file and clicking **Save as Reports\<ReportSetName>\Report.cfg**.

-
8. Close the file.
 9. In the **Profile Manager**, right-click the check mark in the **User** column for the new Report.cfg file and select **Save to<profile name>**.

Generating Reports as Microsoft Excel Files

Navigation title: Generating Reports as Microsoft Excel Files

Information to generate reports as Excel files.

The following requirements must be met:

- Microsoft Excel must be installed on the same machine as Report Server.
- The user account under which the Report Server process is running must have permission to access Microsoft Excel.



Note:

- When you generate reports as Excel files, you are opening a new instance of Excel. For more information about this process, see <http://support.microsoft.com/kb/257757>.
- Although data workbench supports more than 256 columns and 65,536 rows of data, Microsoft Excel does not.

If these requirements are met, Report Server automatically starts Microsoft Excel and output data from certain visualizations, dimension and value legends, and text annotations to a new Excel workbook with one visualization per worksheet.



Note: Data is not exported from graphs, path browsers, process maps, scatter plots, and globes.

Unless you have specified a **Custom Title** for the visualization, the type of window (for example, Movie Table) is used as the worksheet name.

For more information about specifying **Custom Titles** for visualizations, see the [Data Workbench Client Guide](#).

Using a Template File

You also can generate a report as an Excel file using a template Excel (.xls or .xlsx) file. Using a template file can reduce the amount of time that you spend formatting your data each time the report is generated.

This template file must be an .xls or .xlsx file, not an .xlt file.

You can choose to define a template for each individual report, a generic template for all of your reports, or a combination of both. These two items are not mutually exclusive, so you can define a generic template and then define specific templates as well.

To generate a report using a generic template that you use for all reports, you must specify the name of that Excel file in the Default Excel Template parameter in the Report.cfg for that report set file, then place the template file in the same folder as the Report.cfg for that report set (*data workbench installation directory\ProfileName\Reports\ReportSetName*). For information about this parameter, see [Report.cfg Parameters](#) on page 704.

To generate a report using a template that is specific to a report, you must name the Excel file the same as the report workspace (.vw) file, then place the template file in the same folder as the report workspace (.vw) file.

When the report is generated, the existing tabbed sheets in the template (each representing one visualization) are repopulated with the most recent data from the report, while any new windows that are not present in the template as tabbed sheets are ignored. Any other tabbed sheets in the template file remain unchanged.

In addition, if you have a macro defined in the template Excel file that you would like to run automatically when the report is generated, name the macro “VSExport.”

Dynamically Generating Reports

Navigation title:Dynamically Generating Reports

You can generate reports dynamically for the dimension elements that you specify in a lookup file or for a particular number of the dimension elements, such as for the users with the 10 highest order counts.



Note: Adobe discourages the creation of dynamic report sets for daily (or more frequent) report generation. Dynamic report generation can be resource-intensive if you configure reports with large or complex queries.

- [Lookup File Dimension Element Reports](#) on page 665
- [Top Dimension Element Reports](#) on page 666

Lookup File Dimension Element Reports

To configure a report set to dynamically generate and (optionally) distribute reports for the elements of a dimension specified in a lookup file, specify the following parameters in the Report.cfg file:

- **Dimension Name**
- **Lookup File**

For detailed descriptions of these parameters, see [Report.cfg Parameters](#) on page 704.

To create a dynamic report set using a lookup file

1. Create a two-column lookup file for a given dimension. This file must contain two tab-delimited columns, without a header row.
 - Column 1 should contain a list of dimension elements.
 - Column 2 should contain the email addresses of the report recipients. A report for a given element in column 1 is sent to the email address(es) on the same row in column 2. You can enter multiple email addresses by separating them with commas (no spaces).



Note:

- If reports are not to be emailed, column 2 can be empty but must exist.
- This file may contain blank lines.

2. Optional. To enable the emailing of reports, you must do the following in the **Mail Report** section of the Report.cfg file for that particular report set:
 - In the SMTP Server parameter, list the appropriate SMTP server to be used to distribute reports via email.
 - In the Recipients parameter, list at least one email address to enable the reports to be distributed via email. The reports are not mailed to the address listed—you set this parameter only to allow the reports to be emailed. This can be a bogus address, but it must be in an allowable format (for example, jsmith@company.com).
3. Save the lookup file to the report set's folder.
4. Open the Report.cfg file for the report set.
5. In the Dimension Name parameter, type the name of the dimension for which you want to dynamically generate a report.
6. In the Lookup File parameter, type the location and name of the lookup file containing the dimension elements that you want in the report and the recipient email addresses.
7. Repeat these steps for additional report sets.



Note: Dynamic reports are not emailed to recipients until the entire report set is complete.

Top Dimension Element Reports

To configure a report set to dynamically generate reports for the top dimension elements, counting by the specified metric, specify the following parameters in the Report.cfg file:

- Dimension Name
- Top N Metric
- Top N Value
- (Optional) Preload Query Filter

For detailed descriptions of these parameters, see *Report.cfg Parameters* on page 704.

To create a dynamic report set for the top elements

1. Open the report set's Report.cfg file.
2. In the Dimension Name parameter, type the name of the dimension for which you want to dynamically generate a report set.
3. In the Top N Metric parameter, type the name of the metric by which you want to sort the dimension.
4. In the Top N Value parameter, type the number of dimension elements that you want in the report set.
5. (Optional) In the Preload Query Filter parameter, type the name of the desired filter.
6. Repeat these steps for additional report sets.
7. For information about using a dynamic title visualization with your report, see the *Data Workbench User Guide*.

Publish the Report Set to the Data Workbench Server

Navigation title: Publish the Report Set to Data Workbench Server

The final step in creating a report set is to publish the newly created report set to the data workbench server machine.

To upload the report set to the data workbench server

1. On the **Reports** tab, right-click one of the reports that you want to upload and click **Save to server > Yes**.
2. Repeat this step for each report to be uploaded to the data workbench server machine.

Report will now generate any reports that you defined according to the schedule that you specified in the Report.cfg file.

Reviewing Report Set Status

Navigation title: Reviewing Report Set Status

Whenever a Report.cfg file has been configured to generate reports from workspaces, the **Worktop** in data workbench displays information about the particular report, as defined in the Report.cfg configuration file.

- Percent complete.

The percent of processing complete for the report set displays next to the folder name or the Report.cfg file name as "reports X% complete." For example, if the folder is open, the Report.cfg line would read + Report.cfg (reports 10% complete).

- Format (.png, .xls, .xlsx, .jpg)
- Time generated and sent
- Receivers
- Status
- Errors
- Next scheduled report generation
- Processing time

Editing Existing Report.cfg Files

Navigation title: Editing Existing Report.cfg Files

Steps to edit existing Report.cfg files using **Worktop** or a text editor.



Note:

- You must be working online to edit Report.cfg files. To work online, from the **Worktop**, right-click the title bar and click **Work Online**.
- If the **Allow Report Regeneration** parameter in the Report.cfg file is set to True, when you make changes to that file and save that file back to the server, Report automatically re-generates the reports in that set. Although it regenerates the reports, it does not re-send the reports via email. For steps to do so, see *Resending Reports by Email* on page 668.

You can edit an existing Report.cfg from the **Worktop** or using a text editor.

Editing a Report.cfg file using the **Reports** tab of the **Worktop** enables you to edit only those parameters, vectors, and vector items that already exist in the file. If you need to add a parameter or vector to the file, you must edit it using a text editor, such as Notepad.

- *To edit an existing Report.cfg using the Worktop* on page 667
- *To edit an existing Report.cfg using a text editor* on page 667

To edit an existing Report.cfg using the Worktop



Note: You must be working online to edit the Report.cfg from the **Worktop**.

1. In data workbench, on the **Reports** tab, select the subfolder (tab or drop-down subdirectory) for the report set that you want to configure.
2. Click **Report.cfg**. The parameters of the Report.cfg for this report set display.
3. Edit the configuration parameters as desired. For information about these parameters, see *Report.cfg Parameters* on page 704.
4. Save the file by right-clicking **Report.cfg (modified)** at the top of the parameters and clicking **Save to<server location>**.

To edit an existing Report.cfg using a text editor

1. Open the **Reports Manager** by right-clicking within a workspace and clicking **Admin > Profile > Reports Manager**.
2. Click the folder for your report set.
3. Right-click the check mark next to the Report.cfg for this report set and click **Make Local**.
4. In the **User** column, right-click the check mark next to Report.cfg for this report set and click **Open > in Notepad**. The Report.cfg file opens.

The sample Report.cfg shown in *Configure the Report Set* on page 662 contains only the parameters included in the Report.cfg file by default. The following example includes all of the parameters available for the Report.cfg file that you can use as models for your parameter entries:

```
Attachments = vector: 1 items
  0 = Attachment:
    FileName = string: c:\\myimage.jpg
    Content Type = string: image/jpeg
  Alert Threshold = int: 0
  Allow Report Regeneration = bool: true
  Color Set = int: 1
  Command To Execute = string:
```

```

Default Excel Template = string:
Dimension Name = string:
Email Only If Perfect = bool: false
End Date = string:
Every = string: month
Excel Watchdog Timeout (seconds) = double: 300.0
Filter Formula = string:
Gamma Correction = double: 1
Hide Logos = bool: false
Lookup File = string:
Mail Report = MailReport:
    Body XSL Template = string:
    Recipients = vector: 0 items
    SMTP Server = SmtpServerInfo:
        Address = string:
        Password = string:
        User = string:
    Sender Address = string:
    Sender Name = string:
    Subject = string:
Notification Only = bool: false
Output Root = string:
Report Types = vector: 3 items
    0 = string: thumbnail
    1 = string: png
    2 = string: excel
Start Date = string: 7/1/06 19:16 EDT
Thumbnail X = int: 240
Thumbnail Y = int: 180
Top N Metric = string:
Top N Value = int: 0
Use Local Sample Only = bool: false
Workspace Path = string:

```

5. Edit the configuration parameters as desired. For information about these parameters, see [Report.cfg Parameters](#) on page 704.
6. Save and close the file.
7. In the **Reports Manager**, right-click the check mark in the **User** column for the `Report.cfg` file and select **Save to<profile name>**.

Resending Reports by Email

Navigation title:Resending Reports by Email

Steps to resend reports by email.

If the **Allow Report Regeneration** parameter in the `Report.cfg` file is set to True, when you make changes to a `Report.cfg` file and save that file back to the server, Report Server automatically re-generates the reports in that set. It does not resend the reports via email.

1. On the **Reports** tab, select the subfolder (tab or drop-down subdirectory) for the report set that you want to resend.
2. Click **Report.cfg**. The parameters of the `Report.cfg` for this report set display.
3. Change the **Start Date** parameter to the future time at which you want to reports to be resent.
4. Save the file by right-clicking **Report.cfg (modified)** at the top of the parameters and clicking **Save to<server location>**.

The reports in this set regenerate and are resent by email to the specified recipients.

Email Notification of Completed Report

Navigation title:Email Notification of Completed Report

You can configure data workbench to send an email when a report is generated.

Setting this value to True does not send out the report, but rather sends an email notifying the subscribed user that the report has been generated.

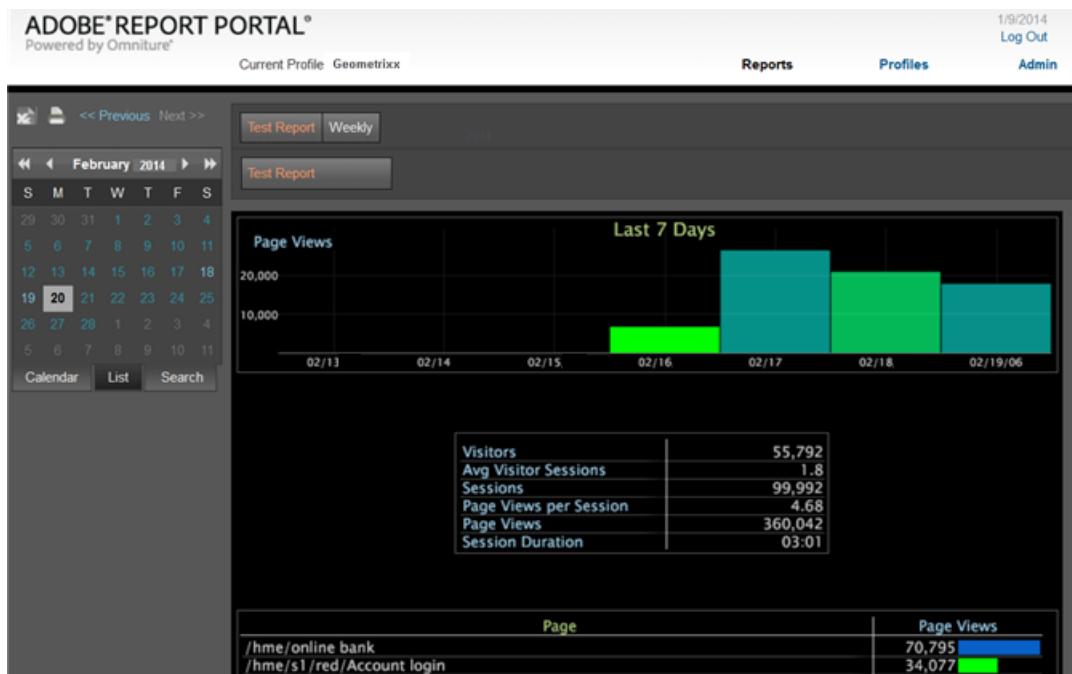
1. From the **Reports** menu, click the **Monthly** tab.
2. Open the `report.cfg` file.
3. In the **Notification Only** field, enter true to send an email when the report is complete.

Using the Report Portal

Navigation title:

The **Report Portal** enables you to view reports generated by Report Server through your web browser.

The **Report Portal** is a web-based application that runs on a Microsoft Internet Information Services (IIS) web server.



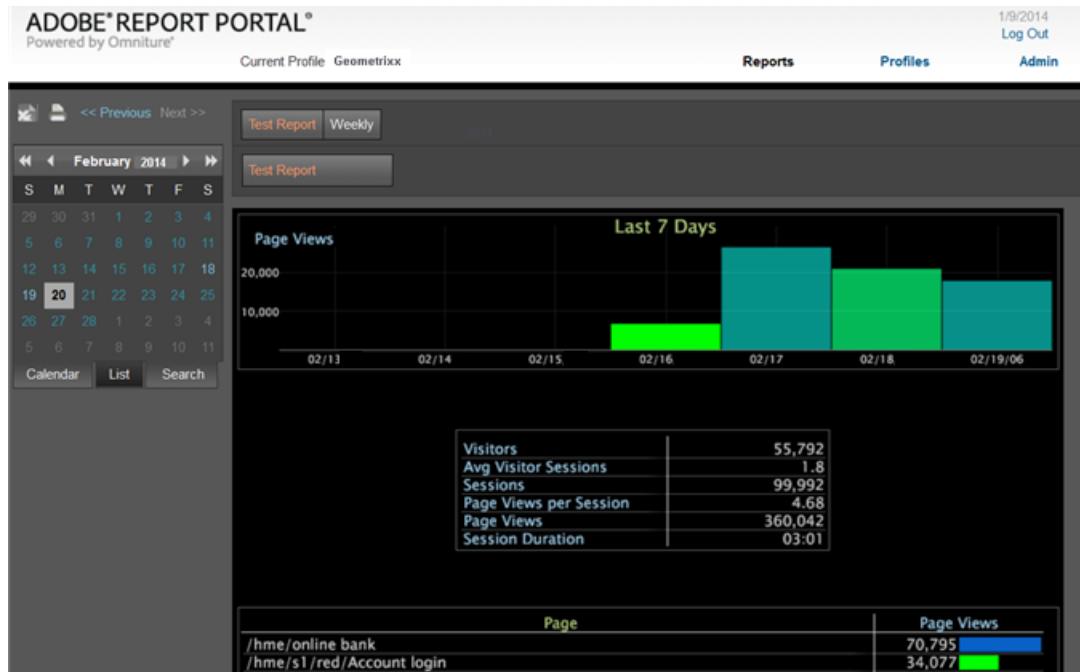
Opening the Report Portal

Steps to open the **Report Portal**.

If you have more than one profile available, you can select the appropriate profile from the drop-down box. If you have only one profile, you will not see a drop-down box. If you have only one profile and authentication is disabled, you are sent directly to the **Report Portal** home page.

1. From your web browser, open the **Report Portal** using the following URI format:
`http://ServerAddress/PortalName`
Example: `http://dataworkbench/ReportPortal`
2. If you have more than one profile available, select a profile from the **Profile** drop-down menu.

3. If **Report Portal** prompts you for log-in credentials, enter an account name and password, then click **login**. The **Report Portal** will open.

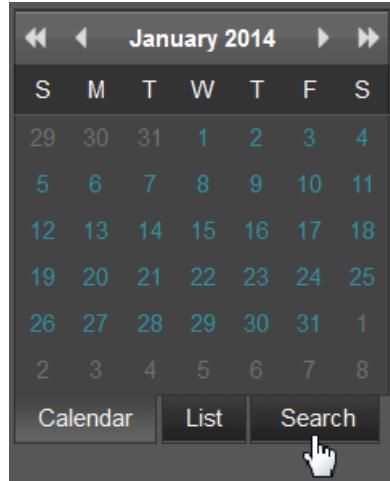


Report buttons	Select reports by clicking a button above the visualization pane.
Calendar	Click the Calendar button in the left pane, then click a to view the reports for the selected day.
List	Click the List button in the left pane to view dates listed by date.
Search	To find specific reports, click the Search button to open the <i>Advanced Search in the Report Portal</i> on page 670 feature.

Advanced Search in the Report Portal

The Report Portal provides a search feature to find reports based on a tree view of folders, tree view of report dates, and an ad hoc search.

1. Click the **Search** tab in the left pane of the Report Portal.



A new web page will open displaying the **Search** page with the **Folder Tree** option as the default view.



2. Select other view options from the buttons displayed at the top of the page.

Search button	Description
Folder tree	Select reports based on the custom hierarchy of report folders. These folders are defined and organized as a Working with Report Sets on page 661.
Date tree	Select folders based on a hierarchy of dates defined by Report Server. The reports are organized based on the date (year/month/day) they were run.
Search	<p>Perform ad hoc searches across all of your reports. If you know the complete name of the report, you can enter it directly into the search box.</p> <p>In addition:</p> <ul style="list-style-type: none"> • Type a <space> key to see a drop-down list of beginning characters or names for all existing reports. • Enter the first name of the report and type a <space> key to see a drop-down list allowing you to auto-fill the name for an existing report.
Deselect all	Click to clear all selected reports.
Show selected	<p>Click to display any reports selected in the Folder tree or Date tree views. The Report Portal will open in a new window and display each report as .png file. If multiple reports are selected, then each individual report can be selected from a tab at the top of the page.</p> <p>Click the Print button on this page to print the selected report or click Download to store locally.</p>

Installing Report Server

Navigation title:

Detailed instructions for installing the Report software and configuring it for report generation and distribution.

Before You Begin

Navigation title:Before You Begin

For some of the features of Report Server to work, you must provide and configure hardware or software before installing.

Basic Report Server Requirements

The reports that are output can either be in the form of .PNG images or .XLS spreadsheets placed in a file system, or as emails. Hardware requirements are identical to the [data workbench client](#).

The following requirements exist for Report Server:

- Access to file system for output of data (network share, or local drive).
- Access to configured SMTP server.
- Microsoft Excel 2010 64-bit or above installed on Report Server. See [Considerations for server-side Automation of Office](#) for additional information.

Additional Requirements

- **Install an Appropriate Graphics Adapter.** To properly render reports, the machine on which you install Report Server must have an appropriate graphics adapter installed.
- **Install Microsoft Excel to Generate Reports as Excel Files.** To generate and distribute reports as Microsoft Excel files (.xls or .xlsx), the machine on which you install Report Server must have the appropriate version of 64-bit Microsoft Excel installed and registered. If Excel has not been registered and Report Server tries to access it for the first time, Excel will display a registration dialog box. If you are not sure whether the copy is registered, start Excel manually and if a registration dialog box appears, complete the registration process.



Note: When you generate reports as Excel files, you are opening a new instance of Excel. For more information about this process, see <http://support.microsoft.com/kb/257757>.

- **Provide access to an SMTP server to distribute reports by email.** If you would like to distribute reports in email, Report Server must be able to connect to an SMTP server, and the appropriate ports to the email forwarding service must be opened.

Installation Overview

Navigation title:Installation Overview

Steps to install and configure Report Server software.

The following tasks must be completed in order:

1. Install the Report Server program files.
2. Download and install the Report Server digital certificate. See [Downloading and Installing the Digital Certificate](#) on page 673.
3. Configure the connection between Report Server and the data workbench server (InsightServer64.exe). See [Configuring the Connection to the Insight Server](#) on page 675.
4. Update Report Server with a language file (.zbin file).
See [Update Report Server with a language file \(.zbin file\)](#) on page 676.
5. Update the access control file on the data workbench server machine to enable Report Server to access the data workbench server. See [Enabling Access to the Data Workbench Server](#) on page 676.
6. Edit the communications file on the master data workbench server machine to display Report Server's status in the **Master Server Detailed Status** interface. See [Displaying Report Server Status](#) on page 677.
7. Register Report as a Windows service. See [Registering Report Server as a Windows Service](#) on page 678.

Installing the Report Program Files

Navigation title:Installing the Report Program Files

Steps to install the Report Server program files.

Report Server is a client of the Data Workbench Server and is licensed to run on a single machine. Use the following procedure to install the Report Server program files on your computer. Install Report Server on a separate machine from the Data Workbench Server.

1. On your Windows machine, create a directory in which to install Report Server.

Example: D:\Adobe\Report

2. Go to the download site provided by Adobe and locate the installation file for Report Server.

3. Extract the contents of the installation file to the directory you created above.

4. To export to Excel in Windows 2008 systems, create one of the following directories:

For Windows 2008 Server 64-bit: C:\Windows\SysWOW64\config\systemprofile\Desktop.

5. Right-click the folder.

6. Select **Properties**.

7. Select **Security**.

8. Click **Edit**.

9. Add **Everyone**.

10. Grant full control for **Everyone**.

Downloading and Installing the Digital Certificate

Navigation title:Downloading and Installing the Digital Certificate

After you have installed the Report Server program files, you must download and install the digital certificate provided to you by Adobe.

Understanding Digital Certificates

Navigation title:Understanding Digital Certificates

Adobe uses X.509 digital certificates to identify and authenticate the client and server components that make up an implementation.

When you install Report Server, you must install the digital certificate that authorizes a named individual (for example, Jane Smith) to use the installed client application.



Note: If you need to migrate Report Server to another machine or another named user, you must obtain a new certificate from Adobe. To do so, contact Adobe Customer Care.

Report Server presents this digital certificate to gain access to a server component. An administrator of the server component can restrict access to server resources based on the common name or organizational unit values that appear in the client's certificate.

The X.509 digital certificates installed with Adobe applications also enable its client and server components to exchange information over Secure Sockets Layer (SSL). SSL secures transmissions over HTTP using a public-and-private key encryption system. Adobe's implementation of SSL supports 1024-bit RSA keys and uses a 128-bit RC4 encryption algorithm.

In addition to security, the digital certificate that you install also functions as a license key that enables you to run the installed Report Server. To function properly, a digital certificate must be node-locked and current, or the application will not start.

Node-locked Certificates

A node-locked certificate is a digital certificate that has been registered to the machine on which it is installed. Node locking permanently associates a certificate with a specific node identifier (a value that uniquely identifies a particular machine). To node lock your certificate, your machine must have Internet access to the Adobe License Server or to a proxy server that has access to the License Server.

If you are installing on a machine that cannot access the Internet, you must obtain and install a special pre-locked certificate as described in [Using Digital Certificates on Machines Without Internet Access](#) on page 674 on this page.

If you are installing on a machine that can access the Internet, your digital certificate is node-locked automatically the first time that you start Report Server. After being node-locked, the certificate cannot be used on any other machine. If you need to migrate Report Server to another machine, you must obtain a new, unlocked certificate from Adobe.

Current Certificates

Besides being node-locked, a digital certificate must be current. To remain current, your certificate must be revalidated on a regular basis, (generally every 30 days, but can vary depending on your agreement with Adobe). If your machine has Internet access, the revalidation process is completely transparent. Report Server automatically connects to the License Server and revalidates the certificate when necessary. If your machine does not have Internet access, you need to manually install updated certificates as described in the following section.

Using Digital Certificates on Machines Without Internet Access

If you are installing on a machine that cannot access the Internet, you must request a pre-locked certificate for your installation of Report Server. A pre-locked certificate is a digital certificate that Adobe manually locks to the node identifier for the machine.

To request a pre-locked certificate, you need to send the node identifier and your certificate number to Adobe Customer Care. To obtain the node identifier for your machine, contact Adobe Customer Care to request the **Adobe Node Identifier** utility. You also can obtain the node identifier from the alert that Report Server issues when it attempts to connect to the License Server and cannot. When you receive the pre-locked certificate, install it as described in the last two steps of [Digital Certificate Installation Procedures](#) on page 674.

When the certificate needs to be re-validated, you must download a new validated certificate from the License Server and reinstall it on your machine (unless your agreement with Adobe states otherwise).

Digital Certificate Installation Procedures

Navigation title:Digital Certificate Installation Procedures

Steps to download and install the digital certificate.

1. Open your web browser to <https://aap.adobe.com>.



Note: Your browser might prompt you to present a digital certificate at this point. If it does, just click **Cancel** to dismiss the dialog box.

2. On the login screen, enter the Account Name and the Password that you received from Adobe, then click **login**.
3. Locate the certificate that has been issued for your instance of Report Server (*Your Name.pem*) and click the icon associated with that certificate.
4. When prompted to save the certificate, click **Save**.
5. Download the file to the Certificates folder in the directory where you installed Report Server.

This folder already contains a certificate file named `trust_ca_cert.pem`. Both certificate files must always be present for Report Server to function.

Configuring the Connection to the Insight Server

Navigation title:Configuring the Connection to Data Workbench Server

Before you can generate reports and alerts, you must configure Report Server to specify the address of the Insight server and identify the profiles that you want it to report against.



Note: Until you configure Report Server as described below, you can not run Report Server successfully. If you attempt to run Report Server with the non-configured file that is installed with the program, Report Server produces a stream of errors.

To configure Report Server

1. With Windows Explorer, navigate to the directory where you installed Report Server.
2. Open the `ReportServer.cfg` file in Notepad and modify the file as desired.

The following sample `ReportServer.cfg` contains only the parameters included in the `ReportServer.cfg` file by default (and highlights the required parameter settings). If you contact the Adobe License Server through a proxy server, you need to add the Licensing vector and its parameters. See [ReportServer.cfg Parameters](#) on page 701 for a detailed description.

```
Fonts = vector: 0 items
Gamma = float: 1.6
Network Location = string: NetworkLocationName
Servers = vector: 1 items
  0 = serverInfo:
    Address = string: ServerIPAddress
    Name = string:
    Port = int: 443
    Proxy Address = string:
    Proxy Password = string:
    Proxy Port = int: 8080
    Proxy User Name = string:
    SSL Client Certificate = string: ReportCertFileName.pem
    SSL Server Common Name = string: ServerCommonName
    Use SSL = bool: true
Result Memory Limit (KB) = double: 100000
Maximum Slice Size = int: 30
Use OpenGL Hardware Rendering = bool: true
Reporting = :
  Profiles = vector: 1 items
    0 = ReportProfile:
      Server = string: ServerCommonName
      Profile = string: ProfileName
Update Interval (minutes) = int: 10
Completion Message Interval (seconds) = int: 600
Status interval (seconds) = int: 600
SMTP Server for Errors = string: SMTPServerHostName
SMTP Server for Errors Username = string: SMTPServerUsername
SMTP Server for Errors Password = string: SMTPServerPassword
SMTP Server for Errors Send From = string: SenderAddress
SMTP Server for Errors Send To = string: RecipientAddresses
```

3. Save and close the file.

Update Report Server with a language file (.zbin file)

For all languages, Report Server 6.0 and later requires the "insight.zbin" file copied to the Report Server root folder.

Update the Report Server language files:

1. Add the renamed "insight.zbin" file to the root ReportServer directory.
2. The Report Server configuration file (reportserver.cfg) requires font settings for double-byte languages. For example, Chinese requires the addition of fonts using SimSun:

Report Server.cfg - Add Fonts

```
Fonts = vector: 2 items
  0 = string: SimSun
  1 = string: Arial
```

3. A parameter for Report Server 6.0 needs to be passed in the command line for localization, for example:

```
ReportServer.exe -Locale -zh-cn
ReportServer.exe -Locale -en-us
```



Note: If a locale is not specified, then the Report Server defaults to English.

Follow the steps to launch the ReportServer as a service with the Locale parameters:

- a. Launch a Command Prompt as an Administrator.
- b. Navigate to the ReportServer install folder.
- c. Type the following command to start the service:
 - For English: ReportServer.exe -RegServer -Locale -en-us
 - For Chinese: ReportServer.exe -RegServer -Locale -zh-cn
4. To verify if the ReportServer is running with the correct parameters:
 - a. Open up Windows Service Manager.
 - b. Right-click Adobe Insight Report Server - Properties.

The path to executable will contain the parameters:

```
ReportServer.exe -Service ReportServer -Locale -en-us
```

Enabling Access to the Data Workbench Server

Navigation title:Enabling Access to the Data Workbench Server

To connect to a data workbench server, Report Server must have permission to access that server.

You grant access to a data workbench server by adding Report Server's common name (as assigned on Report Server's digital certificate) to the server's access control file.



Note: When working in a clustered environment, Report Server should be configured to access the master data workbench server to avoid synchronization issues. In data workbench you can view information about processing servers in your cluster using the **Related Servers** menu item in the **Servers Manager**. For more information about the **Servers Manager**, see the Administrative Interfaces chapter of the *Data Workbench User Guide*.

The following procedure describes how to manually add Report Server to the access control file on a data workbench server. To update the access control file in this way, you must have file-system access on the machine where the data workbench server is installed.

You also can update the server's access control file using the **Server Files Manager** in data workbench. To do this, your data workbench client must have administrative privileges on the server.

For more information about the **Server Files Manager**, see the Administrative Interfaces chapter of the *Data Workbench User Guide*.

To configure access to a data workbench server

1. Navigate to the Access Control folder in the directory where you installed the data workbench server (InsightServer64.exe).
Example: C:\Adobe\Server\Access Control
2. Open Access Control.cfg in a text editor such as Notepad.
3. Locate the **Report Server AccessGroup** and add Report Server's common name to this group as highlighted in the following file fragment. (Type the common name exactly as it appears on Report Server's digital certificate.)

```
. . .
5 = AccessGroup:
Members = vector: 1 items
0 = string: CN: ReportCommonName
Name = string: Report Server
Read-Only Access = vector: 5 items
0 = string: /Profiles/$
1 = string: /Status/
3 = string: /Software/
4 = string: /Addresses/
5 = string: /Users/$
Read-Write Access = vector: 3 items
0 = string: /Profiles/
1 = string: /Users/%CN%/
2 = string: /ReportStatus.vsp
. . .
```

4. Save the file.

Displaying Report Server Status

Navigation title:

The **Detailed Status** interface in data workbench is useful for troubleshooting errors or other issues with the Data Workbench Server and Report Server machines that are clients of Data Workbench Server.

To view Report's status in the **Master Server Detailed Status** interface, you must add a report status server to the **Servers** vector in data workbench server's Communications.cfg file. The following procedure describes how to add the report status server to the Communications.cfg file:

For more information about **Detailed Status** interfaces, see the Administrative Interfaces chapter of the *Data Workbench User Guide*.

To add a Report Status Server

1. Navigate to the Components folder in the directory where you installed the data workbench server (InsightServer64.exe.).
Example: C:\Adobe\Server\Components
2. Open Communications.cfg in a text editor such as Notepad.

-
3. Locate the **Servers** vector and add the report status server to this vector as highlighted in the following file fragment.

```
 . . .
Servers = vector: 17 items
 0 = FileServer:
    Local Path = string: Audit\\
    URI = string: /Audit
 1 = FileServer:
    Local Path = string: Bin\\
    URI = string: /Bin
 2 = FileServer:
    Local Path = string: Components\\
    URI = string: /Components
  . . .
 16 = ReportStatusServer:
    URI = string: /ReportStatus.vsp
```

4. Update the items count for the **Servers** vector (that is, increment the items value by one) as highlighted in the file fragment in the previous step.
5. Save the file.

Registering Report Server as a Windows Service

Navigation title: Registering Report as a Windows Service

Steps to register and run Report Server.

Before you perform this procedure, identify the Windows account under which the Report Server service will run. Make sure that this account has the proper permissions to access the location where generated reports are stored (that is, do not use the **Local System Account**).

Use the procedure below to start Report Server. When you start Report Server for the first time, it automatically registers itself as a Windows service.

When you start Report Server for the first time, it automatically connects to the Adobe License Server to register your digital certificate. To complete the registration process successfully, your machine must be connected to the Internet when you execute the following steps.

1. In Windows, navigate to the directory where you installed Report Server.
Example: D:\Adobe\Report
2. Double-click **ReportServer.exe**.
3. To confirm that Report Server is running correctly, click **Start > Control Panel > Administrative Tools > Services**. This command sequence can vary depending on which version of Windows you are using.
4. In the service list, locate the entry for Report Server and confirm that its status is Started and its startup type is Automatic.
5. Do the following to specify the user account under which Report Server will execute:
 1. Double-click **Report Server** to open the **Properties** window.
 2. Select the **Log On** tab.
 3. Select the **This account** radio button.
 4. Type or browse for the account name. This account must have permission to access the location where generated reports are stored.



Note: If Report Server distributes reports as Microsoft Excel (.xls or .xlsx) files, make sure that the account also has permission to run Microsoft Excel.

-
5. Enter and confirm the password for the account.
 6. Click **OK**.
 6. Right-click the Report Server service and select **Restart** to restart the service under the account you specified.
 7. To check whether Report Server experienced any errors during start-up, click **Start > Control Panel > Administrative Tools > Event Viewer**. This command sequence can vary depending on which version of Windows you are using.
 1. In the left pane of the **Event Viewer** window, select the Applications log.
 2. In the right pane, look for events with Adobe in the **Source** column.
 3. If you find an error from Adobe, double-click the error to display the **Event Properties** window. This window provides detailed information about the error.



Note: After the Report Server service starts, the file `ReportServer.log` is created in the Report Server Trace directory. This file is also useful for troubleshooting issues with Report Server.

You have completed the installation of Report Server. Report Server is designed to run continuously. If you restart the machine, Report Server restarts automatically. If you need to start and stop Report Server manually, you can do so using the **Services** control panel in Windows.

Upgrading and Uninstalling Report Server

Navigation title:

Information about upgrading and uninstalling your Report Server software.

- [Upgrading Report Server](#) on page 679
- [Uninstalling Report Server](#) on page 680

Upgrading Report Server

If you are upgrading to Report Server 5.4, you can use the instructions to upgrade your Report Server software. If you are using Report Server 3.6 or earlier, contact Adobe for assistance with your upgrade.

To upgrade Report Server 5.4, use data workbench to copy an upgrade file to the data workbench server to which Report Server connects. After doing so, Report Server instances automatically upgrade themselves when they connect to that server and load a profile.



Note: Before upgrading Report Server, make sure that you have properly upgraded your data workbench server software as well as the profiles running on the data workbench server. For more information, contact Adobe Consulting Services.

To perform the following procedure, you first must obtain the upgrade file for Report Server.

To upgrade to Report Server 5.4 and later versions

1. Make a backup of all of the files under `E:\Portal` and remove all files and folders within this directory.
2. Copy the contents of the new build into `E:\Portal`.
3. Modify `global.asa`, `email.asp`, and `TopNavigation.xml` as per the instructions in the previous section.
4. Copy the `users.mdb` from your backup.



Note: If you did not previously generate reports with a .png output, you need to go into the individual report folders and modify the `reports.xml` to include a report format of png. Otherwise you may get a 500 error. Your original `reports.xml` would look something like the following:

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<REPORTS>
  <REPORT format="xls">
    <NAME>Dashboard</NAME>
    <PATH>Dashboard.xls</PATH>
    <WEB_PATH>Dashboard.xls</WEB_PATH>
  </REPORT>
</REPORTS>
```

It would need to be modified to the following:

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<REPORTS>
  <REPORT format="xls">
    <NAME>Dashboard</NAME>
    <PATH>Dashboard.xls</PATH>
    <WEB_PATH>Dashboard.xls</WEB_PATH>
  </REPORT>
  <REPORT format="png">
    <NAME>Dashboard</NAME>
    <PATH>Dashboard.png</PATH>
    <WEB_PATH>Dashboard.png</WEB_PATH>
  </REPORT>
</REPORTS>
```

5. In the `report.cfg`, include an output format of png and save. Going forward, it should generate reports in png format.

To upgrade to Report Server 4.0

1. On the data workbench computer, copy the Report Server upgrade file to the `Temp\Software` folder in the directory where data workbench is installed.
2. Start data workbench and load the **Configuration** profile.
3. Click the **Configure Connection to Servers** thumbnail.
4. In the **Servers Manager**, right-click the data workbench server icon and click **Server Files**.
5. In the Software folder, open the `Report Server` folder.
6. Right-click the **Temp** check mark for `ReportServer.exe` and select **Save to > <server name>**.

Uninstalling Report Server

To uninstall Report Server

1. Unregister the **Report Windows** service.
 - a. Open a command prompt and navigate to the bin sub-directory in the folder where you installed the data workbench server (`InsightServer64.exe`) Example: `D:\Adobe\Report\bin`
 - b. At the command prompt, execute the following command to stop and unregister it as a service under Microsoft Windows: `visualreport /unregserver`
2. Delete the Report Server installation directory.

Administering Report Server

Navigation title:

Conceptual information about administrative tasks.

Configuring Report.cfg Files

Navigation title:Configuring Report.cfg Files

In the **Report Portal**, you can view reports generated by Report Server as Excel files (.xls or .xlsx) or .png files.

To display a report set in the **Report Portal**, you must set the following parameters in the `Report.cfg` file for that report set:

- In the **Output Root** parameter, specify the document root of the web server used for your portal.
- In the **Report Types** parameter, specify Excel, png, and/or thumbnail as the report types that you want to generate.

When Report Server generates the reports in the formats that you specified, it places those files in the document root of the web server, which is where during installation you configure the **Report Portal** to access the reports.

For more information about the specific `Report.cfg` parameters, see [Report.cfg Parameters](#) on page 704.

Re-validating the Digital Certificate

Navigation title:Re-validating the Digital Certificate

After installation, the digital certificate issued by Adobe acts as a key that enables you to run Report Server.

Recommended Frequency: As needed

To function properly, a digital certificate must be current.

To remain current, your digital certificate must be re-validated on a regular basis (generally, every 30 days, but this can vary depending on your agreement with Adobe). If your machine has Internet access, the revalidation process is completely transparent. Report Server connects automatically to the Adobe License Server and re-validates the certificate when necessary. If your machine does not have Internet access, you must download a new, validated certificate from the Adobe License Server and install it on your machine using the steps provided in [Downloading and Installing the Digital Certificate](#) on page 673.

Confirming that the Report Server is Running

Navigation title:Confirming that the Report Service is Running

To ensure your implementation is functioning properly, you should confirm that the Report Server service is running.

Recommended Frequency: Every 5-10 minutes

- Click **Start > Control Panel > Administrative Tools > Services**. Report should be listed with a status of “Started.” If the service is not running when it should be, restart the service and check the **Event Viewer (Start > Control Panel > Administrative Tools > Event Viewer)** for errors.

If you need help determining the cause of a problem, please contact Adobe Support.

Monitoring Disk Space

Navigation title:Monitoring Disk Space

You should regularly monitor available disk space so that all Report Server machines continue to perform at the highest level possible.

Each Report Server machine stores the following types of data:

- Operating system data
- Report data

-
- System data



Note: Adobe consultants can evaluate your use scenario to help you project the amount of data storage that your Adobe software application generates and requires. To request such an evaluation, please contact Adobe Consulting Services.

- [Monitoring Report Data Space](#) on page 682
- [Backing Up Operating System, Report, and System Data](#) on page 682

Monitoring Report Data Space

Recommended Frequency: Every 5-10 minutes

Make sure that you provide enough disk space to accommodate your reports in the `Reports` folder within the `Report` installation directory.

Backing Up Operating System, Report, and System Data

Recommended Frequency: Daily

Report, system, and operating system data must be backed up regularly and diligently using your company's normal back up and disaster recovery systems.

Restarting Report Server

Navigation title:

When restarting a Report Server machine, you first must stop the Report Server service.

Recommended Frequency: Only when necessary

To stop the Report Server service

1. Click **Start > Control Panel > Administrative Tools > Services**.
2. Select the Report Server service and click the **Stop** button.

When you restart the Report Server machine, the Report Server service starts automatically.

Reviewing Report Status

Navigation title:Reviewing Report Status

Information about your Report Server status and report set status.

- [Report Server Status](#) on page 682
- [Report Set Status](#) on page 683

Report Server Status

Recommended Frequency: Only when necessary

Report sends status information to the data workbench server every two minutes regarding the status of the Report Server. This information can be seen under the **Report Server Status** node in the **Detailed Status** interface.

To open the Detailed Status visualization

1. In data workbench, right-click in a workspace and click **Admin > Servers**.
2. In the **Servers** interface, right-click the icon of the data workbench server that the Report machine connects to and click **Detailed Status**.
3. Click **Report Server Status**.

If more than one Report is connected to the data workbench server, an entry appears for each **Report Server** in the Status vector. The two-minute interval can be overridden by specifying a value in the Status Interval (seconds) parameter in the **Reporting** node of the ReportServer.cfg file.

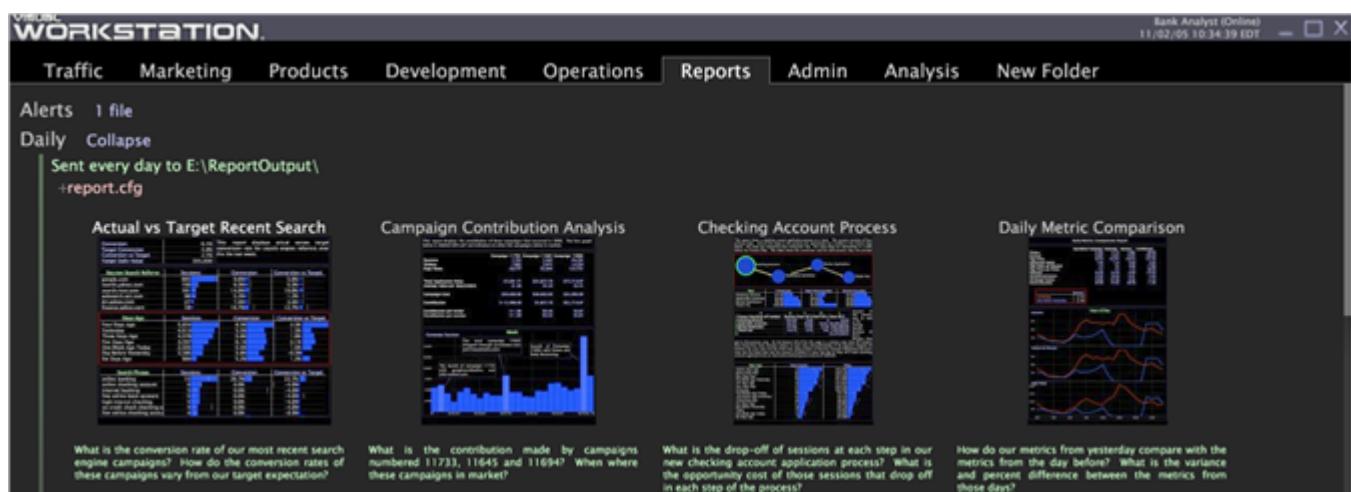
For information about the ReportServer.cfg file, see [Configure the Report Set](#) on page 662. For information about configuring Report, see [Installing Report Server](#) on page 671.

For more information about **Detailed Status**, see the Administrative Interfaces chapter of the *Data Workbench User Guide*.

Report Set Status

Recommended Frequency: Only when necessary

Report transmits status information for each report set to the Data Workbench server. Basic information, such as when a report set is generated and where it is distributed, displays in data workbench above the report set in green text. While running reports, Report Server outputs a message every two minutes indicating the percentage complete of the current queries. This two-minute interval can be overridden by specifying a value in the Completion Message Interval (seconds) parameter in the **Reporting** node of the ReportServer.cfg file.



Note: If an error occurred while running a report, the error is indicated in red text below the thumbnail of that report. You can right-click the workspace to display the full error message.

Working with Accounts

Navigation title:Working with Accounts

Access to and permissions within your **Report Portal** are controlled using individual user and group accounts.

Certain users can be designated as administrators, enabling them to manage all of the accounts.

Both individual user and group accounts specify which profiles and tabs that particular user or group can access. The main difference is that individual accounts enable users to reset their passwords, while group account passwords can be reset only by a **Report Portal** administrator.

The following sections provide you with the information that you need to view, create, and edit individual user and group accounts.

Viewing Accounts

Navigation title:Viewing Accounts

Information about viewing your accounts.

In the **Report Portal**, click the **Admin** tab. The **Admin** page appears listing all individual user and group accounts.

Select Account to Update or Remove					
a b c d e f g h i j k l m n o p q r s t u v w x y z / all / new account					
Account Name	Account Expiration	Admin	Profile Access	Tab Access	
i <input checked="" type="radio"/> test	1/1/2020	True	ALL	ALL	
select			remove		

The following information displays for each account:

This field ...	Indicates ...
Account Name	The name of the individual user or group account. The letter that displays to the left of the user or group name indicates whether this is an individual (i) or group (g) account. Individual accounts enable users to reset their passwords, while groups do not. An administrator is the only person able to reset the password for a group account.
Account Expiration	The date, in MM/DD/YYYY format, until which this user or group is allowed to use Report Portal .
Admin	Whether this user is allowed to create, update, and delete user accounts as well as edit notes associated with each report. A setting of True indicates that this is an admin account.
Profile Access	The profiles that this user is allowed to access (for example, ProductSales). If the user is allowed to access all profiles associated with Report Portal , "ALL" displays.
Tab Access	The tabs that this user is allowed to access (for example, Admin). If the user is allowed to access all tabs in the Report Portal , "ALL" displays.

Adding New Accounts

Navigation title: Adding New Accounts

Steps to add new accounts.

1. In the **Report Portal**, click the **Admin** tab. The **Admin** page appears.

Select Account to Update or Remove					
a b c d e f g h i j k l m n o p q r s t u v w x y z / all / new account					
Account Name	Account Expiration	Admin	Profile Access	Tab Access	
i <input checked="" type="radio"/> test	1/1/2020	True	ALL	ALL	
select			remove		

2. On the right side of the page, click **new account**. The **Create New Account** page appears.

Create New Account

<input type="text" value="account name"/>	
<input type="text" value="email"/>	
<input type="text" value="password"/>	
<input type="text" value="confirm password"/>	
<input type="text" value="ALL"/>	
profile access	
<input type="text" value="ALL"/>	
tab access	
<input style="width: 100px; height: 20px; border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;" type="text" value="individual"/>	▼
account type	
<input style="width: 100px; height: 20px; border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;" type="text" value="active"/>	▼
status	
<input style="width: 100px; height: 20px; border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;" type="text" value="false"/>	▼
admin	
<input type="text" value="expiration date (mm/dd/yyyy)"/>	
<input style="width: 100px; height: 25px; border: 1px solid #ccc; background-color: #f0f0f0; border-radius: 5px; margin-right: 10px;" type="button" value="insert"/> <input style="width: 100px; height: 25px; border: 1px solid #ccc; background-color: #f0f0f0; border-radius: 5px;" type="button" value="cancel"/>	

3. Complete all of the fields on this page as indicated in the following table:

In this field ...	Specify ...
account name	The name of the account that the user must provide when logging on to Report Portal .
email	The email address of the user or group.
password	The password the user must provide when logging on to Report Portal .

In this field ...	Specify ...
confirm password	The password the user must provide when logging on to Report Portal .
profile access	The profiles that this user is allowed to access (for example, ProductSales). To allow access to multiple profiles, separate the names by commas. If the user is allowed to access all profiles associated with Report Portal , type "ALL."
tab access	The tabs that this user is allowed to access (for example, Admin). To allow access to multiple tabs, separate the names by commas. If the user is allowed to access all tabs in the Report Portal , type "ALL." This field, in conjunction with the account type field, is very useful for defining group access rights.
account type	Whether this account is for an individual or a group. Individual accounts enable users to reset their passwords, while groups do not. An administrator is the only person able to reset the password for a group account.
status	Whether this account is active or inactive. The default value is active. To deactivate a user account, select inactive.
admin	Whether to allow this user to create, update, and delete user accounts as well as edit notes associated with each report. The default setting is false. To make this an admin user, select true.
expiration date	The date, in MM/DD/YYYY format, until which this user is allowed to use Report Portal .

- #### 4. Click **insert**.

Editing Existing Users

Navigation title:Editing Existing Users

Steps to edit existing user accounts.

1. In the **Report Portal**, click the **Admin** tab. The **Admin** page appears.

Select Account to Update or Remove																													
a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	/	all	/	new account
Account Name					Account Expiration					Admin					Profile Access					Tab Access									
 test					1/1/2020					True					ALL					ALL									
												<input type="button" value="select"/>					<input type="button" value="remove"/>												

- Click the letter representing the first letter of the account name that you want to edit. For example, if you want to edit the account “Marketing,” you would click the letter “M.”

A list of the account names starting with that letter displays.

3. Select the account name that you want to edit, then click the **select** button. The **Edit Account Info** page appears.

Edit Account Info

test
account name

test
email

old password

new password

confirm password

ALL
profile access

ALL
tab access

individual ▾
account type

active ▾
status

true ▾
admin

1/1/2020
expiration date (mm/dd/yyyy)

[update](#) [cancel](#)

The screenshot displays a web-based form titled 'Edit Account Info'. The form contains several input fields and dropdown menus. At the top, there are fields for 'account name' (containing 'test') and 'email' (containing 'test'). Below these are fields for 'old password', 'new password', and 'confirm password', all of which are currently empty. There are two large dropdown menus labeled 'profile access' and 'tab access', both set to 'ALL'. Underneath these are three more dropdown menus: 'account type' (set to 'individual'), 'status' (set to 'active'), and 'admin' (set to 'true'). A date field shows '1/1/2020' with the placeholder 'expiration date (mm/dd/yyyy)'. At the bottom of the form are two buttons: 'update' and 'cancel'.

4. Change only the fields on this page that need to be updated. The following table provides descriptions of each of these fields:

In this field ...	Specify ...
email	The email address of the user.
old password	The current password, which is needed to proceed when editing an admin account or when resetting the password for a non-admin account.
new password	The new password the user must provide when logging on to Report Portal .
confirm password	The new password the user must provide when logging on to Report Portal .
profile access	The profiles that this user is allowed to access (for example, ProductSales). To allow access to multiple profiles, separate the names by commas. If the user is allowed to access all profiles associated with Report Portal , type "ALL."
tab access	The tabs that this user is allowed to access (for example, Admin). To allow access to multiple tabs, separate the names by commas. If the user is allowed to access all tabs in the Report Portal , type "ALL." This field, in conjunction with the account type field, is very useful for defining group access rights.
account type	Whether this account is for an individual or a group. Individual accounts enable users to reset their passwords, while groups do not. An administrator is the only person able to reset the password for a group account.
status	Whether this account is active or inactive. The default value is active. To deactivate a user account, select inactive .
admin	Whether to allow this user to create, update, and delete user accounts as well as edit notes associated with each report. The default setting is false. To make this an admin user, select true.
expiration date	The date, in MM/DD/YYYY format, until which this user is allowed to use Report Portal .

5. Click **update**.

Installing the Report Portal

Navigation title:Installing Report Portal

To use **Report Portal**, you must install and configure a set of application server pages (ASPs) on IIS.

Before You Begin

The procedures in this chapter describe how to install and configure **Report Portal**. To complete these procedures, you must:

- Have Microsoft IIS installed and know its version.
- Know the names of the report sets whose reports are displayed by **Report Portal**.
- Know the location of the directory in which Report Server saves the output for these report sets. Make certain that the IIS application server has access to this directory.



Note:

- To be viewed using **Report Portal**, reports must follow specific naming conventions. Additionally, the directory into which reports are saved must follow a prescribed structure. For a description of these requirements, see [Ensuring that Your Report Sets are Compatible with Report Portal](#) on page 698.
- Passwords in report portal are now AES-256 bit compliant. If upgrading to Report Portal 2.0, increase the Field Size for Password field from 50 to 150 in the **users.mdb** database. Increasing the field size is required to accommodate passwords with updated encryption.
- If you are upgrading to Report Portal 2.0, increase the Field Size for the **Password** field from 50 to 150 in **users.mdb** database. Increasing the field size is required to accommodate passwords with updated encryption.
- The Report Portal now features stronger hashing algorithms with salting support. If you are upgrading to **Report Portal 2.1**, add a new Text field, **PasswordSalt** with field size of 20 characters in **users .mdb** database. This field is required to store the password salt.

Install the Report Portal Application Files

Navigation title:Install the Report Portal Application Files

Report Portal is made up of a set of application server pages (ASPs) and supporting files.

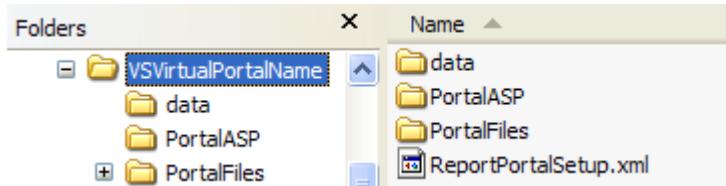
To install the **Report Portal**, you must extract these files from the distribution file that you received from Adobe and install them on the machine where Microsoft IIS is running.

To Install the Report Portal application files

1. If you have not already done so, download the installation package (.zip file) for the **Report Portal** from the Adobe FTP site.
2. On the machine where IIS is running, extract the files in the installation package to any location. This step installs the following subfolders and files in the VSVirtualPortalName folder.

Folder or File	Description
\data\users.mdb	Database containing the list of authorized Report Portal users.
\PortalASP\	Folder containing the ASP files that make up Report Portal .
\PortalFiles\	Folder containing five subfolders (Core, CSS, HTC, Images, and Output) that contain supporting files used by Report Portal .
ReportPortalSetup.xml	Configuration file you use to define the virtual directories associated with Report Portal (used with IIS 6.0 only).

The directory looks like the following example:



Note: The name of the directory may differ from the name shown in the example.

3. Rename the VSVirtualPortalName (or other name) folder to what you want to use as the root virtual directory of your **Report Portal** (hereafter referred to as *PortalName*). For more information about virtual directories, see the next section.

Map the Report Portal Pages to Virtual Directories

Navigation title:Map the Report Portal Pages to Virtual Directories

To configure **Report Portal**, you must map its application files to virtual directories.

A virtual directory defines the address that browser clients use to locate a physical resource on the IIS application server. To access **Report Portal**, clients point their browsers to the virtual directory that you assign to the portal.

The name of the virtual directory that you assign to **Report Portal** must match the name that you used for the VSVirtualPortalName folder in Step 3 of the previous section. For example, if you want to use “Portal” as the name of your **Report Portal**, you must map the portal’s files to a virtual directory named “Portal.” The following example shows the URI that clients would use to access a **Report Portal** assigned to the virtual directory **VisualReportPortal** on a server called myWebServer:

`http://myWebServer/VisualReportPortal`

The following procedures describe how to map **Report Portal** to a virtual directory on IIS 5.0, 6.0, and 7.0 or higher.

Follow the set of procedures for the version of IIS that you are using:

- [Mapping Report Portal to a Virtual Directory \(IIS 7.0 or higher\)](#) on page 690
- [Mapping Report Portal to a Virtual Directory \(IIS 5.0\)](#) on page 693
- [Edit the Session Configuration File](#) on page 695

Mapping Report Portal to a Virtual Directory (IIS 7.0 or higher)

Steps to map the **Report Portal** to a virtual directory (IIS 7.0 or higher).

Currently, most Managed Service clients have servers with the Windows Server 2008 operating system and the IIS 7.0 or higher web server.

Prerequisites

- Make sure that ASP and ASP .Net components are installed for IIS 7.0 or higher.
- Make sure the IIS Web user has **Modify** access to the E:\Portal\data\users.mdb file. You can change that by right-clicking on the **users.mdb** file and under **Properties**, go to the **Security** tab. If you do not see the IIS Web User listed or do not have the ability to add the IIS Web User to the list, simply give the **Users** group the **Modify** access.
- Make sure whatever user account is being used to run the **Application Pools** also has **Modify** access to the E:\Portal\data\users.mdb and the C:\Windows\Temp\ folders.

Installation Steps

1. On the machine where **Report Portal** is installed, start the **IIS Manager**:
Start > Administrative Tools > Internet Information Services (IIS) Manager.
2. Select **Local Machine > Sites > Default Web Site**.
3. Right-click **Default Web Site** and select **Add Virtual Directory**.
4. For an alias, enter **Portal**.
5. For the physical path, enter **E:\Portal\PortalASP**.
6. Click **OK**.

The virtual directory that you created appears under the **Default Web Site**.

7. Add the following virtual directories under the virtual directory that you just created.

Create this alias...	For this physical resource
Core	E:\Portal\PortalFiles\Core
CSS	E:\Portal\PortalFiles\CSS
Images	E:\Portal\PortalFiles\Images
Output	Physical location of the directory in which Report Server saves the output for your report sets. The output folder can be located anywhere and can be named anything. It contains a subfolder for each report set. You can delete the E:\Portal\PortalFiles\Output, but move the profiles.xml to the physical location of the Output file.

8. When finished, verify that IIS displays four new virtual directories. Make sure that the directory structure has one parent folder (with the same name as your portal) and four subfolders.
9. Click on **Application Pools**, then **DefaultAppPool** (assuming that's the one you set up with your portal).
10. Click on **Advanced Settings** and select True for the Enable 32-Bit Applications.
11. To get the **Portal** to work, you need to convert it to an application. After setting up the virtual directories, right click on the Portal virtual directory and select **Convert to Application**.

Additional Tips and Tricks

- You can download the **Portal** from Softdocs under **Softdocs > Report Portal**. You can simply download the `ReportPortal-Release-1-0-0-7.zip`.
- You no longer need the `ReportPortalSetup.xml`, so it can be deleted.
- For the sake of standardization, place the contents of this zip file into `E:\Portal`.
- To determine the SMTP server For managed services clients, you can go look here.
- Put in a request with NetOps to change the domain name entry in IIS for the report server to something friendlier
 - for example, `reports.clientname.insight.omniture.com`, so that your overall portal URL is `http://reports.clientname.insight.omniture.com/Portal`. Configure your `email.asa` file once this change has been put into place.

Mapping Report Portal to a Virtual Directory (IIS 6.0)

Steps to map the **Report Portal** to a virtual directory (IIS 6.0).

Mapping the **Report Portal** to a virtual directory on IIS 6.0 involves three separate tasks:

1. [To Edit the Configuration File](#) on page 691
2. [To Import the Configuration File into IIS](#) on page 692
3. [To Enable Active Server Pages \(ASPs\) on IIS](#) on page 693

You must complete all three tasks.

To Edit the Configuration File

1. On the machine where **Report Portal** is installed, open `\PortalName\ReportPortalSetup.xml` in a text editor such as Notepad.
2. Use the editor's find-and-replace feature to globally replace (Replace All) the string "VSVirtualPortalName" with the name of your portal. For example, if you want to use "VisualReportPortal" as the name of your **Report Portal**, you would search for "VSVirtualPortalName" and replace it with "VisualReportPortal."
3. Locate the following element in this file:

```
<IIsWebVirtualDir Location= "/LM/W3SVC/1/Root/PortalName/Output"
AccessFlags="AccessRead | AccessScript" AppFriendlyName="Output" . . .>
```

4. Set this element's **Path** attribute to the physical location of the directory in which Report Server saves the output for your report sets.

The output folder can be located anywhere, can be named anything, and contains a subfolder for each report set.



Note: This must be the same directory that you specify in the Output Root parameter in the `Report.cfg` file for a report set. For more information, see [Configuring Report.cfg Files](#) on page 681.

The following code sample shows how you would set the **Path** attribute if your reports were saved to `E:\VSReport\ReportOutput`:

```
< . . .
AppIsolated="2"
AppRoot="/LM/W3SVC/1/Root/PortalName/OutputFolder"
DirBrowseFlags="DirBrowseShowDate | DirBrowseShowTime | . . .
Path="E:\VSReport\ReportOutput"
```

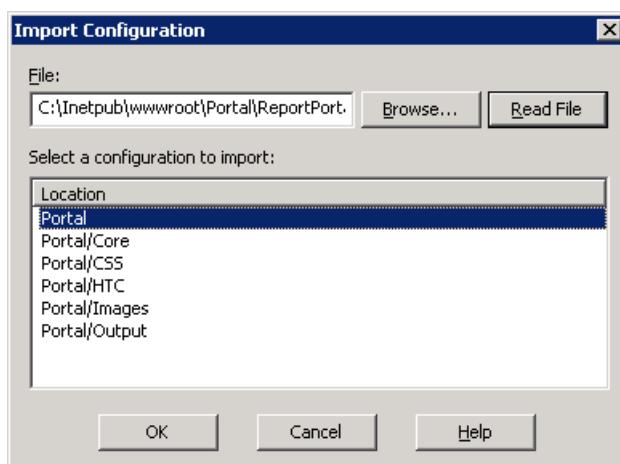


Note: It is critical that the **Path** attribute is set properly.

5. If you changed the default **Path** of the **Output** element, move the **profiles.xml** file from the **\PortalName\PortalFiles\Output** folder to the output directory that you specified in Step 4. In the example above, you would move **profiles.xml** to **E:\VSReport\ReportOutput**.
6. Verify that the **Path** attributes for all of the other **IIsWebVirtualDir** elements are mapped to the correct location by searching for all instances of **C:\Inetpub\wwwroot** and replacing each with the correct path.
7. Save the file. If you want to preserve the original file, you can save the configuration file using a new name.

To Import the Configuration File into IIS

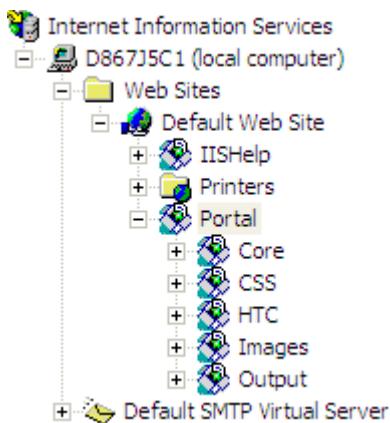
1. On the machine where **Report Portal** is installed, start the IIS Manager using **Start > Administrative Tools > Internet Information Systems (IIS) Manager**.
2. Select **(local computer) > Web Sites > Default Web Site**.
3. Right-click **Default Web Site** and select **New > Virtual Directory (from file)**.
4. Select the **ReportPortalSetup.xml** file and click **Read File**.
5. Verify that six virtual directories are listed for your **Report Portal** as shown in the following example.



If you do not see six virtual directories or if you receive an error message, click **Cancel** and examine the configuration file for errors.

6. Select the first virtual directory in the list (the one that is the parent of the other five) and click **OK**. IIS imports the mappings and add the virtual directories to the Default Web Site.

Make sure that the resulting directory structure has one parent folder (with the same name as your portal) and five sub-directories as shown in the following example.

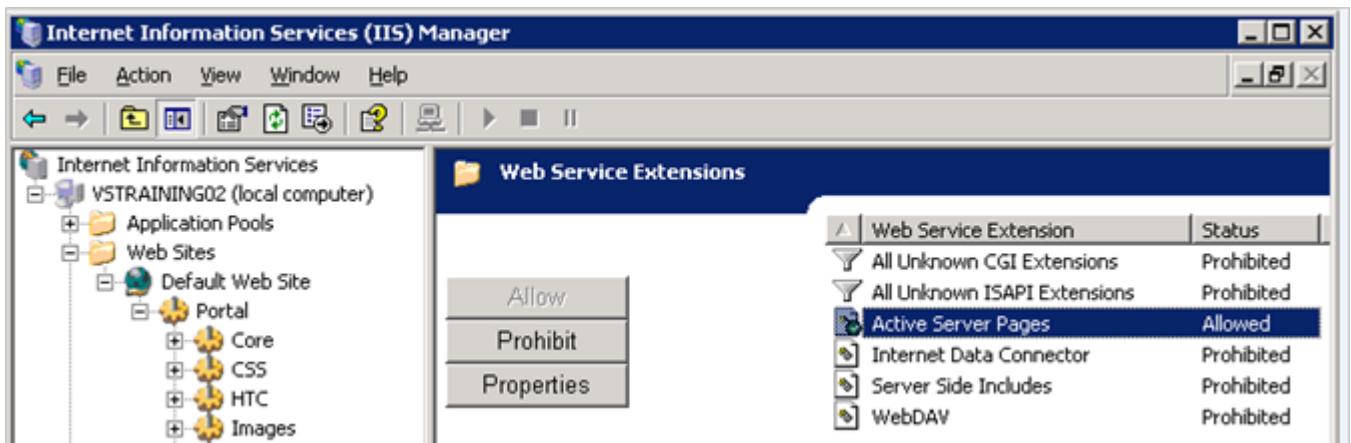


7. Click each virtual directory to ensure that IIS can locate the physical directory it represents. If IIS displays an error, right-click the virtual directory name and verify that the **Local Path** field points to the correct physical directory.

To Enable Active Server Pages (ASPs) on IIS

To use **Report Portal**, ASPs must be enabled on IIS. (By default, ASPs are disabled when IIS 6.0 is installed.) Use the following procedure to verify that ASPs are enabled on your IIS.

1. In the IIS Manager window, select (**local computer**) > **Web Service Extensions**.
2. Verify that the **Active Server Pages** extension is set to **Allowed**.



3. If their Status is Prohibited, select **Active Server Pages** and click **Allow**.
4. Close IIS Manager.

Mapping Report Portal to a Virtual Directory (IIS 5.0)

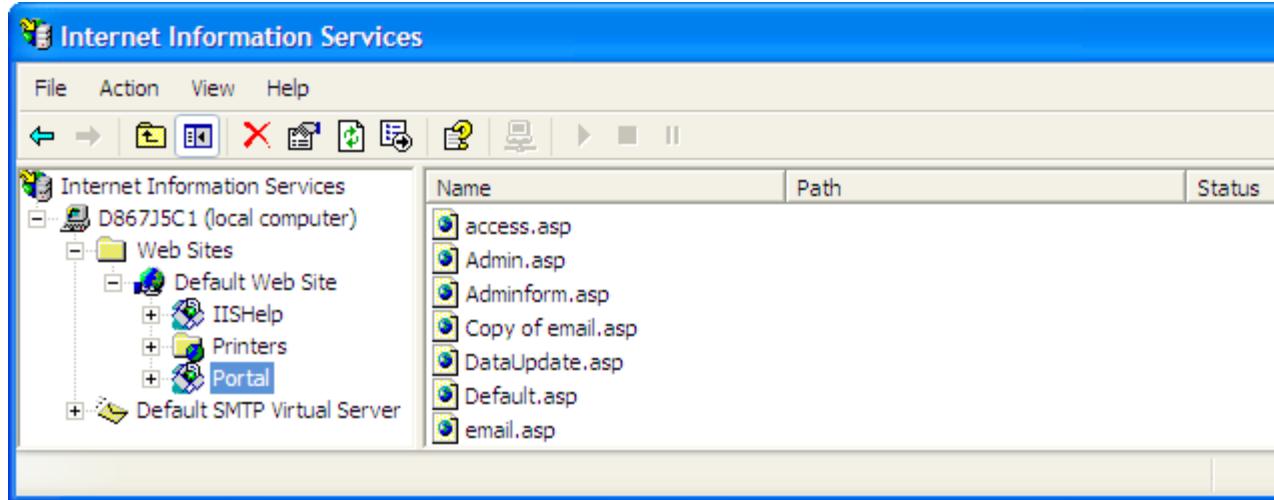
Navigation title: Mapping Report Portal to a Virtual Directory (IIS 5.0)

Steps to map the **Report Portal** to a virtual directory (IIS 5.0).

1. On the machine where **Report Portal** is installed, start the IIS Manager using either **Start > Control Panel > Administrative Tools > Internet Information Services** or **Start > Administrative Tools > Internet Information Services**.
2. Select **Local Machine > Web Sites > Default Web Site**.
3. Right-click **Default Web Site** and select **New > Virtual Directory**.
4. When the **Virtual Directory Wizard** opens, click **Next**.
5. Complete the following steps to define the root virtual directory for **Report Portal**:
 - a. When prompted for an alias, type the name of the **Report Portal**, then click **Next**. For example, if you want to use "Portal" as the name of your **Report Portal**, assign the alias "Portal" to the virtual directory. Click **Next** when finished.
 - b. When prompted for the physical path, browse for and select the **<PortalName>|PortalASP** directory, then click **Next**.

Example: C:\Inetpub\wwwroot\Portal\PortalASP

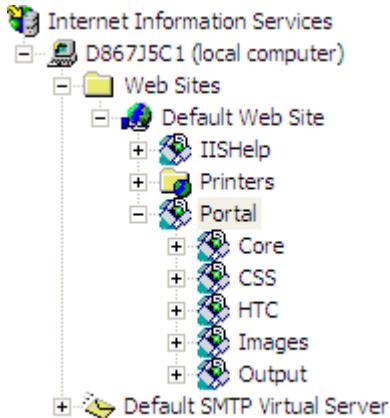
 - c. When prompted for permissions, verify that the following options are enabled:
 - **Read**
 - **Run scripts (such as ASP)**
 - d. Click **Next**, then click **Finish**. The virtual directory that you created appears under the Default Web Site as shown in the following example.



6. Right-click the virtual directory that you just created and select **New > Virtual Directory**.
7. Use the **Virtual Directory** wizard to create an alias for each of the following physical directories. Doing this creates an appropriately named virtual directory for each of these physical resources.

Create this alias . . .	For this physical resource . . .
Core	\PortalName\PortalFiles\Core Example: C:\Inetpub\wwwroot\Portal\PortalFiles\Core
CSS	\PortalName\PortalFiles\CSS Example: C:\Inetpub\wwwroot\Portal\PortalFiles\CSS
HTC	\PortalName\PortalFiles\HTC Example: C:\Inetpub\wwwroot\Portal\PortalFiles\HTC
Images	\PortalName\PortalFiles\Images Example: C:\Inetpub\wwwroot\Portal\PortalFiles\Images
Output	<p>Physical location of the directory in which Report Server saves the output for your report sets. The output folder can be located anywhere, can be named anything, and contains a subfolder for each report set.</p> <p>This must be the same directory that you specify in the Output Root parameter in the Report .cfg file for a report set. For more information, see Configuring Report.cfg Files on page 681.</p> <p>The default location is \PortalName\PortalFiles\Output.</p> <p>Example: C:\Inetpub\wwwroot\Portal\PortalFiles\Output</p> <p>The PortalName\PortalFiles\Output directory contains the profiles.xml file, which must be moved to the output directory that you specify for this alias.</p> <p>It is critical that the Path attribute is set properly.</p>

8. When you are finished, verify that IIS displays six new virtual directories. Make sure that the directory structure has one parent folder (with the same name as your portal) and five subfolders as shown below.



9. When finished, go to [Edit the Session Configuration File](#) on page 695 to continue with the installation process.

Edit the Session Configuration File

Navigation title: Edit the Session Configuration File

The **Report Portal** uses the information in a configuration file called `global.asa` to initialize user sessions.

When you install **Report Portal**, you must edit this file as indicated. The `global.asa` file resides in the `\PortalName\PortalASP\` folder.



Note: Do not change any other parameters in this configuration file.

1. On the machine where IIS is running, open the `global.asa` file in a text editor such as Notepad.
2. Optional. To let users access **Report Portal** without authentication, change the `Session("In")` parameter value to true. The default is false, which authenticates all users who attempt to access **Report Portal**.
3. If your **Report Portal** is installed on a drive other than the C drive, change the value of the `Session("Drive")` parameter to the correct drive location.
4. For the `Session("DBPath")` parameter, change the value to reflect the path to the database that contains the information needed to authenticate **Report Portal** users. Do not include the drive letter, but make sure to include a trailing slash.

Example: `/Inetpub/wwwroot/Portal/data/`

5. Save the file.
6. To verify that the **Report Portal** files have been installed correctly and can be reached through their designated virtual directory, open the following page in your browser:

`http://YourServerAddress/YourPortalName`

Example: `http://localhost/VisualReportPortal`

If the **Report Portal** ASPs have been installed correctly, you should see the portal login page. If you do not see this page, verify that ASPs are enabled on your IIS and double-check your virtual directory mappings.

Edit the Email.asp File

Navigation title: Edit the Email.asp File

Access to and permissions within your **Report Portal** are controlled using individual user and group accounts.

Each time that you add a new account or edit an existing account, a confirmation email can be sent to the email address that you specify for that account (see [Working with Accounts](#) on page 683) and copied to the email addresses that you specify in the `email.asp` file.



Note: Notification emails are sent to account users only when you have specified an email address for the account and properly configured the `email.asp` file. If you do not want notification emails sent for an account, leave the account's email field blank.

This file resides in the `\PortalName\PortalASP` folder.

1. On the machine where IIS is running, open the `email.asp` file in a text editor such as Notepad.
2. Set the following variables:

For this variable . . .	Provide this information . . .
<code>smtpserver</code>	DNS name or IP address of the SMTP server through which messages are sent. For example: <code>mail.hq.omniture.com</code>
<code>smtpserverport</code>	The port on which the SMTP server listens for connections. This is typically port 25.
<code>sendusing</code>	Indicates how the message is to be sent. Values are: 1 - Send messages using the locally installed SMTP service. Use this value if the SMTP service is installed on the computer where the script is running. 2 - Send messages using the SMTP service on the network. Use this value if the SMTP service is not installed on the computer where the script is running.
<code>smtpconnectiontimeout</code>	The amount of time that Report should wait for a response from the SMTP server before it times out the connection.

3. For the `NewUserEmail()` and `UpdateUserEmail()` functions, set the following variables:

For this variable . . .	Provide this information . . .
<code>From</code>	The text that you want to appear in the From header line in your confirmation emails. This value might be the same as the CC value.
<code>CC</code>	Optional. The valid email address of the person or alias who should receive a copy of all messages regarding new and changed user accounts. You can specify multiple email addresses by separating the addresses with commas (no spaces). For example: <code>admin@company.com, joemanager@company.com</code>  Note: The recipients receive copies of emails that contain user passwords.
<code>Subject</code>	The text that you want to appear in the Subject header line in your confirmation emails.
<code>WebPath</code>	The actual path to your portal. For example: <code>http://portal.omniture.com/Example</code>
<code>Body</code>	The text included in the automatically generated emails. For example, following is the default text included in the emails sent to provide login information: <ul style="list-style-type: none">• Your web portal login information is provided below:• UserName: username New Password: password• You may access the portal using the following URL: <code>http://WebPath</code>• After you log into the portal, you can change your password on the Admin tab.

4. Save and close the file.

Edit the Profiles.xml File

Navigation title:Edit the Profiles.xml File

To specify the profiles that you want to be available in **Report Portal**, you must configure the `profiles.xml` file.

The `profiles.xml` file resides in the folder you have designated for output. By default it resides in the `\PortalName\PortalFiles\Output` folder.

To add profile names to the profiles.xml file

1. On the machine where IIS is running, open the `profiles.xml` file in a text editor such as Notepad.
2. Add a profile element and tag for each **Profile** in your portal, as in the following example:

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<PROFILES>
    <PROFILE>
        <NAME>Product Sales</NAME>
    </PROFILE>
    <PROFILE>
        <NAME>Product Marketing</NAME>
    </PROFILE>
</PROFILES>
```

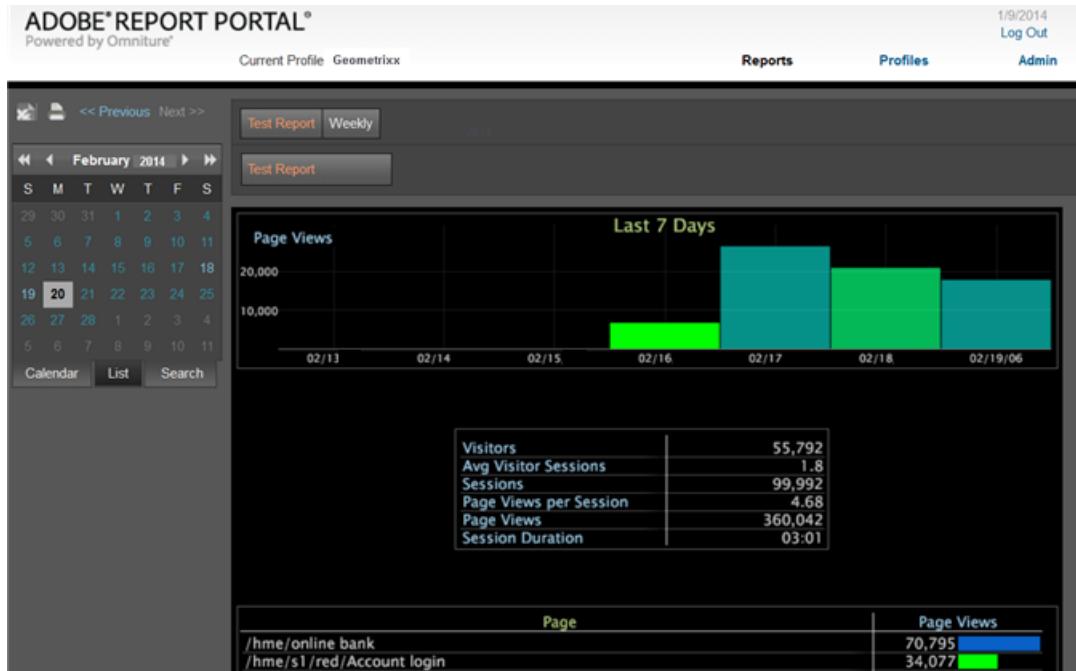
3. Save and close the file.

Customize the Report Portal User Interface

Navigation title:Customize the Report Portal User Interface

Report sets must be configured in a specific way to produce reports that display properly through **Report Portal**.

The user interface for **Report Portal** is designed to display a tab for each report set folder that appears in the output directory and is listed in the `profiles.xml` file, as well as the built-in **Admin** tab, which must be added to the `TopNavigation.xml` file to be displayed. For more information about displaying the built-in **Admin** tab, see [Linking an Output Folder to a Tab in the User Interface](#) on page 699.



- [Ensuring that Your Report Sets are Compatible with Report Portal](#) on page 698
- [Linking an Output Folder to a Tab in the User Interface](#) on page 699

Ensuring that Your Report Sets are Compatible with Report Portal

A report set defines a scheduled job for Report. It consists of two items:

- A folder that defines the collection of workspaces that you want Report to generate as reports.
- A configuration file (`Report.cfg`).

Among other things, the `Report.cfg` file tells Report when to generate the reports and where to save the output files. Report sets reside in the Reports folder on the data workbench server. A profile can display any number of report sets.

To ensure compatibility with **Report Portal**, your report sets must meet the following requirements:

- The output directory for your report sets must contain a configured `profiles.xml` file.
- Each report set must include a top-level report named “`ReportSetName Summary`,” where `ReportSetName` matches the name of the report set. For example, the following **Profile Manager** shows two report sets, “Home” and “Traffic.” Note that each report set defines a summary report (`Home Summary.vw` and `Traffic Summary.vw`, respectively).

File	Base	Traff	Value	Mark	Geog	IP Ge	Bank	User
profile.cfg							✓	
Color Legends\		✓	✓					
Context\	✓	✓	✓	✓	✓	✓	✓	✓
Dataset\	✓	✓		✓	✓	✓	✓	
Dimensions\	✓	✓	✓	✓	✓			✓
Export\	✓							✓
Filters\		✓						
Images\	✓							
Maps\					✓			✓
Menu\	✓	✓	✓		✓		✓	
Metrics\	✓	✓	✓				✓	✓
Reports\								
folder.lock	✓							
folder.useTabs	✓							
Home\								
Home Summary.vw						✓	✓	
Report.cfg						✓	✓	
Traffic\						✓	✓	
Report.cfg						✓	✓	
Traffic Summary.vw						✓	✓	

In **Report Portal**, the summary report appears on the report set's tab. The summary report can contain any workspace, window, or visualization you choose.

- The summary report must be the only report in the top-level folder for a report set. All other reports must be placed in subfolders. If you place other reports in the top-level folder, you can not view them through the portal.

Linking an Output Folder to a Tab in the User Interface

To specify the tabs that you want **Report Portal** to display, you must configure a `TopNavigation.xml` file for each profile. This file determines which report sets appear as tabs in the user interface for a particular profile, as well as the order of those tabs. The `TopNavigation.xml` file resides in the `\PortalName\PortalFiles\Core\TopNav\profileName` folder.

To edit the `TopNavigation.xml` file

- On the machine where IIS is running, open the `TopNavigation.xml` file in a text editor such as Notepad.
- Edit the list of `<TopNav>` elements so that it defines the names and order of the report sets whose output you want **Report Portal** to display, as in the following example:

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<TOPNAV_ELEMENTS>
<TOPNAV>
    <NAME>Monthly Web</NAME>
</TOPNAV>
<TOPNAV>
    <NAME>Weekly Web</NAME>
</TOPNAV>
<TOPNAV>
    <NAME>Admin</NAME>
</TOPNAV>
</TOPNAV_ELEMENTS>
```



Note: The **Admin** tab is a built-in tab that provides additional functionality. If you do not include it in the TopNavigation.xml file, this tab does not display and its functionality is not available.

3. In the `\PortalName\PortalFiles\Core\TopNav\` folder, create a folder for your next profile.
4. Copy the TopNavigation.xml file from the first profile folder and paste it into the new folder.
5. Edit the TopNavigation.xml as necessary, then save the file.
6. Repeat Steps 3-5 for all other profiles available in your portal.

Set Permissions for the Database

Navigation title: Set Permissions for the Database

To enable **Report Portal** to write to the database that contains information needed to authenticate users, you must set the proper permissions for the database.

1. On the machine where IIS is running, navigate to `\PortalName\data\users.mdb`.
2. Right-click the **users.mdb** file and select **Properties**.
3. On the **Security** tab, in Groups or user names, click **Users**.
4. In **Permission for User**, in the Write row, select **Allow**.
5. Click **OK** and close the **Properties** dialog box.

Test the Report Portal

Navigation title: Test the Report Portal

Steps to open the **Report Portal** and ensure that you can successfully view the reports for your profile.



Note: You can not log into the **Report Portal** until reports appear in the Output folder.

1. From your web browser, open the **Report Portal** using the following URI format:
`http://ServerAddress/PortalName`
Example: `http://localhost/VisualReportPortal`
2. If **Report Portal** prompts you for log-in credentials, enter an account name and password (for example, account “test” and password “user” for the default account).
3. When the **Report Portal** appears, verify that:
 - The portal contains a tab for each report set in the output folder.
 - Each tab displays the report set’s summary report.
 - The menu on each tab lists the individual subfolders (if any) for the report set and displays the contents of those folders.

Define Additional Accounts

Navigation title: Define Additional Accounts

Users must have a valid account and provide an account name and password when they access the **Report Portal**.

By default, user authentication is enabled in **Report Portal**.

The list of valid accounts for **Report Portal** is maintained in the database file, `portal.mdb`. **Report Portal** is installed with one account with administrative privileges:

- Account Name: test
- Password: user



Note: For security reasons, Adobe recommends that you change the password for this account after you install **Report Portal**.

To add user accounts to **Report Portal** or change information relating to existing accounts, you use the **Admin** tab on the **Report Portal** user interface.

Each time you add a new account or edit an existing account, a confirmation email is sent as specified in the `email.asp` file in the `\PortalName\PortalASP` folder. For more information, see [Edit the Email.asp File](#) on page 695.

For steps to add additional users, see [Working with Accounts](#) on page 683.



Note: Optionally, you can disable user authentication and allow anonymous access to **Report Portal**. For steps to do so, see the information about the Session("In") parameter in [Edit the Session Configuration File](#) on page 695.

Report Server Parameters Reference

Navigation title:Report Server Parameters Reference

Conceptual information about the parameters used in Report Server configuration files.

Report Server.cfg Parameters

Navigation title:Report Server.cfg Parameters

Information about Report Server.cfg parameters.

The sample Report Server.cfg shown in [Configuring the Connection to the Insight Server](#) on page 675 contains only the parameters included in the Report Server.cfg file by default.

If you contact the Adobe License Server through a proxy server, you need to add the Licensing vector and its parameters to the Report Server.cfg. Following is an example of this vector and its parameters that you can use a model for your Licensing vector:

```
Licensing = serverInfo:  
Proxy Address = string: ProxyIPAddress  
Proxy Password = string: ProxyPassword  
Proxy Port = int: ProxyPort  
Proxy User Name = string: ProxyUserName
```

The following table provides descriptions of the Report Server.cfg file parameters:

Parameter	Description
Excel Extension	Supported Excel extensions include: Excel Extension = string: .xlsx Excel Extension = string: .xls (this is default)
Fonts	Optional. Vector listing the fonts that Report Server should use to render UTF8-based unicode special characters. The number of fonts in the list is unlimited. The first font should always be Lucida Sans Console. If this parameter is not included in the Report Server.cfg file, Report Server uses only Lucida Sans console to display text. Report Server uses the first font in the list to render all characters until it encounters a character that it cannot render. It then uses the second font in the list to render that character. If that font does not render the character, Report Server uses the third font in the list to render that character, and so on, until it reaches the end of the font list. If the correct font is not listed in the vector, Report Server displays the hexadecimal value for the character.

Parameter	Description
	 Note: Do not make changes to this parameter while Report Server is running.
Gamma	Gamma setting for .png file output. The default value is 1.6.  Note: Adobe does not recommend changing this value.
Licensing	Optional. You need to modify the parameters in this vector only if you contact Adobe's license server through a proxy server. Section identifier for parameters you set to contact Adobe's license server through a proxy server.
Proxy Address	The address of a proxy server that Report Server must use to access Adobe's license server.
Proxy Password	Optional. The password associated with the Proxy User Name .
Proxy Port	The port of the proxy server.
Proxy User Name	Optional. The user name used to access the proxy server.
Network Location	The network location that Report Server uses to resolve the server's common name to an IP address. Network locations are defined in the address file located in the Addresses directory on the data workbench server machine.
Servers	Section identifier for parameters you set to configure which data workbench server machines Report Server must connect to generate reports. This includes a number indicating how many items are listed in this vector. For each server, add a serverInfo entry and complete the parameters as necessary.
Address	IP address of the data workbench server machine to which Report Server must connect to generate reports.
Name	The name that Report Server uses internally to identify the data workbench server. This is simply an internal label, so you can use any name you like. For consistency, we suggest that you use the common name as listed on the server's digital certificate.
Port	Port through which Report Server communicates with the data workbench server. For secure connections, this value is usually 443.
Proxy Address	The address of a proxy server that Report Server must use to access the data workbench server. This parameter is needed only when a proxy server is required to connect to your server machines.
Proxy Password	The password to the proxy server. This parameter is needed only when a proxy server is required to connect to your data workbench server machines.
Proxy Port	The port of the proxy server. The default value is 8080. This parameter is needed only when a proxy server is required to connect to your data workbench server machines.
Proxy User Name	The user name to the proxy server. This parameter is needed only when a proxy server is required to connect to your data workbench server machines.
SSL Client Certificate	Name of the SSL certificate file for the Report Server machine.
SSL Server Common Name	Server Common Name listed on data workbench server's digital certificate.
Use SSL	Indicates whether SSL is used for secure communication between the data workbench server and Report Server. The options are true or false. The default value is true.
Result Memory Limit (KB)	The amount of memory (in KB) that you want to make available for reports and alerts. The default value is 50000. When running reports, Report Server checks this value first, then checks the value in the Maximum Slice Size parameter. For example, if you set this parameter to 50,000 and the Maximum Slice Size to 50, Report Server runs only 50 workspaces at a time even if space is available to run more workspaces.

Parameter	Description
	 Note: This limit should never exceed the value set in the data workbench server's Query Memory Limit parameter, and, ideally, should be set a little lower than the Query Memory Limit to provide some leeway to other users who may be running reports at the same time.
Maximum Slice Size	The maximum number of report workspaces that Report Server can run at one time. The default value is 50. For more information about how Report Server uses this setting, see the Result Memory Limit (KB) parameter description.
Update Software	<p>Indicates whether to allow this Report Server's software to be updated by the data workbench server. The options are true or false. The default value is true.</p> <p>Following is an example of this parameter that you can use a model:</p> <pre>Update Software = bool: false</pre>
Use OpenGL Hardware Rendering	<p>Controls whether Report Server uses hardware rendering (such as the machine's graphics card) to produce report output. The options are true or false. The default value is true.</p> <p>This parameter should be set to false only when you are experiencing problems with your graphics card. When set to false, Report Server does not attempt to use hardware rendering and uses software rendering by default.</p>
Reporting	Section identifier for parameters you set to configure reporting.
Completion Message Interval	<p>The frequency (in seconds) with which Report Server prints completion status messages when queries are being run during report or alert generation. The default value is 120 seconds.</p> <p>Example: Workspace queries are 62.145672% complete.</p> <p>Completion messages are written to the <code>reportserver.log</code> and are synched to the server. This setting controls the <code>status.txt</code> files that are sent back and forth for each report set, so that the percent complete can be shown with thumbnails. The messages are sent whenever a report set completes or when the interval is reached, whichever comes first. Setting this higher does not effect the rate at which you see report generated in the client interface by the thumbnails, but does effect how many interim messages you see. Specifying a low value can cause the system to spend a large amount of time synchronizing data, because data is synched from Report Server server to the profile, across all the DPUs and to all connected clients each time a <code>status.txt</code> message changes.</p> <p>The system always sends a <code>status.txt</code> file when a report set completes, regardless of the setting of this configuration parameter.</p>
Profiles	<p>Number indicating how many items are listed in this vector. For each profile for which reports are to be created, add a ReportProfile entry in the Profiles vector and complete the Server and Profile parameters.</p> <p>Server - The name which Report Server uses internally to identify the data workbench server. This name must be the server common name listed on data workbench server's SSL certificate.</p> <p>Profile - Name of the profile for which reports are to be created. This name must match the named profile on the data workbench server machine.</p>
SMTP Server for Errors	<p>The address of the SMTP server from which you want to send Report Server errors via email.</p> <p>Example: <code>mail.mycompany.com</code></p> <p>An SMTP server is required for use of the described capabilities.</p>
SMTP Server for Errors Password	<p>The password for logging in to SMTP server. This parameter is optional unless login is required to send mail.</p> <p>An SMTP server is required for use of the described capabilities.</p>
SMTP Server for Errors Send From	The email address from which you want to send Report Server errors.
SMTP Server for Errors Send To	<p>The email address(es) to which alerts are sent.</p> <p>Example: <code>adm1@company.com, adm2@company.com</code></p>

Parameter	Description
	An SMTP server is required for use of the described capabilities.
SMTP Server for Errors Username	The user name for logging in to the SMTP server. This parameter is optional unless login is required to send mail. An SMTP server is required for use of the described capabilities.
Status Interval	The frequency (in seconds) with which Report Server generates and sends status information to data workbench server to be displayed in Detailed Status . The default value is 120 seconds. It is not recommended to set this to a small value, such as two minutes, because a reporting queue can take hours to run. In that case, you might consider a setting of 600 to 1200 seconds. For more information about Detailed Status , see the Administrative Interfaces chapter of the <i>Insight User Guide</i> .
Update Interval	The frequency (in minutes) with which Report Server monitors all of the profiles listed in the Profiles vector for new reports and alerts. The default value is 10 minutes. The time that you specify is divided among all of the profiles listed. For example, if your interval is set to 10 minutes and you are monitoring two profiles, each profile is monitored for 5 minutes. If a profile is being monitored when a new or modified report or alert is saved to the profile, the report or alert is immediately available for generation. If the Update Interval is configured to monitor more than one profile, it is important that this setting is high enough to load all profiles within the configured time. In systems with many large dimensions configured, for example, where it might take several minutes to retrieve the initial data connection with all the element names, this setting must be long enough for that full synchronization to occur. Otherwise, the system issues a profile sync error.

Report.cfg Parameters

Navigation title:Report.cfg Parameters

Information about Report .cfg parameters.

The sample Report .cfg shown in [Configure the Report Set](#) on page 662 contains only the parameters included in the Report .cfg file by default. The following table provides descriptions of all of the available Report .cfg file parameters.

If you need to add additional parameters to a Report .cfg file, you must do so using a text editor. For steps to do so, including examples of how to define each parameter entry, see [Editing Existing Report.cfg Files](#) on page 667.



Note: The parameters in this table are listed in alphabetical order. When you open the Report .cfg file in Data Workbench, vectors are listed in alphabetical order, followed by individual parameters listed in alphabetical order.

Parameter	Description
Alert Threshold	<i>Optional.</i> This parameter applies only to reports with metric indicators. Number of metric indicators that must appear in the worksheet before an alerting report is sent. If only one metric is being monitored in the metric indicator worksheet, set the threshold to 1. The report is generated when the metric in the sheet evaluates to an up/down arrow or an X. If more than one metric is being monitored in the report, you can select the number of metric indicators that must evaluate to an up/down arrow or an X before the report is generated. For example, if two metrics are being monitored: <ul style="list-style-type: none">• If the threshold is set to 1, the report is generated if either of the metrics in the sheet evaluate to an up/down arrow or an X.• If the threshold is set to 2, both of the metrics must evaluate to an up/down arrow or an X before the report is generated. For more information about metric indicators, see the <i>Data Workbench User Guide</i> .

Parameter	Description
Allow Report Regeneration	<p>Indicates whether Report Server automatically generates or regenerates particular reports when you create or modify those reports. The options are true or false. If set to true, creating or modifying a report workspace causes Report Server to regenerate that report for the most recent run.</p> <p> Note: Changing the <code>Report.cfg</code> file causes Report Server to regenerate all reports controlled by that <code>Report.cfg</code> file.</p>
Attachments	<p><i>Optional.</i> Section identifier for the name and content type of any attachments that go out with reports distributed via email, including the number of attachments.</p> <p>To add a new attachment:</p> <ol style="list-style-type: none"> 1. Open the <code>Report.cfg</code> file in Data Workbench. 2. Right-click Attachments and click Add new child > Attachment.
Content Type	<p>Content type of the file to be attached.</p> <p>Example: image/jpeg</p>
FileName	<p>Location and name of the file to be attached.</p> <p>Example: <code>c:\myimage.jpg</code></p>
Color Set	<p>Identifies the color scheme to be used for <code>.png</code> files. 0 is for a black background; 1 is for a white background; and 2 is for a grayscale image.</p>
Command To Execute	<p><i>Optional.</i> A batch command or executable that runs after the report set is generated. If launch of the command shell interpreter is required, precede the command with cmd /c.</p>
Default Excel Template	<p><i>Optional.</i> File name of the generic Excel template file (<code>.xls</code> or <code>.xlsx</code>) that you want to use when generating reports as Excel files. This parameter supports full file paths, such as <code>c:\templates\mytemplate.xls</code>.</p> <p>This file is used for all Excel reports unless a template has been defined specifically for a particular report. See Using a Template File on page 664.</p>
Dimension Name	<p><i>Optional.</i> Name of the dimension for which you want to dynamically generate a report. If you enter a dimension name in this parameter, you must enter a value in either the Lookup File parameter or the Top N Metric and Top N Value parameters. The dimension named in this parameter must exist in the dataset for which reports are being created.</p>
Email Only If Perfect	<p><i>Optional.</i> Lets the user specify that a report set should be sent out only when no errors occurred during the run. The options are true and false. The default value is false.</p>
End Date	<p><i>Optional.</i> The last date and time that you want the report set to run. This time is based on the As Of time of the dataset.</p> <p>Format: MM/DD/YYYY hh:mm time zone, using the 24-hour syntax for time</p> <p>Example: 08/01/2007 12:01 EDT</p> <p>For more information about time zone settings, see the <i>Dataset Configuration Guide</i>.</p>
Every	<p>Frequency of the report set generation: day, week, or month.</p>
Excel Watchdog Timeout (seconds)	<p><i>Optional.</i> The number of seconds that you want Report Server to wait for Microsoft Excel to respond when generating a report as an Excel file before Report Server decides that Excel is not responding and terminates the process. Using this parameter enables Report Server to terminate Excel when it becomes unresponsive and continue processing your non-Excel reports. The default is 300.0. To disable this functionality, set this parameter to 0.0.</p> <p>Make sure that the value you define is long enough to allow the report to be exported to Excel. Otherwise, Report Server may prematurely terminate Excel and your report will not generate.</p>

Parameter	Description
Filter Formula	<p><i>Optional.</i> Filter that is applied to every workspace in the report set.</p> <p>For more information, see the syntax for creating filters.</p>
Gamma Correction	<p>Gamma setting for .png file output. The default is 1.6.</p> <p> Note: Adobe recommends that you do not change this value.</p>
Hide Logos	<p>Indicates whether Report Server hides the logos when generating your reports. The options are <code>true</code> or <code>false</code>. If set to <code>false</code>, your report is generated with the Report logo. The default is <code>false</code>.</p>
Lookup File	<p><i>Optional.</i> When this parameter is populated, Report Server runs in dynamic mode and generates reports for each element of the dimension specified in the Dimension Name parameter. This file must contain two tab-delimited columns, without a header row.</p> <ul style="list-style-type: none"> Column 1 contains a list of dimension elements. Column 2 contains the email addresses of the report recipients. A report for a given element in column 1 is sent to the email address on the same row in column 2. You can enter multiple email addresses by separating them with commas (no spaces). If reports are not to be emailed, this column can be empty but must exist. <p> Note: If you enter a value in this parameter, you must enter a value in the Dimension Name parameter.</p>
Notification Only	<p>This Report Server setting allows you to configure data workbench to send an email when a report is generated. Setting this value to <code>true</code> does not send out the report, but rather sends an email notifying the subscribed user that the report has been generated.</p>
Mail Report	<p>Section identifier for distributing reports via email. To distribute reports via email, complete the following parameters for the Mail Report entry. All reports in the report set are emailed in one message to the email addresses specified in the Recipients parameter.</p> <p> Note: Report Server sends an email only when it has generated at least one report.</p> <p>To enable the emailing of reports, you must complete at least the following parameters for this entry:</p> <ul style="list-style-type: none"> • SMTP Server • Recipients • Sender Address • Notification Only <p> Note: To send out reports by email after re-generating a report set, see Editing Existing Report.cfg Files on page 667.</p> <p>The Notification Only value is available in 5.4x and 5.5x releases.</p> <p>For a large set of recipients to be notified (more than 20), it is highly recommended that you use email distribution lists.</p>
Body XSL Template	<p><i>Optional.</i> Path of the XSL template file to be applied to the <code>reports.xml</code> file. Using this parameter enables Report Server to send your reports within the distributed email instead of as attachments. The resulting text is used as the body of the email message.</p> <p>See Report Sample Files on page 708 for a sample file.</p> <p>For information about Extensible Stylesheet Language (XSLT), see The Extensible Stylesheet Language Family.</p>
Recipients	Email addresses of the people to whom you want to send the report.
Sender Address	Email address of the sender.

Parameter	Description
Sender Name	Optional. Name of the sender.
SMTP Server	Address of the SMTP server machine and the password and user name required for authentication.
Subject	<i>Optional.</i> Subject line describing the email to be sent.
Notification Only	Lets you configure data workbench to send an email when a background report is generated. Setting this value to True does not send out the report, but rather sends an email linking the subscribed user to the report location.
Output Root	<p><i>Optional.</i> Output location of the generated report sets. The default is the <profile name>\Reports folder within the Report Server installation directory.</p> <p>To configure Report Server to output reports to a portal, set the Output Root to the document root of the web server used for the portal.</p>
Preload Query Filter	<p><i>Optional.</i> This parameter applies only to the Top Dimension Element report type.</p> <p>The name of the filter that you want to apply to the query that must be run to determine the top N dimension elements before the report can be generated. The default is Broken_Session_Filter. For more information about the Broken Session Filter, see the <i>Data Workbench User Guide</i>.</p>
Report Types	<p>Format(s) in which you want to generate your output. You can use any or all of the following options to output the report set in multiple formats at one time:</p> <ul style="list-style-type: none"> • Excel creates an Excel workbook with one visualization per worksheet. As a general rule, use Excel files for email distribution. See Generating Reports as Microsoft Excel Files on page 664. For information about using a template file, see Using a Template File on page 664. • png creates Portable Network Graphic files. As a general rule, use .png files for display in a web browser (portal). • Thumbnail creates a thumbnail (.jpg file) of the workspace. The default size is 240x180. To change the default size, edit the Thumbnail X and Thumbnail Y parameters. <p>To add a new report type when editing Report.cfg in data workbench, right-click Report Types, click Add new child, and select the desired report type.</p>
Start Date	<p>The first date and time that you want the report set to run. This time is based on the As Of time of the dataset. Format: MM/DD/YYYY hh:mm time zone, using the 24-hour syntax for time.</p> <p>For more information about time zone settings, see the <i>Dataset Configuration Guide</i>.</p> <p> Note: The reports start to run when the timestamps of the data in the profile match the date and time specified.</p> <p>Example:</p> <p>If the start date is 08/08/2006 12:00 EST, reports run for data with a timestamp of 08/08/2006 12:00 EST and later.</p> <ul style="list-style-type: none"> • Daily reports will run for 08/08/2006 and each day thereafter for data with hh:mm = 12:00 EST. • Weekly reports will run for 08/08/2006 and for every 7th day thereafter for data with hh:mm = 12:00 EST. • Monthly reports will run for 08/08/2006 and for the 8th day of each month thereafter for data with hh:mm = 12:00 EST. <p>The Report Time metric affects “Last N” reporting dimensions, such as “Last 7 Days,” “Yesterday,” and “3 weeks ago.” For queries in Report Server, the Report Time metric (Report Time.metric) identifies the date and time for which the reports are being run. This is initially the date and time specified in the Start Date parameter, which then increments by the period specified by the Every parameter. For queries in data workbench, the Report Time metric is based on midnight of the As Of metric (As Of.metric). Because of the difference in the definitions of the Report Time metric, if you query a workspace that uses a Last N dimension, you can receive different results in data workbench and Report Server for the same workspace.</p>
Thumbnail X	<i>Optional.</i> Integer controlling the size (in pixels) of the X axis of thumbnails generated as output.
Thumbnail Y	<i>Optional.</i> Integer controlling the size (in pixels) of the Y axis of thumbnails generated as output.

Parameter	Description
Top N Metric	<p><i>Optional.</i> See the description for the Top N Value parameter.</p>  Note: If you enter a value in this parameter, you must enter a value in the Dimension Name parameter and the Top N Value parameter.
Top N Value	<p><i>Optional.</i> When this parameter is populated, Report Server runs in dynamic mode and generates reports for the top number (specified in this parameter) of elements for the dimension specified in the Dimension Name parameter, counting by the metric specified in the Top N Metric parameter.</p> <p>Example: If you enter Page in the Dimension Name parameter, Sessions in the Top N Metric parameter, and 5 in this parameter, the generated report lists the five top pages with the highest number of sessions.</p>  Note: If you enter a value in this parameter, you must enter a value in the Dimension Name parameter and the Top N Metric parameter.
Use Local Sample Only	<p><i>Optional.</i> Indicates whether you want Report Server to generate reports using only the local sample of the dataset. Setting this parameter to true enables you to view a sample of the report set (without placing a load on a data workbench server) to see how the output looks without taking all the time needed to fully process the data. This operates as a test function. The options are true or false. The default is false.</p>
Workspace Path	<p><i>Optional.</i> Location of a collection of workspaces for a given report set. This is useful for maintaining a single copy of workspaces that need to be generated and distributed multiple ways, using <code>Report.cfg</code> files for multiple report sets. The root directory for this path can be any profile folder. Do not enter a slash (\) at the start of the path string.</p> <p>Example: You can save the common workspaces for Set A and Set B in the <code>Reports\Common</code> folder, then define the <code>Report.cfg</code> files for two different report sets, each with different generation and distribution settings. In both <code>Report.cfg</code> files, you would set the Workspace Path parameter to <code>profile name\Reports\Common</code>.</p>
XSL Output File	<p><i>Optional.</i> Path of the output file that is created when the XSL Template is applied to the report index.</p>
XSL Template	<p><i>Optional.</i> Path of the XSL template file to be applied to the report index. The resulting transformed <code>.xml</code> is written to the specified XSL Output File. See Report Sample Files on page 708 for a sample file.</p>  Note: Unless you use an <code>.xsl</code> template when generating your reports, all reports are distributed by email as attachments.

Report Sample Files

Navigation title:Report Sample Files

Conceptual information about sample templates for formatting reports.

Sample Report Index

Navigation title:Sample Report Index

Code sample of the Report index.

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<REPORTS>
  <REPORT>
    <NAME>Last 6 Months</NAME>
    <PATH>By Period\Last 6 Months .png</PATH>
    <WEB_PATH>By Period/Last 6 Months .png</WEB_PATH>
  </REPORT>
  <REPORT>
    <NAME>Last 7 Days</NAME>
    <PATH>By Period\Last 7 Days.png</PATH>
```

```

<WEB_PATH>By Period/Last 7 Days.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>Last 8 Weeks</NAME>
    <PATH>By Period\Last 8 Weeks.png</PATH>
    <WEB_PATH>By Period/Last 8 Weeks.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>Last Week by Day</NAME>
    <PATH>By Period\Last Week by Day.png</PATH>
    <WEB_PATH>By Period/Last Week by Day.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>This Week by Day</NAME>
    <PATH>By Period\This Week by Day.png</PATH>
    <WEB_PATH>By Period/This Week by Day.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>Page Summary</NAME>
    <PATH>Page\Page Summary.png</PATH>
    <WEB_PATH>Page/Page Summary.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>Pages to Date</NAME>
    <PATH>Page\Pages to Date.png</PATH>
    <WEB_PATH>Page/Pages to Date.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>To Date by Weekday</NAME>
    <PATH>Page\To Date by Weekday.png</PATH>
    <WEB_PATH>Page/To Date by Weekday.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>Referrer Summary</NAME>
    <PATH>Referrer\Referrer Summary.png</PATH>
    <WEB_PATH>Referrer/Referrer Summary.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>Referrers to Date</NAME>
    <PATH>Referrer\Referrers to Date.png</PATH>
    <WEB_PATH>Referrer/Referrers to Date.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>To Date by Weekday</NAME>
    <PATH>Referrer\To Date by Weekday.png</PATH>
    <WEB_PATH>Referrer/To Date by Weekday.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>Session Summary</NAME>
    <PATH>Session\Session Summary.png</PATH>
    <WEB_PATH>Session/Session Summary.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>To Date by Weekday</NAME>
    <PATH>Session\To Date by Weekday.png</PATH>
    <WEB_PATH>Session/To Date by Weekday.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>To Date</NAME>
    <PATH>Session\To Date.png</PATH>
    <WEB_PATH>Session/To Date.png</WEB_PATH>

```

```

</REPORT>
<REPORT>
    <NAME>To Date by Weekday</NAME>
    <PATH>Visitor\To Date by Weekday.png</PATH>
    <WEB_PATH>Visitor/To Date by Weekday.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>To Date</NAME>
    <PATH>Visitor\To Date.png</PATH>
    <WEB_PATH>Visitor/To Date.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>Visitor Summary</NAME>
    <PATH>Visitor\Visitor Summary.png</PATH>
    <WEB_PATH>Visitor/Visitor Summary.png</WEB_PATH>
</REPORT>
<REPORT>
    <NAME>Traffic Summary</NAME>
    <PATH>Traffic Summary.png</PATH>
    <WEB_PATH>Traffic Summary.png</WEB_PATH>
</REPORT>
</REPORTS>

```

Sample XSL Style Sheet

Navigation title: Sample XSL Style Sheet

Code sample of XSL style sheet.

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet version="1.0"
xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<xsl:template match="/">
    <html>
        <body>
            <h2>Reports</h2>
            <xsl:for-each select="REPORTS/REPORT">
                <a>
                    <xsl:attribute name="href">
                        <xsl:value-of select="PATH" />
                    </xsl:attribute>
                    <xsl:value-of select="NAME" />
                </a><p/>
            </xsl:for-each>
        </body>
    </html>
</xsl:template>
</xsl:stylesheet>

```

Sample Mail XSL Style Sheet

Navigation title: Sample Mail XSL Style Sheet

Code sample of Mail XSL style sheet.

```

<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet version="1.0"
xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
<xsl:template match="/">
    <html>
        <body>
            <h2>Reports</h2>
            <xsl:for-each select="REPORTS/REPORT[@format != 'xls']">

```

```
<img><xsl:attribute name="src">cid:<xsl:value-of  
select="PATH" /></xsl:attribute></img><p/>  
</xsl:for-each>  
</body>  
</html>  
</xsl:template>  
</xsl:stylesheet>
```

Data Workbench Page Tagging

Conceptual information about page tagging, including how to acquire data and the types of date you can acquire.

What Kind of Data Can I Acquire?

Navigation title:What Kind of Data Can I Acquire?

Sensor enables you to acquire web request data (event or log data) as well as extended measurement data.

Extended measurement data is not available to your web servers as a part of their normal operation.

The following topics are described:

Web Request Data

Sensor enables web request data (event or log data) to be acquired and transported automatically to a central location for storage and processing for analysis. Unless you specifically choose to filter out certain types of requests and not collect data about those request types, **Sensor** captures data about all GET requests made of the web servers on which it is installed.

Sensor automates this data acquisition process for all GET requests that are made on your servers and has significant business and technical benefits over alternative methods of acquiring website request data. These benefits include the following:

- Requests that are unnecessary for analysis and reporting can be filtered out before you incur costs for their acquisition, transportation, storage, and processing.
- Site administrators do not have to rotate log files in batch, either manually or via script.
- **Sensor** aggregates log files at a central location to allow easy access for processing.
- **Sensor** organizes and stores log files in a common data-preserving format, removing the need to preprocess them before they can be used for analysis and reporting purposes.
- Instances of certain content types can be included in the log files even though most requests for a certain content type are automatically filtered out.
- **Sensor** compresses log file entries, which requires significantly less storage space, reducing costs and allowing the data to be kept available for analysis for longer periods of time.
- **Sensor's** fault tolerant features allow system and network faults while still ensuring the delivery of the log data to a central repository.
- **Sensor** allows the implementation of controlled experiments with web content, processes, and marketing campaigns.
- **Sensor** time stamps log entries in 100ns units, allowing new types of analytic functionality.
- **Sensor** allows site owners to add data (measurements) to the log entries after initial implementation for consideration in analysis and reporting.

For more information about acquiring extended measurement data, see [Acquiring Baseline Measurements](#) on page 715.

Extended Measurement Data

Sensor also supports the use of page tags (or embedded object requests) to acquire measurement data that is not available to your web servers as a part of their normal operation. Page tags are commonly used to measure:

- The viewing of a logical page in a dynamic website.
- The viewing of content or ads on a third-party controlled website.
- The viewing of content that is served from a browser cache or CDN.
- Detailed information about a visitor's browser, including measurements such as page load time, screen resolution, what form fields the visitor has filled in, and so on.
- Other data that is not otherwise sent by browsers to your web servers.

Sensor collects any information placed in any GET request made to a web server that is running **Sensor**. Such requests may come from embedded object requests of any sort, either to simply measure that the request was made at a certain time by a certain browser or to pass other measurement data into the data collection stream so that it may be processed for analysis and reporting purposes.

Sensor provides the best of both client-side and server-side data acquisition worlds—it acquires your server-side web log data and collects client-side, third-party site, or cache-busting measurements taken by embedded object requests. In other words, **Sensor** acquires both the request data normally known to your web servers (server-side measurements) and any additional measurement data that you collect through the use of page tags (client-side measurements) that send their data to any web servers running **Sensor**. Such web servers can be dedicated to collecting client-side measurements but are not required to be.

For more information about acquiring extended measurement data, see [Acquiring Extended Measurements](#) on page 721.

How Do I Acquire this Data

Navigation title:How Do I Acquire this Data?

You must install and run **Sensor** on each web server that serves the content for your site to collect all of the requests that are seen by those servers.

These requests make up 90% or more of the requests made to your site and 90% or more of the data that is needed for the complete analysis of your site's traffic. **Page Tags** should then be used to collect the remaining 10% or less of the traffic data that is not known to your web servers. The following, however, are valid configurations for the collection of web request data from your site, in order of preference, based on our operational experience:

1. **Sensor** is installed on each web server that you control and that supports your site. Content from third-party sites, content served from cache, and certain types of dynamic content should be tagged, and such page tags should send the data that they collect to a web server at your location that is running **Sensor**. You may add an additional web server if the level of page tag request traffic justifies such, or in special cases, dedicate a web server to collect these page tag requests.
2. **Sensor** is installed on two web servers, also referred to as data collection servers in this guide, at your location that are dedicated to collecting page tag request data from tagged pages. All content on your site is tagged and all page tags are directed to the two data collection servers.
3. **Sensor**'s data collection services are provided by an outsourcer that runs data collection servers to collect all of your web request data. In this case, all content on your site is tagged and the page tags send their data to the outsourced data collection servers.

For more information about **Sensor**, see the *Data Workbench Sensor Guide*.

Implementing Reference Page Tags

Navigation title:Implementing Reference Page Tags

Information about editing and adding Reference Page Tag Execution Scripts and Page Tag Execution Calls, and understanding the v1st Cookie.

Editing the Reference Page Tag Execution Script

Navigation title: Editing the Reference Page Tag Execution Script

The **Reference Page Tag** consists of a **Page Tag Execution Script** that resides on a web server, and when called results in the collection of all client-side data for the page requested by the site visitor.

You can modify the **Reference Page Tag Execution Script** to collect additional information that may be identified during requirements gathering meetings with the Adobe Consulting Services team. The **Reference Page Tag Execution Script** is relatively small in size to avoid large download additions to your web pages.

The following **Reference Page Tag Execution Script** code is provided to you in a file named `zig.js`:

```
//REFERENCE PAGE TAG
// CONSTANTS
var ct = "<img src=\"";
var cd = "[PATH_TO_WEB_SERVER]"; //this should contain the domain of
                                //the web site that will host the
                                //page tag

var cu = "[PATH_TO_WEB_PAGE_TAG_CODE]/zag.gif?Log=1";
                                //this should contain the full path to
                                //the zag.gif file (excluding domain)
                                //and include the query string of log=1

var ce = ">";
var c = {};
c["sw"] = screen.width;
c["sh"] = screen.height;
c["cd"] = screen.colorDepth;
var co = "";

for ( cKey in c ) {
co = co+"&"+cKey+"="+escape(c[cKey]);
}
document.write(ct,cd,cu,co,ce);

var d = {};
d["dt"] = document.title;
d["dr"] = document.referrer;
d["cb"] = new Date().getTime();
var vo = "";

if (typeof v != "undefined") {
for ( vKey in v ) {
vo = vo+"&"+vKey+"="+escape(v[vKey]);
}
}
for ( dKey in d ) {
vo = vo+"&"+dKey+"="+escape(d[dKey]);
}
document.write(ct,cd,cu,vo,ce);
//END REFERENCE PAGE TAG
```

To facilitate data collection through the use of the **Reference Page Tag**, complete the following steps:

1. Create or place the 1 pixel by 1 pixel image file named `zag.gif` into a directory present on your web server.

-
2. Modify the cd variable to reference the appropriate domain of your website or Adobe managed services domain from which the `zag.gif` file is referenced. The reference to the file is created through the execution of the `zig.js` file functions. For example:

```
/www.mysite.com
```

3. Modify the cu variable to reference the appropriate path to the location of the `zag.gif` file. For example

```
/scripts
```

4. Ensure appropriate cache-control headers are established for the `zag.gif` and `zig.js` files.

Adding Reference Page Tag Execution Calls

Navigation title: Adding Reference Page Tag Execution Calls

The **Reference Page Tag Execution Call** is inserted into web pages for which you want to collect measurement data.

It should be included in the body of the HTML document and can be placed within a global include footer if applicable. The **Reference Page Tag Execution Call** can be modified by your team to collect additional information that might be identified during requirements gathering meetings with the Adobe Consulting Services team.

To facilitate data collection through the use of the **Reference Page Tag**, complete the following steps:

1. Copy the following code into your HTML document body:

```
<!--//BEGIN REFERENCE PAGE TAG-->
<script language="javascript">
var vlc = "0" //Capture Link Click 1=TRUE, 0=FALSE
var v = {};
</script>

<!--//MODIFIY PATH TO ZIG.JS-->
<script language="javascript" src="/path/to/zig.js"
type="text/javascript"></script>
<!--//END REFERENCE PAGE TAG-->

<noscript>

</noscript>
<!-- END REFERENCE PAGE TAG-->
```

2. Modify the path to the location of the `zig.js` and `zag.gif` files. For example:

```
/www.mysite.com/scripts/zig.js
/www.mysite.com/images/zag.gif
```

Please ensure that the appropriate HTTP Cache-Control headers have been set on your web server to ensure that the `zig.js` and `zag.gif` files are not cached by the browser. You can set the HTTP Cache-Control header information using one of two methods. The first method is to set an HTTP header via the web server. The second method is to set an HTTP header for each specific page or embedded object using script. With the scripting method, the web page must have been created using a programming language such as JSP or ASP. The page then is scripted so that it sends the appropriate header information. Two obvious drawbacks accompany this method: 1) all pages must be coded to send the header, and 2) the pages cannot be static HTML, which has some effect on web server performance.

Web sites running on Microsoft IIS can add the appropriate HTTP header through the Microsoft Management Console. Websites served from Netscape iPlanet Web Servers can accomplish this by editing the `obj.conf` file within the site's configuration directory. The Apache Web Server provides webmasters the ability to customize HTTP headers using the included `mod_headers` module where AOLServer becomes customizable through the

use of Tcl modules. Before implementing HTTP Cache-Control headers, you should refer to the documentation specific to your web server platform.

In general, the HTTP header should be structured as follows:

```
Cache-Control: no-cache  
Pragma: no-cache  
Expires: -1
```

Understanding the v1st Cookie

Navigation title:Understanding the v1st Cookie

Site uses cookies to uniquely identify visitors to your website and track their behavior over time.

The first time that a particular browser (considered a visitor) makes a request of your website, **Sensor** works with your web server to set a persistent cookie, cs(cookie)(v1st), which is interpreted internally within the system as x-trackingid. This persistent cookie is set in addition to any other cookie that your site may otherwise set. This cookie optimizes your ability to track your visitors over multiple sessions, which enables many types of analysis that are otherwise impossible.

Sensor assigns a 64-bit numeric key in the cookie to identify new visitors that make a request of the site as a unique identifier. When the **Sensor** sees a request from a browser, it checks to see whether “cs(cookie)(v1st)”, a first-party persistent cookie set by **Sensor**, exists in the request data. If the cs(cookie)(v1st) is not present, **Sensor**, through your web server, tells the browser to set it. Unlike other solutions, **Sensor** is able to set this cookie on the visitor’s first request.

Below is the standard persistent cookie header sent to the browser on its first request of your site by your web server, at the direction of **Sensor**. The format can be defined at the time of configuration if a different name or expiration date is desired. For example:

```
Set-Cookie:v1st=3D80DCA944D60E16; path=/; expires=Wed, 19 Feb 2020 14:28:00  
GMT
```

This cookie is set just once, on the very first request made to your site by that visitor. It then is collected from that visitor each time that browser makes a request (either page or embedded object request) of your site in the future. The cookie is very small in size to minimize the amount of bandwidth used to transmit it to your servers with each request from that browser to your site.

Accepting a persistent cookie is at the browser’s discretion. Most web users understand what cookies do and also recognize that first-party cookies provide a valuable benefit to them in allowing site content to be customized to their preferences. These first-party cookies are not blocked by the default security settings of the popular browsers. If a user does choose to block first-party cookies, their page view requests are still logged, but the measurement data from those requests cannot be reliably correlated to a particular visitor or their sessions on the site. Many sites, especially sophisticated dynamic sites, already use first-party cookies, which are in many cases necessary to enable web applications to operate for the visitor. A step back from a persistent cookie is a session cookie, which allows a series of requests to be knit together into a session, but does not allow inter-session visitor tracking. **Site** is capable of sessionizing visitor data based on session cookies or by IP number, but both methods significantly detract from the types and value of analysis that can be conducted with **Site** or any other web activity analysis and reporting system.

Acquiring Baseline Measurements

Navigation title:Acquiring Baseline Measurements

Conceptual information about acquiring page request data, embedded object requests, and measurement data through cookies.

Acquiring Page Request Data

Navigation title:Acquiring Page Request Data

Sensor acquires all measurement data that is carried on the page requests (GET requests) made to the web servers on which it has been installed.

Sensor acquires this measurement data through the web server's application programming interface, directly from the instance or instances of the web server software running on your web server. **Sensor** does not access the web server generated log files. In fact, after **Sensor** and the data workbench server have been installed and tested, the web server's native logging feature can be disabled without affecting data collection. In many cases, disabling the logging of files to the local disks of the web server machines themselves improves the page serving capacity of those web servers because of the relatively large amount of fixed disk I/O required to log this information to the local disk of the web server machine.

Sensor collects measurement and web request data directly from each web server process and virtual web server process (if applicable) and temporarily writes the data to a Queue File, a fault-tolerant memory queue with fixed disk backing, on the web server machine. The Sensor Transmitter service (or daemon depending upon the platform) retrieves data from the Queue File and then compresses and encrypts it before transmitting it to the data workbench server for long-term storage. With **Sensor**, data is accumulated on your web server machines in the Queue File only if you are having a network or other problem that prevents its transmission. The Queue File allows for the efficient local storage of hours to days of web request data to protect the data if a network or system fault does not allow the data to be transmitted to the data workbench server in real time.

Sensor collects measurement data from each physical and logical web server process, filters it by content type, compresses it, encrypts it, and streams it to the data workbench server.

The following table contains the fields of log information that are acquired by **Sensor** for each GET request that is not filtered out based on **Sensor**'s configuration file:

W3C Name	Data Collected	Explanation	Explanation
x-trackingid	Tracking Identifier (unique visitor)	Identifier read from a cookie placed in the user's browser by Sensor on Visitor's initial request	V1st=3C94007B4E01F9C2
Date	Timestamp	Time at which request was processed by server (at 100ns precision; accuracy depends on server environment and NTP)	2002-11-21 17:21:45.123
Time			
sc(content-Type)	Content Type	Type of object returned from server	text/html
sc-status	HTTP Response Status Code	Numerical code generated by the server that notes the status of the HTTP server's response	404
cs-uri-stem	URI Stem	The stem portion of the URI requested by the client	pagedir/page.asp
c-ip	Client IP	IP Address of the requesting client	127.0.0.1
s-dns	Server Domain Name	Domain name of the web server processing the request	www.domain.com
cs(referrer)	Referring URL	Contents of the HTTP referrer field sent by the client	http://www.referringsite.com
cs(user-agent)	User Agent	Device used to make a request to the HTTP server	Mozilla/4.0+(compatible;+MSIE+6.0;+Windows+NT+5.1)
cs(cookie)	Client Cookies from Domain	Contents of all of the user's cookies for the site	KL_TC1 1038058778312 KL972x1038058778312282052 KL_PVKL972 0
cs-uri-query	Query String	The query string portion, if any, of the URI requested by the client	PAGENAME=dynamic1&link=3001

Understanding the Query String

Navigation title:Understanding the Query String

The query string (cs-uri-query) is often used by web applications and site developers to pass information from page to page because of the stateless nature of HTTP.

In many cases, information may be passed in the query string when it is acquired by **Sensor** at the web server. Such information can be used by **Site** to illuminate the true structure of the site, and the path of visitors through it, as well as other information.

In some dynamic websites, name=value pairs (variables) in the query string are important for determining the actual page being requested by a visitor. In such cases, URLs may be structured in the following or similar manner:

```
http://www.myserver.com/pageserved.asp?PAGENAME=HOME
```

In this example, PAGENAME is actually the indicator of what page will be served to the requester of this URL. Many web log analysis tools and services limit a site operator's ability to define what a page is in their site based on what query string variables occur in the query strings of the site's URLs. The data workbench server and data workbench can be configured to use such query names to define unique pages. This is important because many systems would interpret the following URLs as the same page, but **Site** does not.

```
http://www.myserver.com/pageserved.asp?PAGENAME=HOME
```

```
http://www.myserver.com/pageserved.asp?PAGENAME=HOME2
```

Similarly, site developers and applications often add many query string variables into a site's URLs that have nothing to do with identifying the actual page that is being requested. Examples are shown below:

```
http://www.myserver.com/pageserved.asp?PAGENAME=HOME&CAMPAIGN=10001
```

```
http://www.myserver.com/pageserved.asp?PAGENAME=HOME&CAMPAIGN=10002
```

```
http://www.myserver.com/pageserved.asp?PAGENAME=HOME&CAMPAIGN=10003
```

In this example, the query string variable CAMPAIGN= has been added to the URL. This CAMPAIGN variable is being used to indicate which marketing campaign caused a visitor to select this URL. **Site** can be configured to use this CAMPAIGN information, yet separate it from the definition of what page a visitor viewed so that in your list of pages for reporting and analysis purposes you would simply see the following:

```
http://www.myserver.com/pageserved.asp?PAGENAME=HOME
```

Filtering By Content Type

Navigation title:Filtering By Content Type

The purpose of **Sensor**'s content-type filtering capability is to eliminate the need to store and process information that is not useful for analysis purposes.

Much of the request data that is available through a web server's API is not useful in business analysis. Storage and processing are expensive and **Sensor**'s content-type filtering allows you to avoid unnecessary storage and processing.

To maximize the web log data processing performance and reduce the amount of measurement data that must be stored, **Site** acquires measurement data (request data, log entries, log data, and so on) for all web content-type requests, except for specifically listed content types (such as cascading style sheets, image requests, and so forth), which are filtered out before they are transmitted to the data workbench server by **Sensor**. This filtering can be disabled for an entire web server, and it also can be overridden for a particular content object by adding the name-value pair "Log=1" to the query string of a particular embedded object (for example,
`http://www.mysite.com/advertisement.gif?Log=1`).

Performing Controlled Experiments

Navigation title:Performing Controlled Experiments

Sensor enables controlled website experimentation, allowing you to run controlled experiments on random subsets of your visitor population with minimal disruption to your ongoing operations.

Controlled experiments allow experimentation with hypotheses for improving website financial performance, improving customer experience, or influencing customer behavior. Marketing programs, content personalization initiatives, process flows, and creative or content concept hypotheses can be tested on a controlled percentage of randomly selected site visitors and then analyzed in data workbench to determine the merit of the hypotheses.

Acquiring Embedded Object Requests (Page Tags)

Navigation title:Acquiring Embedded Object Requests (Page Tags)

After the HTML of a page is requested by a browser, the browser requests the embedded objects referenced in the HTML of that page from a web server to fill in the page shown by the browser.

Such embedded object requests are most commonly requests for image files or JavaScript files, though there are hundreds or maybe thousands of types of embedded objects used on the Internet today. Many of these embedded object requests are not generally useful in analyzing or reporting on the business activity of an Internet site; many such requests are therefore not desirable for acquisition unless they have a specific business purpose, such as presenting an advertisement or taking another measurement of site activity.

For example, an image may be an advertisement, and you may want to know that the advertisement was impressed upon a visitor. A JavaScript snippet may be in use to take a measurement that the particular browser has a certain characteristic and pass it back to a **Sensor** for acquisition. Each page on a site may have 10 or 100 embedded object requests in it. If a site stores the log information for each of these requests, the amount of data storage needed to keep the log data available for future analysis is multiplied by the number of embedded object requests for each page requested. For this reason, **Site** lets you keep the requests that are important for analysis and discard others before you incur unnecessary storage costs.

By using the override feature provided in the content-type filtering capabilities of **Sensor** (appending “Log=1” to the query string of an embedded object request URL), that particular embedded object request and the related measurement data can be acquired without requiring the site manager to store all requests of that type (for example, all <image> requests).

Sensor collects the measurement data in the following table for each embedded object request made of the web server, assuming that **Sensor** is not configured to filter it out or that the filter has been overridden. The collected information is related to the visitor and session and subsequent sessions through the x-trackingid or cs(cookie) log field entries.

W3C Name	Data Collected	Explanation	Example
x-trackingid	Tracking Identifier (unique visitor)	Identifier read from a cookie placed in the user's browser by Sensor on initial request	V1st=3C94007B4E01F9C2
Date Time	Timestamp	Time at which request was processed by server (at 100ns precision; accuracy depends on server environment and NTP)	2002-11-21 17:21:45.123
sc(content-Type)	Content Type	Type of object returned from server	text/html
sc-status	HTTP Response Status Code	Numerical code generated by the server that notes the status of the HTTP server's response	200
cs-uri-stem	URI Stem	The “stem” portion of the URI requested by the client	pagedir/page.asp
c-ip	Client IP	IP Address of the requesting client	127.0.0.1
s-dns	Server Domain Name	Domain name of the web server processing the request	www.domain.com
cs(referrer)	Referring URL	Contents of the HTTP referrer field sent by the client	http://www.referringsite.com
cs(user-agent)	User Agent	Device used to make a request to the HTTP server	Mozilla/4.0+(compatible;+MSIE+6.0; +Windows+NT+5.1)

W3C Name	Data Collected	Explanation	Example
cs(cookie)	Client Cookies from Domain	Contents of all of the user's cookies for the site	KL_TC1 1038058778312 KL972 x1038058778312282052 KL_PVKL972 0
cs-uri-query	Query String	The query string portion, if any, of the URI requested by the client	PAGENAME=dynamic1&link=3001

Acquiring Measurement Data Through Cookies

Navigation title:Acquiring Measurement Data Through Cookies

As a part of the Baseline Measurement data collected, **Sensor** collects the domain cookies sent from a visitor's machine when making a request from your web server.

This includes both persistent and session cookies that your website sets when a visitor interacts with your system.

In most cases, websites set persistent cookies to identify visitors or capture user input for use within subsequent visitor sessions. Any information written to and stored within persistent cookies can be captured and used alongside all other measurement data within the data workbench server.

An example of such a persistent cookie could involve a customer identifier in the form of a numeric key present within a domain specific cookie residing on the visitor's machine. In addition to identifying the user as a return visitor, the persistent cookie could also be used to further identify the visitor as a returning customer or to tie the visitor to information contained within a customer database to allow offline customer demographic information to be displayed within **Site** and used for interactive analysis.

Session cookies can be a good mechanism to collect user input through form fields or other dynamic interactive elements within your website. In the case of a website implementing forms to capture user-specific input data, the information remains in the session cookie only for as long as the session is active. When a user leaves your website or subsequently ends a session, the information is no longer stored on the user's computer. However, the information entered is captured by **Sensor** and made available as measurement data within **Site**.

Following is an example of using a session cookie to capture a single form variable entered by a visitor.

```
<html>
<head>
<title>Cookie Collection </title>

<script language="JavaScript">
function AppendFormValues()
{
    var item = document.testform.elements[i];
    var formitem = "v_"+i;
    var formvalue = item.value;
    cookie += formitem + "=" + formvalue + "&";
    document.cookie = cookie;

    testform.submit();
}

</script>
</head>
<body>
<form name="testform" method="post" action="nextpage.asp">
<input type="text" size=15 name="name"><br />enter name
<br><br>

<a href="javascript:AppendFormValues();">Click Here To </a><br /><br />
```

```

<br /><br /><br />

</body>
</html>

```

In this example, a function is called to set a session cookie on the visitor's machine with the name of the field and the value entered into the form field. As the form is submitted, and the subsequent web page is requested, the session cookie set is passed to the web server and collected by **Sensor**. The following data is therefore available within the data workbench server for use in data analysis:

Data Collected	Explanation	Example
v_1	Value associated with the v_1 cookie. This value represents the NAME entered into the form field that resulted in the session cookie being set.	v_1=John Smith

Session cookies may also be utilized to iteratively capture form fields or a multitude of embedded JavaScript variables present within an HTML page. In the following example, JavaScript is used to recursively capture any form field present within an HTML file and set a session cookie with the appropriate name=value pairs.

```

<script language="JavaScript">

function AppendFormValues()
{
    var cookie="formcookie=";
    for (i=0; i<document.testform.length; i++){
        if (document.testform.elements[i].type == "radio") {
            if (document.testform.elements[i].checked){
                var item = document.testform.elements[i];
                var formitem = "v_"+i;
                var formvalue = item.value;
                cookie += formitem + "=" + formvalue + "&";
            }
        }
        else if (document.testform.elements[i].type == "select") {
            var item = document.testform.elements[i];
            var formitem = "v_"+i;
            var optionindex = eval(document.testform.elements[i].selectedIndex);
            var formvalue = document.testform.elements[i].options[optionindex].value;

            cookie += formitem + "=" + formvalue + "&";
        }
        else{
            var item = document.testform.elements[i];
            var formitem = "v_"+i;
            var formvalue = item.value;
            cookie += formitem + "=" + formvalue + "&";
        }
    }
    document.cookie = cookie;
    document.testform.submit();
}

</script>

```

In this example, a session cookie is set on the visitor's machine with the name and value of every form field that exists within the form. This includes input fields, check boxes, radio buttons, select boxes, and text areas. As you may notice in this example, because the number of form fields is unknown, it is necessary to capture all form name and field values as a single string, delimited by an ampersand. This step must be taken because of a limit on the number of cookies a user may have on his or her computer at one point in time. Microsoft Internet Explorer allows only twenty (20) session cookies to be present before it begins dropping the oldest.

Acquiring Extended Measurements

Navigation title:Acquiring Extended Measurements

Conceptual information on acquiring extended measurements, and tracking link clicks and exits to external links.

Acquiring Dynamic Page Names

Navigation title:Acquiring Dynamic Page Names

For some sites, it is necessary to use embedded object requests to pass information to the web server so that details about what page was actually served may be acquired by **Sensor** and used for reporting and analysis.

This might be required if the page's URL, as seen by the web server, is not indicative of the page content that is shown to the browser. This case often results from the use of personalization or dynamic content management systems in which the actual content served in a page is determined on the fly by the URL, the cookie, related data and application logic.

The implementation of an embedded object to gather additional measurements should have minimal impact on your overall site performance. Adobe suggests that you embed a JavaScript file as the object used to collect the extended attributes. (Note that a JavaScript file can be embedded without any potential impact to the layout and presentation of your web page as may result with the use of an embedded image.) To accurately capture the information passed within the embedded object, Adobe also suggests that a common name be used. For naming purposes, Adobe refers to this object as `zig.js`. The `zig.js` file should be created within the appropriate directory on a web server on which **Sensor** is installed. This file needs to exist so that the request does not return a 404 error code. The contents of the file itself are not important. You should use a blank file named `zig.js` to minimize the amount of network traffic incurred as a result of the request.

For **Sensor** to collect a usable name for the page that was actually served, a few lines of JavaScript code must be added to the dynamic pages that you want to track or to which you want to add a unique page name. This code embeds a snippet of JavaScript in the page, which causes a tertiary embedded object request to be made to the web server as the page is loading. That request sends details about the specific page that was served back to the web server. The name of the page that was actually served is carried back to the web server as a variable in the query string of the embedded object (in this case JavaScript) request.

In general, the object request embedded in each such HTML page should look like the following:

```
<!-- BEGIN REFERENCE PAGE TAG-->
<script language="javascript">
var vlc = "0" //Capture Link Click 1=TRUE, 0=FALSE
var v = {};
v["_pn"] = "Application Form";
</script>

<script language="javascript" src="http://www.myserver.com/path/to/zig.js"
type="text/javascript"></script>

<noscript>

</noscript>

<!-- END REFERENCE PAGE TAG-->
```

`Log=1` ensures that **Sensor** logs the request in spite of the **Sensor** content type filtering rules to the contrary, such as the filtering out of JavaScript and image requests before they are stored. The declared `v_pn` variable identifies the name of the actual page content being served so that **Site** knows the name of the page the visitor actually viewed. The `v_pn` value could be established manually or by other script or application code.

After the value is collected, you can configure the data workbench server to use the contents of the query string variable (name=value pair, for example, `v_pn=Application Form`) appended to the `zag.gif` URI (for example, `http://www.mysite.com/pageserved.asp?v_pn=Application%20Form`), as an augmentation

of the `zag.gif` URI. In addition to the baseline measurements acquired with every HTTP request, an extended measurement would be acquired with this request.

Data Collected	Explanation	Example
<code>v_pn=</code>	Value associated with the <code>v_pn</code> query string variable	<code>v_pn=Application Form</code>

Acquiring Additional Information

Navigation title:Acquiring Additional Information

Query string variables may be added to a JavaScript request to collect additional measurements when a request is made.

These variables may be added manually or by script in the page itself.

Additional information that can be acquired from a page may be added to the embedded object via script using the following code as an example:

```
<!-- BEGIN REFERENCE PAGE TAG-->
<script language="javascript">
var vlc = "0" //Capture Link Click  1=TRUE, 0=FALSE
var v = {};
v["_pn"] = "Application Form";
v["_1"] = "99.99";
v["_2"] = "visa";
</script>

<script language="javascript" src="http://www.myserver.com/path/to/zig.js"
type="text/javascript"></script>
<noscript>


</noscript>
<!-- END REFERENCE PAGE TAG-->
```

In this example, the script variables for `v_1` and `v_2` can be derived from another function within your web page. The variables have been inserted as examples. In addition to the baseline measurements acquired with each request, the following extended measurements would be acquired with the request of the URL above:

Data Collected	Explanation	Example
<code>v_pn=</code>	Value associated with the <code>v_pn</code> query string variable	<code>v_pn=Application Form</code>
<code>v_1=</code>	Value associated with the <code>v_1</code> query string variable	<code>v_1=99.99</code>
<code>v_2=</code>	Value associated with the <code>v_2</code> query string variable	<code>v_2=visa</code>

Acquiring Document Objects

Navigation title:Acquiring Document Objects

Using the JavaScript Document Object Model, additional scripting methods can be employed to augment the request for the `zig.js` file.

Information such as the value of META tags, ID values of DIV tags, and so forth, can be referenced by name and collected as variables for use in analysis. For example, to dynamically capture the information contained within the META element of the HTML document, you can use the following JavaScript syntax:

```
<!-- BEGIN REFERENCE PAGE TAG-->
<script language="javascript">
var m0 = document.getElementsByTagName('META')[0]; //define the first
instance of META
var metacontent = m0.getAttribute('content'); //get the 'content' value of
META
```

```

var vlc = "0" //Capture Link Click 1=TRUE, 0=FALSE
var v = {};
v["_1"] = metacontent;
</script>

<script language="javascript" src="http://www.myserver.com/path/to/zig.js"
type="text/javascript"></script>

<noscript>

</noscript>

<!-- END REFERENCE PAGE TAG-->

```

Data Collected	Explanation	Example
v_1=	Value associated with the METAVALUE query string variable. This value represents the data within the META element of the HTML document.	v_1=This page serves content related to the order thank you page.

After the data is collected, you can configure the data workbench server to process this measurement data for the purposes of analysis and reporting.

Measuring Advertisement Impression

Navigation title:Measuring Advertisement Impression

Marketing your website may involve the placement of advertisements in the form of image or other rich media files (served from your web server) on third-party websites.

In such cases, you might want to measure both the impression of the advertisement on a browser and the subsequent click-through, if one occurs, to the advertisement's target URL on your website.

For advertisements in the form of images, appending Log=1 to the query string results in the image request, and thus the advertisement impression, being captured by **Sensor** for analysis purposes.

```

<!--REFERENCE IMPRESSION TAG-->
<IMG SRC="http://www.mysite.com/images/zag.gif?Log=1&v_ic=CAMPAIGN
1&v_ica=72890ab&v_icr=http://money.cnn.com/markets/" />
<!--END REFERENCE IMPRESSION TAG-->

```

Data Collected	Explanation	Example
v_ic=	Value denoting the Impression Campaign	v_ic="CAMPAIGN1"
v_ica=	Value denoting the Impression Campaign Asset	v_ica="72890ab"
v_icr=	Value denoting the Impression Campaign Referrer	v_icr="http://money.cnn.com/markets/

In addition to appending Log=1 to the image request, an identifier should be added to the URL leading from the advertisement to the target page within your website to track the advertisement that led to the click-through and to track the click-through back to the particular campaign for that advertisement.

```

<a
href="http://www.mysite.com/path/to/landingpage?Log=1&v_c=CAMPAIGN&v_ca=72890ab&v_cr=http://money.cnn.com/markets/">
Click Here
</a>

```

Data Collected	Explanation	Example
v_c=	Value denoting the Click-through Campaign	v_c="CAMPAIGN1"
v_ca=	Value denoting the Click-through Campaign Asset	v_ca="72890ab"
v_cr=	Value denoting the Click-through Campaign Referrer	v_cr="http://money.cnn.com/

Data Collected	Explanation	Example
		markets/

Tracking Link Clicks

Navigation title: Tracking Link Clicks

Steps used to facilitate the collection of Link Clicks through the use of the Reference Page Tag.

Through the deployment of the **Reference Page Tag**, it is possible to collect measurement data denoting the links (or href values) that visitors click while visiting particular pages. Typically, this collection does not involve the implementation of additional link identifiers into your HTML pages.

To facilitate collection of Link Clicks through the use of the **Reference Page Tag**, complete the following steps:

1. Copy the following code into the existing file named zig.js:

```
//REFERENCE LINK AND FORM CLICK PAGE TAG
//INITIATE FUNCTIONS ONLOAD
function addEvent(obj, evType, fn){
    if (obj.addEventListener){
        obj.addEventListener(evType, fn, false);
        return true;
    } else if (obj.attachEvent){
        var r = obj.attachEvent("on"+evType, fn);
        return r;
    } else {
        return false;
    }
}
addEvent(window, 'load', startCapture);
addEvent(window, 'load', startCapture);
function startCapture(){
//TO CAPTURE LINK CLICKS
    if (vlc == "1"){captureLink();}
//TO CAPTURE FORM FIELD CLICKS
    if (vfc == "1"){captureForm();}
}
//BEGIN LINK CAPTURE PAGE TAG
function captureLink(){
    if (document.links[0]){
        if (document.links){
            var links = document.links, link, k=0;
            while(link=links[k++]) {
                link.onclick = captureLinkName;
            }
        }
    }
}
function captureLinkName() {
    var lc=new Image();
    this.parent = this.parentNode;
    lc.src='zag2.gif?linkname=' + escape(this.name) + "&cd=" + new Date().getTime();
}
//END LINK CAPTURE PAGE TAG
//BEGIN FORM CLICK CAPTURE PAGE TAG
function captureForm(){
    if (document.forms[0]) {
        if(document.forms){
            for(i=0; i<document.forms[0].elements.length; i++){

```

```

        var forms = document.forms[0].elements[i];
        forms.onfocus = captureFormName;
    }
}
}

function captureFormName() {
    var fc=new Image();
    fc.src='zag3.gif?fieldname=' + escape(this.name) + "&cd=" + new
Date().getTime();
}
//END FORM CLICK CAPTURE PAGE TAG

```

2. Create or place the 1 pixel by 1 pixel image file named `zag2.gif` into a directory present on your web server.
3. Modify the `lc.src` variable to reference the appropriate domain of your website from which the `zag2.gif` file is referenced.
4. Ensure appropriate cache-control headers are established for the `zag.gif` and `zig.js` files.
5. Within the HTML files you desire to collect link click values from, the **Reference Page Tag Execution Call** must be modified to inform the **Page Tag Execution Script** to capture link clicks for that page. To do so, change the `vlc` variable value to "1" as highlighted in the following code example:

```

<!-- BEGIN REFERENCE PAGE TAG-->
<script language="javascript">
var vlc = "1" //Capture Link Click 1=TRUE, 0=FALSE
var vfc = "0"; //Capture Form Field Click 1=TRUE, 0=FALSE
var v = {};
</script>

<script language="javascript" src="http://www.myserver.com/path/to/zig.js"
type="text/javascript"></script>

<noscript>

</noscript>

<!-- END REFERENCE PAGE TAG-->

```

Data Collected	Explanation	Example
<code>v_ln=</code>	Value denoting the Impression Campaign	<code>v_ln="About%20Us"</code>

Tracking Exits to External Links

Navigation title: Tracking Exits to External Links

Capturing activity across third-party website links to enable Exit Target analysis.

Web pages can contain links to third-party websites, and activity across those links can be captured to enable Exit Target analysis, especially in the case that the third-party site is responsible for paying referral fees when such referrals are received. Because the click event is written to the log files of the third-party system by default, modifications need to be made to the link for the click event to be captured locally. The third-party link present within your website must be modified as follows:

```
<A
HREF="http://www.myserver.com/PageExit.htm?v_eurl=http://www.othersite.com">
```

The referenced `PageExit.htm` file must be created and should be structured to contain the following script:

```
<html>
<head>
```

```

<script>
function getExitURLQuery(variable)
{
    var query = window.location.search.substring(1);
    var vars = query.split("&");
    for (var i=0; i<vars.length; i++)
    {
        var pair = vars[i].split("=");
        if (pair[0] == variable)
        {
            return pair[1];
        }
    }
}
</script>

<script>
location.replace(getExitURLQuery("v_eurl"));
</script>

</head>
</html>

```

By making the request for the PageExit.htm file, the v_eurl value is collected for analysis purposes. Additionally, when PageExit.htm is loaded, it immediately redirects to the specified v_eurl target location.

Data Collected	Explanation	Example
v_eurl	Value associated with the v_eurl query string variable. This value represents the target URL of the link present within the HTML page.	v_eurl=www.othersite.com

Acquiring Form Field Input

Navigation title:Acquiring Form Field Input

Information on acquiring form field input, including general information and .asp-specific information.

General Information

Navigation title:General Information

Values entered into a form in a web page can be collected and appended in the query string of the subsequently requested page (on form submission) through the use of JavaScript.

This is shown in the following example. Include this JavaScript after any form validation scripts used in your HTML pages.

```

<html>
<head>
</head>
<script language="JavaScript">

function AppendFormValues( )
{
    for(var i = 0; i < document.formname.length; i++)
    {
        var item = document.formname.elements[i];
        var formitem = "v_"+i;
        var formvalue = item.value;
        formvalues += formitem + '=' + formvalue + '&';
    }
    document.formname.action = document.formname.action + '?' + formvalues;
}

```

```

}
</script>
<body>
<form name="formname" action="thankyou.asp" method="POST"
onSubmit="AppendFormValues();">
<input name="NAME" size="50" value=""></input>name<br/>
<input name="CITY" size="50" value=""></input>city<br/>
<input name="STATE" size="50" value=""></input>state<br/>
<input name="ZIP" size="10" value=""></input>zip<br />
<input type="submit" name="submit" value="submit"/>
</body>
</html>

```

This example appends the values entered into the form by the browser user to the subsequent “thankyou.asp” page indicated in the FORM Action value as follows:

```
http://www.myserver.com/thankyou.asp?v_1=John Smith&v_2=Los Angeles&v_3=California&v_4=90210
```

The following extended measurements would be acquired with this request in addition to the baseline measurements collected by **Sensor**:

Data Collected	Explanation	Example
v_1	Value associated with the NAME form field	v_1=John Smith
v_2	Value associated with the CITY form field	v_2=Los Angeles
v_3	Value associated with the STATE form field	v_3=California
v_4	Value associated with the ZIP form field	v_4=90210

ASP-Specific Information

Web pages are often structured using ASP (Active Server Pages) programming language.

ASP is a Microsoft technology that runs within IIS (Internet Information Services). When a browser requests an ASP file, IIS passes the request to the ASP engine. The ASP engine reads the ASP file, line by line, and executes the scripts in the file. Finally, the ASP file is returned to the browser as plain HTML. ASP provides RESPOND or REQUEST objects which, in addition to other uses, allow the response or request of user queries or data submitted from HTML forms.

In certain cases, you may not want to append the values entered into forms to the URL that is displayed within the Address bar of a user’s browser or that is viewable within the HTML code itself. Simple server-side JavaScript allows you to append form field names and their respective values to the log file without making them available within the user’s browser or embedding them into the HTML file. To capture the actual form values entered into particular forms within your website, a few lines of code must be added to append the form values to the log request.

Within the processing page of a form, include the following code to append the entered form values to the request data (in addition to writing the submitted form values to an external database or other location):

```

var sName= Request.Form("Name");
var sCity= Request.Form("City");
var sState= Request.Form("State");
var sZip= Request.Form("Zip");

Response.AppendToLog(" &v_1=" + sName);
Response.AppendToLog(" &v_2=" + sCity);
Response.AppendToLog(" &v_3=" + sState);
Response.AppendToLog(" &v_4=" + sZip);

```

This process would append the form values as defined to the request data for the **Form Processing** page. Within the log data, the appended values would be available as query strings of the **Form Processing** page as illustrated below. For example, v_1, v_2, v_3 and v_4 would now be query strings containing the data entered into the appropriate form fields. The syntax described in the example above can be duplicated for any additional form fields and values that you want to capture.

```
http://www.myserver.com/path/to/formprocessingpage.asp?v_1=John+Smith&v_2=Los+Angeles&v_3=California&v_4=90210
```

If you want every form field and value to be captured and available for analysis, you can use the following syntax:

```
var formvalues = Response.Form;
Response.AppendToLog(formvalues);
```

This example would take all form fields present within the HTML along with their respective values and append them as query strings to the log entry for the **Form Processing** page. It should be noted that this would include any hidden fields present within the form.

The log data would be augmented as detailed in the following table:

Data Collected	Explanation	Example
v_1	Value associated with the NAME query string	v_1=John Smith
v_2	Value associated with the CITY query string	v_2=Los Angeles
v_3	Value associated with the STATE query string	v_3=California
v_4	Value associated with the ZIP query string	v_4=90210

Tracking Visitor Activity Within Flash Rich Media Content

Navigation title: Tracking Visitor Activity Within Flash Rich Media Content

Websites architected using Flash require special attention with respect to the capture of visitor actions performed within the rich media content.

Using Flash ActionScript, you can make simple changes to your existing Flash movies to allow the tracking of all visitor interactions with the movie, such as button clicks or mouse movements.

To facilitate Visitor activity tracking within your Flash movie, please follow the steps listed below:

1. Add the following ActionScript code to your movie. This code represents a function that can be called by events within the Flash movie that you want to track.

```
// FLASH TAG CODE BEGIN
var FLASHTAGURI = "[PATH_TO_WEB_SERVER]/flashtag.txt";
function tag(PAGENAME, VARIABLES) {
    loadVariablesNum(FLASHTAGURI + "?" + "PAGENAME=" + PAGENAME + "&" + VARIABLES, 0);
}
// FLASH TAG CODE END
```

2. Create a blank file named flashtag.txt and place the file on your web servers.
3. Within the function in Step 1, replace the [PATH_TO_WEB_SERVER] placeholder with the fully qualified or relative path to the location of the flashtag.txt file. For example:

```
var FLASHTAGURI = http://www.mysite.com/flashtag/flashtag.txt";
```

4. Add the following ActionScript code to all events to be tracked. This code represents a function call used to capture data about the event:

```
on(release) {tag(" [PUT_PAGE_NAME_HERE] ", "[PUT_ADDITIONAL_VAR_HERE]");}
```

This example illustrates the use of the on(release) event; however, the tag() function may be referenced through any event that you may want to track, such as an on(press), on(rollover), on(rollout), or on(keypress) event.

The [PUT_PAGE_NAME_HERE] placeholder should be replaced with a string that represents the name of the page or event that you are tracking. The [PUT_PAGE_NAME_HERE] variable can be modified either manually or through variable reference to denote a unique name for the page or event within the Flash application. The value replacing the [PUT_PAGE_NAME_HERE] placeholder may consist of a simple name or may be structured to represent a hierarchical structure similar to a full URI. For example:

```
on(release) {tag( "/about_us/index.swf" , "[ PUT_ADDITIONAL_VAR_HERE ]" );}
```

Adobe recommends that, prior to code deployment, you compile a written specification for page names and event names to facilitate the alignment of business requirements and development tasks and reduce the potential for additional development cycles.

5. If desired, additional variables may be collected and associated with pages or events in the Flash movie. To do so, replace the [PUT_ADDITIONAL_VAR_HERE] placeholder with a set of name=value pairs separated by an ampersand (&). For example:

```
on(release) {tag( "/about_us/index.swf" , " var1=value1&var2=value2" );}
```

The variables can be modified either manually or through variable reference to denote additional attributes to be collected and associated with the page or event. If there are no applicable additional variables to collect, remove [PUT_ADDITIONAL_VAR_HERE].

Your setup of visitor tracking within Flash rich media content is now complete. When the event is invoked, the tag (PAGENAME , VARIABLES) function will be called, resulting in an HTTP request being made for the following file. This function will be called in addition to other functions that may be triggered as defined within your Flash movie:

http://www.mysite.com/flashtag/flashtag.txt?PAGENAME=/about_us/index.swf&var1=value1&var2=value2

The HTTP request resulting from the Flash Tag ActionScript function results in the following information being collected with respect to each event within the Flash movie. The last row in the table (W3C Name cs-uri-query) represents the information collected for the additional variables specified in your function call.

W3C Name	Data Collected	Explanation	Example
x-trackingid	Tracking Identifier (unique visitor)	Identifier read from a cookie placed in the user's browser by Sensor on Visitor's initial request	v1st=3C94007B4E01F9C2
Date	Timestamp	Time at which request was processed by server (at 100ns precision; accuracy depends on server environment and NTP)	2002-11-21 17:21:45.123
Time			
sc(content-Type)	Content Type	Type of object returned from server	Text/html
sc-status	HTTP Response Status Code	Numerical code generated by the server that notes the status of the HTTP server's response	200
cs-uri-stem	URI Stem	The stem portion of the URI requested by the client	/flashtag/flashtag.txt
c-ip	Client IP	IP Address of the requesting client	127.0.0.1
s-dns	Server Domain Name	Domain name of the web server processing the request	www.mysite.com

W3C Name	Data Collected	Explanation	Example
cs(referrer)	Referring URL	Contents of the HTTP referrer field sent by the client	
cs(user-agent)	User Agent	Device used to make a request to the HTTP server	Mozilla/4.0+(compatible;+MSIE+6.0; +Windows+NT+5.1)
cs(cookie)	Client Cookies from Domain	Contents of all of the user's cookies for the site	KL_TC1 1038058778312 KL972x1038058778312282052 KL_PVKL972 0
cs-uri-query	Query String	The query string portion, if any, of the URI requested by the client	PAGENAME=about_us&sw=var1-value1&var2-value2

P3P Considerations for Third-Party Page Tagging

Navigation title:P3P Considerations for Third-Party Page Tagging

Conceptual information about third-party tagging and preventing cookie-blocking using P3P.

Understanding Third-Party Page Tagging

In most implementations, the Adobe persistent cookie is viewed as a first-party cookie. First-party cookies are defined as those associated with the host domain.

Suppose a user visits <http://www.example.com/>. Assuming that a Sensor is installed and operational on the web server hosting the domain, an Adobe persistent cookie is set and viewed as a first-party cookie. You may, however, want to collect measurement data from a third-party site through the use of embedded objects, which are requested and loaded from your server that is running **Sensor** instead of from the third-party server hosting the page or advertising program. For example, <http://www.example2.com/> serves a web page with an embedded object request served from <http://www.example.com/>. The request for the embedded object from <http://www.example.com/> results in a cookie being set; however, in this context, the web browser or user-agent views the cookie as a third-party cookie.

In newer web browsers such as Microsoft's IE6, privacy features filter cookies based on their compact policies sent in the HTTP response header from the web server. In the default IE6 settings, which most users never change, third-party cookies are blocked when they have nonexistent or unsatisfactory compact policies. Most sites that are experiencing cookie-blocking problems have third-party cookies on their site that do not have the appropriate compact policies being sent in the HTTP response header. Additionally, some cookie-blocking problems occur when a site is framed by another site. For example, an online store that is part of an online shopping portal may appear in a frame provided by the portal. From the perspective of the browser, the store content may appear to be third-party content when framed by the portal. However, if a visitor goes directly to the online store without going through the portal, the content will be first-party content. Thus, the online store finds their cookies are blocked only when visitors come in through the portal.

Web-based mail systems cause a similar problem. If a website visitor emails a web page to a friend who uses a web-based mail system, the email message appears as third-party content to the friend's browser because it is framed by the email system. If there are any cookies associated with the page that was emailed, they are treated as third-party cookies by IE6.

Using P3P to Prevent Cookie-Blocking

P3P provides a standard way for websites to encode their privacy policies in a computer-readable XML format. P3P-enabled web browsers and other P3P user agents automatically fetch P3P privacy policies, parse them, and compare them with a user's privacy preferences.

To prevent IE6 from blocking cookies on your site, you need to ensure the following:

1. All of the cookies that are being set in a third-party context have P3P compact policies associated with them.
2. The appropriate compact policy is sent using a custom HTTP response header.
3. Those compact policies are considered satisfactory by IE6.
4. If the third-party cookies are being set by another company, you may need to ask them to enable P3P and set P3P compact policies. Any host that sets a P3P compact policy also must have a corresponding full P3P policy. Users can change their IE6 settings so that cookies are blocked under other conditions as well; however, placing satisfactory compact policies on third-party cookies prevents most IE6 cookie-blocking.

The following is an example of such a P3P header:

```
P3P: policyref=" http://www.myserver.com/w3c/p3p.xml ", CP="NOI DSP COR PSA  
PSD OUR IND COM NAV"
```

In this example, the file p3p.xml is used to reference an associated policy.xml file residing on your web server that denotes the kinds of information your website collects, dispute resolution methods that your organization adheres to, how the data collected is used, who owns the data, and other standard information related to Internet Privacy. The three character codes following the “CP” are the compact policy codes that emulate what is stated within your policy.xml file.

All compact policies and policy XML files should be tailored for the respective organization for which they are being deployed, accurately specifying their internal privacy policies with regard to website data collection. A multitude of P3P policy editors can be found online along with more in-depth information relative to implementing an appropriate privacy policy within your website.

For more information on how Internet Explorer 6 handles P3P Headers, please visit:

<http://msdn.microsoft.com/library/default.asp?url=/library/en-us/dnpriv/html/ie6privacyfeature.asp>

Page Tag Variable Mapping Worksheet

Navigation title:Page Tag Variable Mapping Worksheet

Tables containing list of constant and custom variables.

Constant Variables

Variable Name	Variable Short Description	Additional Information
ct	constant tag	MODIFIED WITHIN JS FILE
cd	constant domain	MODIFIED WITHIN JS FILE
cu	constant uri	MODIFIED WITHIN JS FILE
ce	constant end	MODIFIED WITHIN JS FILE
c	constant	DO NOT SET
sw	screen width	DO NOT SET
sh	screen height	DO NOT SET
cd	color depth	DO NOT SET
d	dynamic	DO NOT SET
dt	document title	DO NOT SET
dr	document referrer	DO NOT SET
cb	cache busting	DO NOT SET

Variable Name	Variable Short Description	Additional Information
v	visual custom	DO NOT SET
v_jd	visual custom	DO NOT SET

Custom Variables

Variable Name	Variable Short Description	Additional Information
v_pn	Page Name	v_pn="Application_Form"
v_pc	Page Category	v_pc="Application_Process"
v_pe	Page Error	v_pe="Validation_Error"
v_c	Click-through Campaign	v_c="Campaign 1"
v_ca	Click-through Campaign Asset	v_ca="72890ab"
v_cr	Click-through Campaign Referrer	v_cr=http://money.cnn.com/markets/&v_cp=CNNMy_Mkt"
v_ic	Impression Campaign	v_c="Campaign 1"
v_ica	Impression Campaign Asset	v_ca="72890ab"
v_icr	Impression Campaign Referrer	v_cr=http://money.cnn.com/markets/&v_cp=CNNMy_Mkt"
v_ln	Link Name	v_ln="About%20Us"
v_eurl	Exit URL	V_eurl="http://www.offsite.com/
v_1	custom variable 1	
v_2	custom variable 2	
v_3	custom variable 3	

Data Workbench Metrics, Dimensions, and Filters

Metrics, dimensions, and filters provide a framework within which calculations are made about the data processed into a data workbench dataset.

The results of the calculations defined using this framework are shown in workspaces, dashboards, reports, or other outputs. In short, any number that you see in or from an application is the result of a query of a dataset that involves a metric, a dimension, and a filter.

At the most basic level, a metric describes what is being calculated from and about the dataset, a dimension breaks down the data in the dataset into categories, and a filter describes a selected portion or subset of the data in the dataset.

When the Data Workbench Server processes data to create a dataset, dimensions of the data are created and then updated on an ongoing basis as new data is read and processed by the server. Metrics and filters are calculated from these dimensions of data.



Caution: If you redefine an internal metric, the system will behave unexpectedly because of the wrong value. Your reports will not generate unless a metric reads 100%. It is recommended that you do not change metric definitions.

An Example

Imagine a dataset that contains information about all of the people in the world. This dataset contains, at a minimum, all of the people in the world and their ages. A useful metric to calculate from this dataset would be Average Age. Evaluating this metric would result in one number: the average age of the world's population.

Adding a dimension to the dataset makes this information more useful and manageable. If the dataset also contains each person's country of residence, defining a Country dimension would provide a way to segment the people into groups for each country in the world. Evaluating the Average Age metric over the Country dimension would result in a list of numbers, one for each country, each representing the average age of the people in that country.

The application of a filter (or selection filter) in a metric formula can give more detailed information or allow the definition of a new metric based on existing metrics and dimensions. Evaluating the Average Age metric with a filter of "where country equals Sweden" results in one number: the average age of people in Sweden. A metric based on this filter could be Swedish Average Age.

For example:

```
Swedish_Average_Age=Average_Age[country = 'Sweden']
```

How Metrics, Dimensions, and Filters Relate

In general, evaluating a metric over a dimension results in that metric being evaluated for each dimension element (or element). In the example above, the Country dimension has an element for each country of the world. Evaluating Average Age over Country would yield the average age for each of the elements (countries), including the element Sweden.

It is important to note that when you evaluate a metric over a dimension, you will receive the same numerical result for a specific dimension element regardless of whether you evaluate that metric for the whole dimension or you define a filter corresponding to that specific dimension element. Using the previous example, when looking for the average age of people in Sweden, either of the following methods would yield identical results:

- Evaluate the Average Age metric over the Country dimension and then look at the number for the dimension element Sweden.
- Evaluate the Average Age metric with a filter of "people in Sweden" (expressed as `Average_Age[Country='Sweden']`).

Filters are syntactical expressions that reference one or more dimensions and dimension elements. As you saw in the above example, using the expression `[dimension=element]` is an easy way to specify a filter.

It is just as easy to apply such a filter to define a new metric using an expression such as `New_Metric=Metric[Filter]`. Such a filter can be used to define a new metric based on a specific dimension element. To use the above example, `Average_Age[Country='Sweden']` specifies a metric for the average age of people in Sweden. If we were to give this metric a name, such as `Swedish_Average_Age`, we could use it in other calculations as a metric. For example, evaluating

`Swedish_Average_Age/Average_Age` would result in a single number: the ratio of the average age of people in Sweden to that of people in the rest of the world.

If the dataset with information about all of the people in world also includes a dimension Eye Color, the expression `Swedish_Average_Age[Eye_Color='green']` would result in the average age of Swedes with green eyes. You might also obtain this same result without using an intermediate metric definition by applying a different filter: `Average_Age[Country='Sweden' AND Eye_Color='green']`. In this case, the AND operator specifies a filter expression using two other basic filter expressions.

Expanding Metrics, Dimensions, and Filters

Data workbench enables you to create new metrics, filters, and dimensions to expand the information available through an application or to easily customize the definitions in the application to meet your needs. The open-ended nature of metrics, dimensions, and filters enables you to answer more questions as you add new metrics and dimensions.

Newly defined metrics can be evaluated over existing dimensions, and, similarly, existing metrics can be evaluated over newly defined dimensions. Newly defined metrics also can be used as input in other metrics. Because metrics are built on the client side (that is, their construction does not require dataset processing), they have a minimal impact on system performance. You can increase your analysis capabilities by constructing many levels of metrics.

For more information about the syntax used to manipulate metrics, dimensions, and filters, see [language syntax](#).

Dimension Types

Several types of dimensions are available in the data workbench server. As such, it is important to know the dimension type when using a dimension to create metrics, filters, or derived dimensions.

Insight Server can create and maintain the following types of dimensions:

Dimension Types	Description
Countable	A dimension type in which the number of elements in the dimension can be counted by the system. Countable dimensions must be derived from other Countable dimensions. Countable dimensions can be parents of other dimensions or children of other countable dimensions. Examples: Visitor, Session, Page View, Booking, and Order.
Simple	A dimension that has a one-to-many relationship with a parent countable dimension. A simple dimension can be thought of as representing a property of elements of its parent dimension. Example: Visitor Referrer is a simple dimension with a parent of the Visitor dimension. Each Visitor can have only one Visitor Referrer (their first HTTP referrer), but many Visitors might have the same Visitor Referrer. Therefore the Visitor Referrer is “one-to-many” with the Visitor dimension.
Numeric	A dimension that has ordered, numerical values and a one-to-many relationship with a parent countable dimension. A numeric dimension can be thought of as representing a numeric property of elements of its parent dimension. Numeric dimensions are often used to define “sum” metrics. Example: The numeric dimension Session Revenue defines the revenue, in dollars, for each Session. Each Session has a single amount of revenue, but any number of Sessions might have the same revenue, so Session Revenue is “one-to-many” with Session. A metric “Revenue” might be defined as <code>sum(Session_Revenue, Session)</code> , giving the total amount of revenue for the selected Sessions.
Many-to-Many	A dimension that has a many-to-many relationship with a parent countable dimension. A many-to-many dimension can be thought of as representing a “set” of values for each element of its parent dimension. A many-to-many dimension is equivalent to an (anonymous) countable dimension with its parent and a Simple dimension with a parent of the anonymous countable dimension. Example: The many-to-many dimension Search Phrase has a parent of Session. Each Session can use zero or more Search Phrases, and a Search Phrase can be used in any number of Sessions.
Denormal	A dimension that has a one-to-one relationship with a parent Countable dimension. The element names of the denormal dimension can carry information about the corresponding elements of the parent dimension. A denormal dimension can be thought of as storing an arbitrary string value for each element of the parent. Denormal dimensions can be used with Insight Server’s segment export capability to output details about a subset or “segment” of a countable dimension. In addition, denormal dimensions can be referenced in metric formulas and worksheet visualizations and can be used (with certain restrictions) to define filters. Example: The denormal dimension EMail Address has a parent of Visitor. Each Visitor has an EMail Address, and each element of the EMail Address dimension is associated with a single Visitor. Even if two visitors have the same e-mail address, their addresses will be different elements of the EMail Address

Dimension Types	Description
	dimension. A Segment Export can reference the EMail Address dimension to output the EMail Address of each visitor in a Segment.
Time	A dimension that enables you to create a set of periodic or absolute local time dimensions (such as Day, Day of Week, Hour, Hour of Day, and so on) based on a timestamp field that you specify. When defining time dimensions, you also can choose a day other than Monday to be used as the start of a week by specifying the Week Start Day parameter. Example: The time dimension Session Time has parent of Session. Therefore, the dimension defines a set of time dimensions (Day, Day of Week, Hour, Hour of Day, Month, and Week) whose elements correspond to the times at which visitors' sessions on the site began.
Derived	Derived dimensions, rather than being defined in the dataset configuration based on the data being processed, are defined in the profile based on other dimensions or metrics. Many derived dimensions are created automatically to drive different types of visualizations. Example, when a user builds a site or process map, Insight Server silently creates a "Prefix" dimension. Others, such as the reporting time dimensions, are defined by files in the Dimensions directory of a profile.

About Built-in Metrics, Dimensions, and Filters

This section provides detailed information about the built-in metrics, dimensions, and filters provided in data workbench.

Almost all metrics are defined in the metrics directory of a profile. A few built-in metrics are provided to convey technical information about the status of data processing through its two phases called log processing and transformation which are defined by Insight Server's log processing and transformation configuration files, respectively.



Note: The log processing phase does not mean that all event data comes from log files. Although it sometimes does, event data can come from another source, such as an ODBC data source. For more information about data sources, see the [Dataset Configuration Guide](#).

Built-in Metrics

Data workbench includes built-in metrics.

The following table lists the available built-in metrics for data workbench:

Built-in Metric	Description
As Of	The As Of time of the dataset in 100 nanoseconds intervals since January 1, 1600 00:00 UTC. The As Of timestamp is the latest time in the dataset for which all data has definitely been processed.
Log Bytes Read	The total amount of compressed data (in bytes) that has so far been processed during the log processing phase.
Log Bytes Total	The total amount of compressed data (in bytes) in log files matching the configured log sources criteria and start and end date range of the dataset. If log processing is paused in the Log Processing Mode configuration file then Log Bytes Total will be the same as Log Bytes Read.
Log Processing Progress	The percentage completion of the log processing phase of Insight Server data processing.
Total Decoded Log Entries	The number of entries in the log files that were successfully decoded as a part of the log processing phase of Insight Server data processing.
Total Filtered Log Entries	The number of entries in the log files that after being decoded were accepted by the robot filter and the log entry conditions and other filters that are applied as a part of the log processing phase of Insight Server data processing.
Total Log Entries	The number of entries in the log files that have so far been processed by the log processing phase of Insight Server data processing.

Built-in Metric	Description
Total Processed Log Entries	The number of filtered log entries that have been incorporated into the Adobe dataset.
Transformation Progress	The percentage completion of the transformation phase of Insight Server data processing.
Uncompressed Log Bytes Read	The total amount of uncompressed data (in bytes) that has so far been processed during the log processing phase.

Built-in Dimensions

Data workbench includes built-in dimensions.

The following table lists the available built-in dimensions for data workbench:

Name	Details	Level	Description
None	Derived	N/A	<p>Has a single element “” which relates to all elements of all dimensions. Evaluating a metric over None is like evaluating it over no dimension.</p> <p>The Filter [None=“ ”] is equivalent to [True].</p>
One (Hidden)	Numeric	N/A	<p>The element “1”, which also has ordinal value = 1, relates to all elements of all dimensions. The One dimension is normally used to construct counts by using this syntax:</p> <p><code>sum(one, Countable_Dimension)</code></p>

Built-in Profile Filters

Profile filters constrain the scope of the data available from a dataset.

In addition to the log processing and transformation filters that can be applied before or during the creation of a dataset, other profile filters are available that impact the scope of the data available for selection from a dataset. This section describes only the latter type of special filters.

The following profile filters are available to the user after a dataset has been created:

- Data Subsetting Filter
- Broken Session Filter



Note: Additional filters can be created and applied through their presence in a profile's Filters directory.

Data Subsetting

A data subset acts as a data filter by allowing you to select only the dimension elements of data that are of interest to you.

Creating data subsets decreases the amount of time required to calculate exact answers to your queries. If the data subset that you specify is small enough, Data Workbench can retrieve all needed data from Insight Server and answer questions about the subset quickly and accurately. This is especially useful if you know that, for example, analyzing a day of data or sessions from a particular referrer will require several hours.

Users can create data subsets themselves, or they can access data subsets defined in an inherited or working profile. When a user creates a subset of the dataset (by selecting the desired data in Insight and then right-clicking in the workspace and clicking Data > Subset), a Data Subset.filter file is created in the Filters folder within the User profile directory. This filter defines the data subset that you have selected and saves the subset for future use.



Note: You can create multiple data subsets and can toggle them to view different portions of the data. Remember to turn off Data Subsetting when you want to view all data. Otherwise, your metric values will not be representative of all data in the dataset.

Broken Session Filter

The Broken Session Filter is a metric formula that can be easily modified to meet any filtering requirements. In the default Site profiles, the Broken Session Filter is configured to include all visitors that have a Visitorized Flag that is set to 1. The value 1 indicates the presence of a tracking cookie for that visitor.

Following is the text of the default Broken Session Filter.filter file provided by Adobe in the release package for the Site profiles.

```
entity = derived_filter:  
  formula = string: Visitorized_Flag="1"  
  model = ref: wdata/model
```

By default, workspaces have the Broken Session filter applied to both their selection and their benchmarks, and it can be toggled by right-clicking in the workspace and clicking Data > Broken Session Filter.

The Broken Session Filter can be referenced in filter expressions as Broken_Session_Filter, even if it is not enabled for the current workspace. See [filter expressions](#) for additional information.

Baseline Profiles

A set of standard profiles for each application has been created to allow for one or more of the profiles to be installed at any given time.

These baseline profiles include the definitions of standard metrics, dimensions, filters, reports, workspaces, and dashboards. Adobe continuously updates these profiles and makes them available to its licensees through its Software Support program. Additionally, users of Adobe applications can define additional profiles and use them with or instead of the profiles provided by Adobe.

The profile management system allows the overriding of a profile's configurations by higher level profiles. This feature should be used by users or administrators who want to override any of the definitions incorporated in these profiles after they are installed. Installing new releases of these profiles will overwrite any prior versions. As a result, any changes made to these profiles directly would be overwritten when the new profiles are installed.

Profiles for vertical markets or specific types of companies or industries also may be available from Adobe ClientCare. The following are descriptions of the baseline profiles:

- The **Base profile** includes configuration files that ship with Insight Server. The Base profile should not be modified by the user or Administrator. Any alterations made in the Base profile are subject to being overwritten when Adobe releases a next version of Insight Server or other software application.
- The **Traffic profile** includes a set of fundamental metrics, dimensions, and filters for web analytics. It also includes template workspaces, reports, and dashboards that facilitate the analysis of, reporting on, and general understanding of overall Internet site activity trends and patterns. This profile functions "out-of the box" with a baseline installation of Site.
- The **Value profile** includes a set of metrics and dimensions and template workspaces, reports, and dashboards associated with the built-in Site business value and conversion model. This profile allows users to identify value events on the site and associate a monetary value to those events.

This profile expands your site analysis capabilities by providing a Business Value Model, an advanced method for measuring and tracking the amount of value generated by your site. Value events and their relative value are defined through a simple drag and drop interface within Site. Site uses these definitions to calculate the business value generated by each session, and this information is used in turn to define metrics such as value, value events, conversion, and so on. These metrics enables you to answer questions such as:

-
- What is the most profitable path through the site?
 - What referrer or campaign has generated the most value?
 - What is the average number of purchases on the site per day? (How many value events occur per day on average?)

After you have defined a Business Value Model in Site, you can use the Value metrics and dimensions in your analysis.

- The **Marketing profile** includes set of metrics and dimensions and template workspaces, reports, and dashboards associated with analysis of Internet marketing campaigns, including search analysis and extended referrer analysis.

Adobe also provides the following optional profiles to use:

- The **IP Geo-location profile** includes dimensions and layer files related to the analysis of visitors' locations based on IP geo-location data provided to Adobe by Quova, Inc., and incorporated into data workbench.
- The IP Geo-intelligence profile includes dimensions and layer files related to the analysis of visitor's locations based on IP geo-intelligence data provided to Adobe by Digital Envoy, Inc., and incorporated into data workbench.

For information about the IP-Geo-location and IP Geo-intelligence profiles, contact your Adobe support personnel. The following sections describe the metrics and dimensions defined in each of the baseline profiles.

Traffic Profile Metrics

The Traffic profile contains the following metrics to identify visitor traffic.

Metric	Metric Formula and Level	Description
Entries	Formula: <code>Page_VIEWS[no shift(None, Page_View, Session, -1)]</code> Level: Page View	The number of sessions that entered the site on each page. This metric is evaluated over the Page dimension only.
Exit Rate	Formula: <code>Exits/Page_VIEWS</code> Level: Page View	The percentage of sessions that exited the site from each page. The Exit Rate metric can only be evaluated over the page dimension.
Exits	Formula: <code>Page_VIEWS[no shift(None, Page_View, Session, 1)]</code> Level: Page View	The number of sessions that exited the site from each page. This metric is evaluated over the Page dimension only.
LVCI90	Formula: $(\text{raw(Visitors)} - ((\text{raw(Visitors}) + .69)^{0.5} * 1.281551 - 1.2269)) * (\text{Visitors}/\text{raw(Visitors)})$ Level: Visitor	A measure of the lowest number of possible visitors as reported by Insight. Mathematically, it specifies the lowest number of visitors with a 90% probability.
Page View Duration	Formula: <code>sum(exact_page_duration, page_view)*0.1/Page_VIEWS[any Exact_Page_Duration]</code> Level: Page View	The average length of time (MM:SS) spent on a particular page or group of pages. This metric is evaluated over the Page dimension only.
Page Views per Session	Formula: <code>Page_VIEWS / (Sessions by Page_View)</code>	The average number of page views in each session that has page views.

Metric	Metric Formula and Level	Description
	Level: Session	
Page Views	Formula: sum(One, Page_View) Level: Page View	The number of page views. A page view is a request for a defined page (access to images and other types of filtered content are not counted).
Pct of Page Views	Formula: $\text{Page_Views} / \text{total}(\text{Page_Views})$ Level: Page View	The percentage of page views.
Pct of Sessions	Formula: $\text{Sessions} / \text{total}(\text{Sessions})$ Level: Session	The percentage of sessions.
Pct of Visitors	Formula: $\text{Visitors} / \text{total}(\text{Visitors})$ Level: Visitor	The percentage of visitors.
Pct Referred Visitors	Formula: Referred_Visitors/Visitors Level: Visitor	The percentage of visitors that were referred to this site from another site.
Referred Sessions	Formula: Sessions[Referrer<>'None' and Referrer<>'bookmarks'] Level: Session	The number of sessions that were referred to this site from another site.
Referred Visitors	Formula: Visitors[Visitor_Referrer<>'None' and Visitor_Referrer<>'bookmarks'] Level: Visitor	The number of visitors that were referred to this site from another site.
Retention	Formula: Sessions[shift(None, Session, Visitor, 1) = ""] / Sessions Level: Session	The percentage of sessions that are not the visitors last sessions.
Session Duration	Formula: (sum(Exact_Page_Duration, Session)*.1/Sessions)[Session_Duration <= '01:00:00'] Level: Session	The average length of time (MM:SS) a visitor spends in a session.  Note: You can use this metric with the Segment Export feature.
Sessions by Page View	Formula: Sessions by Page_View Level: Session	The number of sessions that had a page view.
Sessions	Formula: sum(One, Session) Level: Session	A count of visitor sessions. A session is a period of activity for one visitor of a web site. Individual sessions for each visitor are identified using cookies, timeouts, and other heuristics.

Metric	Metric Formula and Level	Description
SCI80	Formula: confidence(Sessions) * 1.281551 / Sessions Level: Visitor	A measure of confidence of the Sessions metric as reported by data workbench. Mathematically, it is a +/- percentage specifying the range within which the actual answer will lie 80% of the time. As a rule-of-thumb, doubling the SCI80 percentage will give a range within which the actual answer will lie 99% of the time.
UVCI90	Formula: (((raw(Visitors) + .68)^0.5 * 1.281551 + 1.2269) + raw(Visitors)) * (Visitors/raw(Visitors)) Level: Visitor	A measure of the highest number of possible visitors as reported by Insight. Mathematically, it specifies the highest number of visitors with a 90% probability.
VCI80	Formula: ((raw(Visitors) + .68)^0.5 * 1.281551 + 1.2269) / raw(Visitors) Level: Visitor	A measure of confidence of the Visitors metric as reported by Insight. Mathematically, it is a +/- percentage specifying the range within which the actual answer will lie 80% of the time. As a rule-of-thumb, doubling the VCI80 percentage will give a range within which the actual answer will lie 99% of the time.
Visitors by Page View	Formula: Visitors by Page_View Level: Page View	The number of visitors that had a page view.
Visitors by Session	Formula: Visitors by Session Level: Session	The number of visitors that had a session.

Traffic Profile Dimensions

The Traffic profile contains the following dimensions to help identify visitor actions.

The dimensions in the following table are defined in the transformation dataset include files in the Traffic\Dataset\Transformation directory.

Name	Type	Level	Description
Day	Simple	Session	The day of the first log entry of the session.
Day of Week	Simple	Session	The day of week of the first log entry of a session.
Exact Page Duration (Hidden)	Numeric	Page View	The duration in milliseconds of page view as measured by subtracting the time of the next page view from the time of that page view. The exact duration of the last page view in a session is always 0.
Hour	Simple	Session	The hour of the first log entry of a session.
Hour of Day	Simple	Session	The hour of day of the first log entry of a session.
Month	Simple	Session	The month of the first log entry of a session.
Page View	Countable	Session	A log entry or “Event Data Record” satisfying the Page View Condition.
Referrer	Simple	Session	The second level domain of the referrer of the first log entry of the session (or None if it is an internal referrer domain).
Session	Countable	Visitor	A period of related contiguous activity by a Visitor. It is delimited by a 30-minute period of

Name	Type	Level	Description
			inactivity and an external referrer domain or a maximum 48-hour Session Duration. Both of those timeouts and the set of domains that are considered internal can be configured in the Transformation.cfg file.
URI	Simple	Page View	The URI stem of a page view. The URI dimension can be redefined using the ReplaceURI, PrependURI, and AppendURI transformations.
Visitor	Countable	N/A	A unique value of x-trackingid. Usually corresponds to a unique browser if persistent cookies are used. The x-trackingid can be otherwise based on IP number or some other unique identifier such as x.509 common name.
Week	Simple	Session	The week of the first log entry of a session.

The dimensions in the following table are defined in the Dimension directory of the Traffic profile:

Dimension	Type	Level	Description
Browser	Simple	Visitor	The type of user agent used by the visitor, including the version number (for example, MSIE 6.0.)
Browser Type	Simple	Visitor	The type of user agent used by the visitor (for example, MSIE).
Search Engine	Simple	Session FIRST ROW	The first search engine in a visitor's session when a visitor has searched from a named search engine.
Next URI	Derived (ShiftDim based on URI dimension)	Page View	The URI of the next URI after the currently selected URI. This derived dimension is used to conduct analysis about what visitors do next after any given URI.
Onetime vs Repeat	Derived (SegmentDim based on Repeat Visitors and One-time Visitors metrics)	Visitor	The type of visitor: one-time or repeat. One-time visitors have had only one session on the site, while repeat visitors have had more than one session.
Page	Derived (RenameDim based on URI dimension)	Page View	The name of each page visited during a session. Initially, each page's name is the same as the URI but can be changed for easier interpretation.
Referrer	Inherited from the built-in Referrer dimension	Session	The second level domain name of the web site that first referred a session to the site (as supplied by the visitor's browser.)
Search Phrase	Simple	Session FIRST ROW	The first search phrase in a visitor's session as passed on by a search engine when a visitor has searched from a named search engine.
Search Term	Many-to-Many	Session	Each search term as passed on by a search engine when a visitor has a search referrer click-through from a named search engine.
Session Duration	Derived (MetricDim based on Exact Page Duration metric)	Session	The duration of a session in 30-second increments.
Session Number	Simple	Session	The number of times a visitor has visited the site. For example, first-time visitors have a Session Number of "1", second-time visitors have a Session Number of "2", etc.
Session Page Views	Derived (MetricDim based on Page Views metric)	Session	The number of page views in a session. For example, selecting "3" in the Session Page Views

Dimension	Type	Level	Description
			dimension would select all sessions with exactly 3 page views.
Search Engine	Simple	Session	The second level domain name of the first web site that is a named search engine that referred a visitor to the site during a session (as supplied by the visitor's browser.)
Time Dimensions	Simple	Session	The hour, day, hour of day, week, day of week, month, quarter or year in which a Session began.
Time Reporting Dimensions	Derived (LastN and Rename dimensions based on built-in time dimensions)	Session	Dimensions including Days Ago, Days of Last Month, Days of Last Week, Days of This Month, Days of This Week, Hours of Today, Hours of Yesterday, Last 12 Months, Last 2 Months, Last 2 Weeks, Last 24 Hours, Last 3 Months, Last 4 Weeks, Last 6 Months, Last 7 Days, Last 7 Days and Today, Last 8 Weeks, Last 9 Months, Last Month, Last Week, Months Ago, This Month, This Week, This and Last 14 Days, This and Last 6 Months, This and Last 8 Weeks, Today, Today Reporting, Today and Last 30 Days, Weeks Ago and Yesterday provide a convenient way to build a Workspace or Report which always shows a portion of data in the Dataset. Each is based upon when a session began.
URI	Inherited from the built-in Dimension URI	Page View	The URI of each page viewed.
Visitor Page Views	Derived (MetricDim based on Page Views metric)	Visitor	The number of page views to date for a visitor. For example, selecting "3" in the Visitor Page Views dimension would select all visitors with exactly 3 page views across all of their sessions.
Visitor Referrer	Inherited from the built-in dimension Referrer	Visitor	The second level domain name of the web site that first referred a Visitor to the site for their first session (as supplied by the visitor's browser).

Value Profile Metrics

Metric	Formula	Level	Description
Conversion	Sessions [not Session_Value=�]/Sessions	Session	The percentage of sessions which generated business value (as defined by the Business Value Model).
Pct of Value	Value/total(Value)	Session	The percentage of overall value that was generated from the selected set of sessions.
Value	sum(Session_Value, Session)*0.01	Session	The total business value generated, in dollars (as defined by the Business Value Model).
Value Events	Sessions [not Session_Value=�]	Session	The count of sessions that generated business value (as defined by the Business Value Model).
Value Events per Visitor	(Value_Events/Visitors) by Visitor	Visitor	The average number of sessions for each visitor that generated business value (as defined by the Business Value Model).
Value per Visitor	(Value/Visitors) by Visitor	Visitor	The average business value generated, in dollars, by each visitor.

Value Profile Dimensions

Dimension	Type	Level	Description
Session Value	Derived (MetricDim based on Value metric)	Visitor	The value of the session as defined by the Business Value Model.
Value Segment	Derived (MetricDim based on Value metric)	Visitor	The total value that a visitor has generated in all their sessions. The default increment used for each segment of value is \$50, but can be changed.

Marketing Profile Dimensions

The Campaign dimension is defined in the Site Marketing profile to provide campaign analysis capabilities.

Dimension	Type	Level	Description
Campaign	Renamed Simple	Session FIRST ROW operation	The campaign identifier extracted from a value in the visitor's first request.

Samples of Custom Marketing Profile Dimensions

You can incorporate additional dimensions of data for further analysis. These dimensions are added by incorporating additional information into the stream of data that is collected for analysis. For example, the following table contains some of the custom marketing dimensions that have been added in deployments for customers in various industries:

Dimension (Custom)	Description
Email Campaign Date	Parses the campaign date (first value) from email campaign query strings.
Email Campaign Detail	Collects the value string attached to the email campaign query string variable.
Email Campaign Segment	Parses the campaign segment (third value) from email campaign query strings.
Email Campaign Type	Parses the campaign type (second value) from email campaign query strings.
Marketing Campaign Detail	Collects the value string attached to marketing campaign query string variables.
Marketing Campaign Owner	Parses the campaign owner (fourth value) from marketing campaign query strings.
Marketing Campaign Source	Parses the campaign source (first value) from marketing campaign query strings.
Marketing Campaign Type	Parses the campaign type (second value) from marketing campaign query strings.
PPC Campaign Detail	Collects the value string attached to the ppc query string variable.

Data Workbench Geography Overview

Navigation title: Data Workbench Geography

Data workbench **Geography** is a geographical imagery and analysis software product that enables you to display topographical imagery in data workbench.

For information about installing data workbench**Geography**, see [Installing Data Workbench Geography](#) on page 744.

Data workbench **Geography** consists of the **Geography** profile, which includes default imagery layers and their related lookup files as well as additional dimensions (City, Coordinates, Country, DMA, and Domain) for use in your analysis.

You also can add other layers and lookup files to expand your imagery capabilities. For more information about imagery layers, see [Working with Imagery Layers](#) on page 746.

If you have licensed data workbench**Geography**, you can subscribe to an optional data service that enables you to map IP addresses on the globe or create visualizations showing geographically related dimensions of data (such as area code, time zone, and MSA) that are based on IP address. For information about subscribing to these services, please contact Adobe. For information about installing and working with data services, see [Working with Data Services](#) on page 760.

Installing Data Workbench Geography

Navigation title:Installing Data Workbench Geography

This chapter provides the steps to install data workbench**Geography** on a data workbench server.

To complete these steps, you must have access to the files on the data workbench server.

If you have subscribed to one of the two optional data services, you also must install the data service profile and its lookup files. For more information about these services, see [Working with Data Services](#) on page 760.

For information about the system requirements for installing, configuring, and operating data workbench**Geography**, see the *Minimum System Requirements* document.

Adobe distributes the data workbench**Geography** software product as a .zip file. This file contains two folders (Lookups and Profiles). To install data workbench**Geography** on a data workbench server, you must perform the following steps:

- Install the **Geography** profile. See [Installing the Geography Profile](#) on page 744.
- Install the **Geography** lookups. See [Installing the Geography Lookup Files](#) on page 746.
- (Optional) Install the `Terrain_Images.cfg` file. See [Installing the Terrain Images.cfg File](#) on page 746.



Note: You must install the files for data workbench**Geography** on the data workbench server machine on which you are processing and running your dataset profile. If you are running a data workbench server cluster, you must install the files on the master server. For information about dataset profiles, see the *Dataset Configuration Guide*.

Installing the Geography Profile

Navigation title:Installing the Geography Profile

The **Geography** profile provided with data workbench**Geography** is an internal profile that provides additional functionality to your Adobe application.

As with all other internal profiles provided by Adobe, the **Geography** profile should not be changed. All customization must occur in your dataset or role-specific profiles or other profiles that you create.

The **Geography** profile includes several transformation dataset include files (located in the `Dataset\Transformation\Geography` folder) that define geographical dimensions. Following is a list of the transformation dataset include files provided with the **Geography** profile:

- `City.cfg`
- `Coordinates.cfg`
- `Country.cfg`
- `DMA.cfg`
- `Domain.cfg`

Each of the files is named for the extended dimension that it defines. An additional file, `IPLookup.cfg`, defines several geographical data fields that are used to define dimensions in the other transformation dataset include files.

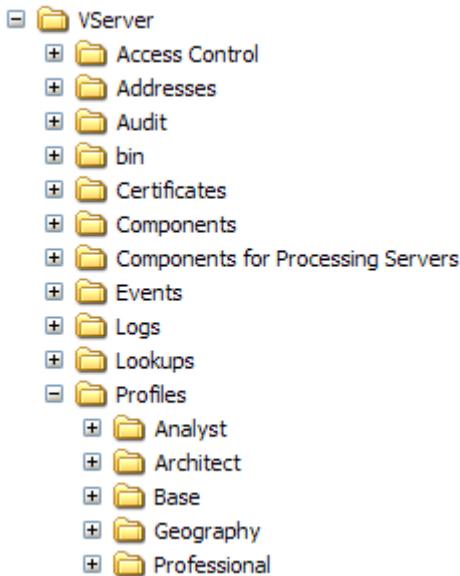
For information about transformation dataset include files, see the *Dataset Configuration Guide*.

To install the Geography profile on the data workbench server



Note: The following installation instructions assume that you have installed data workbench and established a connection between data workbench and the data workbench server on which you are installing data workbench**Geography**. If you have not done so, see the *Data Workbench User Guide*.

1. Open the Profiles folder from the .zip file provided to you by Adobe.
2. Copy the Geography folder to the Profiles folder in your data workbench server installation directory. You want to end up with a ...\\Profiles\\Geography folder on your data workbench server as shown in the following example. The names of the other folders within the Profiles folder may differ from the ones shown.



3. Use the following steps to update the profile.cfg file for each profile with which you want to use data workbench**Geography**.

1. Open the **Profile Manager**.
2. Right-click the check mark next to profile.cfg and click **Make Local**. A check mark for this file appears in the **User** column.
3. Right-click the newly created check mark and click **Open > from the workbench**. The profile.cfg window appears.
4. In the profile.cfg window, right-click **Directories** and click **Add new > Directory**.

To add the new directory to the end of the list of directories, right-click the number or name of the last directory in the list and click **Add new > Directory**.

5. Type the name of the new directory: **Geography** .
6. Right-click (**modified**) at the top of the window and click **Save**.
7. In the **Profile Manager**, right-click the check mark for profile.cfg in the **User** column, then click **Save to > <profile name>**.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe (including the **Geography** profile), as your changes are overwritten when you install updates to these profiles.

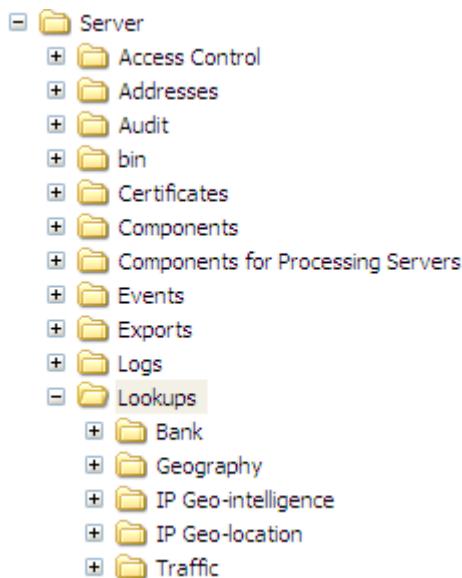
Installing the Geography Lookup Files

Navigation title:Installing the Geography Lookup Files

Steps to install the **Geography** lookup files on the data workbench server.

The lookup file (Lookups\Geography\DMA.txt) provided with the **Geography** profile is a two-column tab-delimited text file that lists DMA (designated market area) codes and the corresponding names.

1. Open the Lookups folder from the .zip file provided to you by Adobe.
2. Copy the Geography folder to the Lookups folder in your data workbench server installation directory. You want to end up with a ...\\Lookups\\Geography folder on your data workbench server as shown in the following example. The names of the other folders within the Lookups folder may differ from the ones shown.



Installing the Terrain Images.cfg File

Navigation title:Installing the Terrain Images.cfg File

Steps to install the Terrain Images.cfg file on the data workbench server.



Note: You need to install the Terrain Images.cfg file only if you want to make new terrain image layers to display on the globe visualization.

1. Copy the Terrain Images.cfg file from the .zip file provided to you by Adobe.
2. Paste the file in the Components folder in your data workbench server installation directory.

For information about configuring Terrain Images.cfg, see [Working with Terrain Image Layers](#) on page 748.

Working with Imagery Layers

Navigation title:Working with Imagery Layers

Conceptual information about creating and implementing layers.

Understanding Imagery Layers

Navigation title:Understanding Imagery Layers

Conceptual Information about the **Geography** profile layers, types of imagery layers, and creating new layers.

Types of Imagery Layers

Data workbench **Geography** enables you to view the following types of imagery layers in data workbench:

- **Terrain image layer:** This type of layer displays terrain imagery of the Earth, over which geographical data can be displayed. The globe visualization in data workbench is an example of a terrain image layer. See [Working with Terrain Image Layers](#) on page 748.
- **Element point layer:** This type of layer displays one point on the globe for each element of a dimension. See [Working with Element Point Layers](#) on page 754.
- **Vector layer:** This type of layer displays vector data (line art) on the globe. See [Working with Vector Layers](#) on page 757.

In data workbench, you can select which of these layers you want to display for a particular analysis task.

Geography Profile Layers

The **Geography** profile provides you with a set of default imagery layers, which are stored in the Profiles\Geography\Maps folder within the data workbench server installation directory:

- **Blue Marble 2km:** This terrain image layer creates a 3-D map of the world, which is what displays when you add the globe visualization to a workspace. When this layer is not selected, the globe is not visible but the other layers still display. The Blue Marble 2km.layer file references the Blue Marble 2km.tsi file.

For information about working with the globe visualization, see the *Data Workbench User Guide*.

- **Zip Points:** This element point layer enables you to map locations in your dataset using a United States ZIP Code. The Zip Points.txt lookup file (provided by Adobe) contains a list of all United States ZIP Codes and each ZIP Code's latitude and longitude. The Zip Points.layer file references the Zip Points.txt file and the Zipcode.dim file and contains the configuration parameters needed to display the locations on the globe. Each element of the ZIP Code dimension (Zipcode.dim) that you define within your dataset is mapped on the globe using the latitude and longitude listed for that ZIP Code in the Zip Points.txt lookup file.

For information about defining dimensions, see the *Dataset Configuration Guide*.

- **Boundaries:** This vector layer provides the major world political boundaries, such as countries, as well as the boundaries of natural physical features of the Earth, such as lakes and islands. The Boundaries.layer file references one or more of the mwcoast.vec, mwisland.vec, mwlake.vec, mwnation.vec, mwriver.vec, mwstate.vec, US states.vec, and world boundaries.vec files.
- **IP Coordinates:** This element point layer uses dynamic points to enable you to map locations in your dataset using IP addresses. The IP Coordinates.layer file references the Coordinates dimension (Coordinates.dim) and specifies the Visitors metric as the metric to use to determine the size of the points on the globe for each coordinate.

Your **Geography** profile or other profiles in your installation may contain additional imagery layers that Adobe provided or your company created.

Creating New Layers

You can create new imagery layers by copying the appropriate type of layer file included in the **Geography** profile into any Profiles\profile name\Maps folder, then renaming and editing the file as appropriate. All new layers must meet the following requirements:

- The .layer file must adhere to the format of one of the supported layer types.
- The .layer file must reference the appropriate lookup and dimension files, if necessary.
- The referenced lookup file also must be stored within the data workbench server installation directory, and its path must be specified accurately in the .layer file.

For more information about the format and parameters for each type of layer file and its associated files, see the section in this chapter for the appropriate layer type.

Working with Terrain Image Layers

Navigation title: Working with Terrain Image Layers

In data workbench, a terrain image layer displays terrain imagery of the Earth.

Terrain image layers are stored in the **Geography** profile, in a custom format. These image layers can be generated by Adobe, or data workbench server can transform your user-supplied terrain imagery into terrain layers suitable for use on the globe visualization.



Note: To work with terrain image layers, you must install the `Terrain_Images.cfg` file provided by Adobe. For installation instructions, see [Installing Data Workbench Geography](#) on page 744.

To define a terrain image layer, you must have the following:

- **One or more terrain image files** containing the images to be displayed on the globe.
- **A `Terrain_Images.cfg` file** that specifies the terrain image file(s) to be used for the layer(s). The `Terrain_Images.cfg` file enables you to add one or more sources to create a terrain image layer. The format of your terrain image file determines the type of source that you should add. The following table provides descriptions of the available terrain image layer sources, including the supported terrain image file formats:

Terrain Image Layer Sources

Type	Description
Raw unprojected bitmap	<p>Creates terrain image layers from 24-bit headerless RGB files that are latitude-longitude aligned (unprojected), where north is the top of the image, and east is the right.</p> <p>Supported image format(s): RAW</p> <p> Note: This source requires projection information. For information about projection formats, see Specifying Projection Information for Terrain Images on page 751.</p>
General image, unprojected	<p>Creates terrain image layers from 24-bit, latitude-longitude aligned (unprojected) image formats, where north is the top of the image, and east is the right.</p> <p>Supported image format(s): BMP, JPG, PNG, TIFF</p> <p> Note: This source requires projection information. For information about projection formats, see Specifying Projection Information for Terrain Images on page 751.</p>
Image with embedded projection	<p>Creates terrain image layers from image formats that embed geodetic data in the image file. The projection information is extracted from the image.</p> <p>Supported image format(s): Erdas (IMG), GeoTIFF</p> <p> Note: This source usually does not require projection information but supports the addition of such information if needed. For information about projection formats, see Specifying Projection Information for Terrain Images on page 751.</p>

To define a terrain image layer

1. In data workbench, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the **Servers Manager** workspace.
2. Within the **Servers Manager** window, right-click the icon of the desired data workbench server and click **Server Files**.

-
3. In the **Server Files Manager**, click **Components** to view its contents. The `Terrain Images.cfg` file is located within this directory.
 4. Right-click the check mark in the server name column for `Terrain Images.cfg`, then click **Make Local**. A check mark appears in the **Temp** column for `Terrain Images.cfg`.
 5. Right-click the newly created check mark in the **Temp** column and click **Open > from the workbench**. The `Terrain Images.cfg` window appears.
 6. In the **Terrain Images** window, click **component** to view its contents.
 7. Right-click **Sources** > **Add new** and choose one of the following source types:
 - **Raw unprojected bitmap**. (Once added, this source type is labeled `RawTerrainSource` in the **Terrain Images** window.)
 - **General image, unprojected**. (Once added, this source type is labeled `GDALTerrainSource` in the **Terrain Images** window.)
 - **Image with embedded projection**. (Once added, this source type is labeled `GDALTerrainSource` in the **Terrain Images** window.)
 8. Edit the parameters for the source as necessary using the following sample file and table of parameters as guides.

-Terrain Images.cfg (modified)	
-component	TerrainImageBuilder
-Sources	
-0	RawTerrainSource
-Projection Info	LatLonProjection
Lat0	90
Lat1	-90
Lon0	-180
Lon1	180
Gamma	0.75
Height	10800
Source Image	earth.raw
Tile Compression Quality	80
Tile Compressor	JPEG
Width	21600
-1	GDALTerrainSource
-Projection Info	LatLonProjection
Lat0	90
Lat1	-90
Lon0	-180
Lon1	180
Gamma	1.0
Source Image	earth.png
Tile Compression Quality	100
Tile Compressor	
-2	GDALTerrainSource
Gamma	1.6
Source Image	2007-*.*.img
Tile Compression Quality	75
Tile Compressor	JPEG
Source Image Location	MapSources\
Temp Image Storage	Temp\
Write Layers To	Profiles\Geography\Maps\

Terrain Images.cfg: Sources

Parameter	Description
Gamma	Optional for all sources. Specifies gamma correction to be applied to the source image. This may be desirable due to the fact that data workbench normally runs with a high gamma setting. The default value is 1.
Height	Required for raw unprojected bitmap images. The height of the source image in pixels.
Projection Info	Required for raw unprojected bitmap images and general images, unprojected, but supported for images with embedded projection. Data workbenchGeography supports latitude-longitude projections and Transverse Mercator (TM) projections for terrain image layers. The default projection format is the latitude-longitude projection (LatLonProjection). For information about projection formats, see Specifying Projection Information for Terrain Images on page 751.

Parameter	Description
Source Image	Required for all sources. The name of the source image file. This may be a file name or a wildcard pattern. Using a pattern may be useful if, for instance, images for the same region at different dates are uploaded, with no change in the associated metadata. Therefore a pattern like "Tysons Corner *.raw" would create layers from Tysons Corner 050211.raw, Tysons Corner 050218.raw, and so on as new images are added, with no additional configuration necessary if the parameters for the files are otherwise identical.
Tile Compression Quality	Optional for all sources. For JPEG compression, an integer from 0 to 100 specifying how to balance image size and quality. (The default value is zero.) A higher number results in better image quality, but produces larger images and longer download times for data workbench users. Compressing images below 70 may result in image degradation.
Tile Compressor	Optional for all sources. Specifies what compression method is used to write output files. The only currently supported methods are RAWRGB (the default, resulting in no compression) and JPEG. Use JPEG compression to reduce the size of layers that are transmitted during profile synchronization.
Width	Required for raw unprojected bitmap images. The width of the source image in pixels.

9. Edit the Source Image Location, Temp Image Storage, and Write Layers To parameters using the following table as a guide. These parameters apply to all of the terrain image sources that you define in the Sources section of this file.

Terrain Images.cfg: Additional Parameters

Parameter	Description
Source Image Location	Required. The directory that is scanned for images to translate into terrain layers. If it is not an absolute path, it is interpreted relative to the data workbench server installation directory.
Temp Image Storage	Optional. The name of a directory that is used for storage of temporary files used in the translation of source images to terrain layers. If it is not an absolute path, it is interpreted relative to the data workbench server installation directory. The default location is the Temp directory.
Write Layers To	Required. The directory to which terrain layers are output. Ordinarily, this is the Maps subdirectory of a profile directory, so that the Globe visualization can find the layers.

10. Save the file by right-clicking (**modified**) at the top of the window and clicking **Save**.
 11. To save updated file to the data workbench server machine, in the **Server Files Manager**, right-click the check mark for **Terrain Images.cfg** in the **Temp** column, then click **Save to > <server name>**.

Specifying Projection Information for Terrain Images

Navigation title:Specifying Projection Information for Terrain Images

Data workbench **Geography** supports both latitude-longitude projections and Universal Transverse Mercator (UTM) projections for all terrain image layer sources.

Projection information is required for raw unprojected bitmaps and general images, unprojected. You can specify projection information for images with embedded projection information, though it is usually not required because the parameters of the projection are determined automatically from geodetic data embedded in the image itself. The following sections provide details about specifying these projection formats in the **Terrain Images.cfg** file.

Latitude-Longitude Projections

Navigation title:Latitude-Longitude Projections

The latitude-longitude projection format (LatLonProjection) in the **Terrain Images.cfg** file is defined by four parameters for latitude and longitude.

To specify a LatLonProjection for unprojected images (raw unprojected bitmaps and general images, unprojected), you can enter settings for the LatLonProjection within the **Terrain Images.cfg** window in data workbench. See [To Specify a LatLonProjection for Unprojected Images](#) on page 752.

To specify a LatLonProjection for images with embedded projection information, you must open the `Terrain Images.cfg` file in a text editor such as Notepad, set the Projection Info parameter to “LatLonProjection”, and add settings for the LatLonProjection. See [To Specify a LatLonProjection for Images Within Embedded Projection Information](#) on page 752.

- [To Specify a LatLonProjection for Unprojected Images](#) on page 752
- [To Specify a LatLonProjection for Images Within Embedded Projection Information](#) on page 752

To Specify a LatLonProjection for Unprojected Images

1. Open the `Terrain Images.cfg` file in data workbench and add a terrain image layer source as described in [Working with Terrain Image Layers](#) on page 748.
2. Edit the Projection Info parameters using the following parameters table as a guide:

LatLonProjection Parameters

Parameter	Description
Lat0	The latitude of the top edge of the image, in degrees, where 90 is the North Pole and -90 is the South Pole.
Lat1	The latitude of the bottom edge of the image.
Lon0	The longitude of the left-hand edge of the image, in degrees, where positive numbers are east and negative numbers are west longitudes.
Lon1	The longitude of the right-hand edge of the image.

3. Save the file by right-clicking (**modified**) at the top of the window and clicking **Save**.
4. To save the locally made changes to the data workbench server machine, in the **Server Files Manager**, right-click the check mark for `Terrain Images.cfg` in the **Temp** column, then click **Save to > <server name>**.

To Specify a LatLonProjection for Images Within Embedded Projection Information

1. In the **Server Files Manager**, click **Components** to view its contents. The `Terrain Images.cfg` file is located within this directory.
2. Right-click the check mark in the server name column for `Terrain Images.cfg`, then click **Make Local**. A check mark appears in the **Temp** column for `Terrain Images.cfg`.
3. Right-click the newly created check mark in the **Temp** column and click **Open > in Notepad**. The `Terrain Images.cfg` file appears in a Notepad window.
4. Edit the Projection Info parameters using the following sample file fragment as a guide. Be sure to specify the projection type as highlighted below. For descriptions of the parameters, see the LatLonProjection Parameters table in the previous procedure.

```
Projection Info = LatLonProjection:  
    Lat0 = double: 90  
    Lat1 = double: -90  
    Lon0 = double: -180  
    Lon1 = double: 180
```

Universal Transverse Mercator Projections

Navigation title:Universal Transverse Mercator Projections

The Universal Transverse Mercator (UTM) projection is defined by eight parameters.

When specifying a Universal Transverse Mercator projection for a terrain image layer, your terrain image files must be aligned with false (projected) north towards the top of the image, and false east to the right of the image.

To specify a UTM projection for any terrain image source, you must open the `Terrain_Images.cfg` file in a text editor such as Notepad, set the `Projection Info` parameter to “`TransverseMercatorProjection`”, and add settings for the UTM projection.

To specify a Universal Transverse Mercator projection

1. In the **Server Files Manager**, click **Components** to view its contents. The `Terrain_Images.cfg` file is located within this directory.
2. Right-click the check mark in the *server name* column for `Terrain_Images.cfg`, then click **Make Local**. A check mark appears in the **Temp** column for `Terrain_Images.cfg`.
3. Right-click the newly created check mark in the **Temp** column and click **Open > in Notepad**. The `Terrain_Images.cfg` file appears in a Notepad window.
4. Edit the `Projection Info` parameters using the following sample file fragment and parameters table as guides. Be sure to specify the projection type as highlighted below.

```
Projection Info = TransverseMercatorProjection:  
Ellipsoid Inverse Flattening = double: 294.9786982139006  
Ellipsoid Semimajor Axis = double: 6378206.4000000004  
False Easting = double: 500000  
False Northing = double: 0  
Northwest Corner Coordinates = v3d: (550339, 5.42059e+006, 0)  
Prime Meridian = double: -123  
Scale Factor = double: 0.9996  
Southeast Corner Coordinates = v3d: (555099, 5.41356e+006, 0)
```

TransverseMercatorProjection Parameters

Parameter	Description
Ellipsoid Inverse Flattening, Ellipsoid Semimajor Axis	The parameters of the ellipsoid used for the projection. The semimajor axis is specified in meters.
False Easting	The false easting of the central meridian of the projection, in meters. For UTM, this is always 500,000.
False Northing	The false northing of the equator in the projection, in meters. For UTM, this is 0 for northern hemisphere zones and 10,000 for southern hemisphere zones.
Northwest Corner Coordinates, Southeast Corner Coordinates	The coordinates (in projected meters) of the top left and bottom right corners of the image.
Prime Meridian	The longitude of the central meridian of the projection, specified in degrees east of Greenwich. Negative numbers may be used to specify degrees west.
Scale Factor	The ratio of the radius of the projection cylinder to the semimajor axis of the ellipsoid. For Universal Transverse Mercator (UTM) projections, this is always 0.9996.

Making a New Terrain Image Layer Available

Navigation title: Making a New Terrain Image Layer Available

Steps to make a new terrain layer available to display on the globe visualization.

1. In the `Profiles\profile name\Maps` folder within the **Insight Server** installation directory, place the layer file and the supporting image files.
2. Edit the `order.txt` file in the `Profiles\profile name\Maps` folder to reflect the order in which you want the layers to display. By default, layers appear in lexicographic order by their names.



Note: When editing the `order.txt` file, take care not to cover up map layers that you want to show.

For more information about using `order.txt` files, see the Configuring Interface and Analysis Features chapter of the *Data Workbench User Guide*.

-
3. In data workbench, select the desired profile by right-clicking the workspace title bar and clicking **Switch Profile** > *<profile name>*.
 4. Right-click the workspace title bar and click **Work Online**. An X appears next to **Work Online**.
 5. Open a workspace and on a globe visualization, right-click and select the new layer. An X appears next to the layer name.

Working with Element Point Layers

Navigation title:Working with Element Point Layers

In data workbench, an element point layer displays one point on the globe for each element of a dimension.

The size of the point displayed on the globe is determined by evaluating a specified metric over the dimension whose elements you want to plot. Therefore, larger metric values correspond to larger points on the globe.

Defining Element Point Layers Referencing Lookup Files

Navigation title:Defining Element Point Layers Referencing Lookup Files

When creating an element point layer that references a lookup file to obtain latitude and longitude data, the location of the point is obtained by retrieving each element and its associated latitude and longitude from the lookup file.

Instead of using a lookup file, you can use the **Dynamic Points** functionality, which embeds the latitude and longitude of a location in the name of each element of a dimension. See [Defining Element Point Layers Using Dynamic Points](#) on page 756.

To define an element point layer that references a lookup file, you must create or already have available the following:

- **A dimension** defined in the `Transformation.cfg` file or a transformation dataset include file. For information about transformation configuration files, see the *Dataset Configuration Guide*.
- **A lookup file** containing the data used to plot each data point. This file must contain at least three columns of data for each data point: the key, the longitude, and the latitude. For more information about the required format of the lookup file, see [Element Point Lookup File Format](#) on page 754.
- **A layer file** that specifies the location of the lookup file and identifies the related dimension and metric as well as the key, longitude, and latitude column names in the lookup file. For more information about the required format of the layer file, see [Element Point Layer File Format](#) on page 755.



Note: The `Zip Points.layer` file, provided with the **Geography** profile, is an element point layer that identifies the `Zipcode.dim` file, the `Sessions.metric` file, the `Zip Points.txt` lookup file, and the names of the key, longitude, latitude, and name columns in the lookup file.

Element Point Lookup File Format

Navigation title:Element Point Lookup File Format

Information about the element point layer columns.

The element point layer lookup file must contain at least the following three columns:

- **Key column:** This column should contain common key data, which enables the data workbench server to connect the data in the lookup file to that in the dataset. The **Key** column must be the first column in the lookup file. Each row in this column identifies an element of the dimension.
- **Longitude column:** This column should contain the longitude for each data point in the **Key** column.
- **Latitude column:** This column should contain the latitude for each data point in the **Key** column.
- **Name column:** (Optional). If you want to specify a name to be displayed on the map for each data point, you can include a **Name** column in the lookup file.

Each row in the `Zip Points.txt` lookup file contains a ZIP Code in the first column followed by the longitude, latitude, and associated city name.

```
tude, and associated city name.  
ZIP_CODE LATITUDE LONGITUDE NAME  
00210 +43.005895 -071.013202 PORTSMOUTH, NH  
00211 +43.005895 -071.013202 PORTSMOUTH, NH  
...
```

Element Point Layer File Format

Navigation title:Element Point Layer File Format

Formatting information about the element point layer file.

Each element point layer `.layer` file that references a lookup file must be formatted using the following template:

```
Layer = ElementPointLayer:  
  Data Paths = vector: 1 items  
    0 = Path: Maps\\Lookup File Name.txt  
  Longitude Column = string: Longitude Column Name  
  Latitude Column = string: Latitude Column Name  
  Name Column = string: Location Column Name  
  Key Column = string: Key Column Name  
  Dimension = ref: wdata/model/dim/Dimension Name  
  Metric = ref: wdata/model/metric/Metric Name  
  Scale = double: Scale  
  Color = v3d: RGB Color Vector  
  Rendering Mode = int: Mode Number
```

Element Point Layer Parameters: Lookup Files

Parameter	Description
Data Paths	Path to the lookup file containing latitude and longitude data.
Longitude Column	The name of the column in the lookup file containing the longitude data.
Latitude Column	The name of the column in the lookup file containing the latitude data.
Name Column	Optional. The name of the column in the lookup file containing the names of the locations represented by the latitude and longitude data.
Key Column	The name of the column in the lookup file containing the common key data, which enables the data workbench server to integrate the data in the lookup file into the dataset. This must be the first column in the lookup file. Each row in this column is an element of a dimension. This dimension must be defined in the <code>Transformation.cfg</code> file or a transformation dataset include file and specified in the Dimension parameter of this file. For more information about transformation configuration files, see the <i>Dataset Configuration Guide</i> .
Dimension	The name of the dimension (defined in a transformation configuration file) containing elements that correspond to the data rows in the Key column.
Metric	The name of the metric that is evaluated over the dimension specified in the Dimension parameter.
Scale	Optional. Value used to size the points in the layer. The default value is 100. Larger values make the points bigger, and smaller values make them smaller.
Color	Optional. The RGB color vector, which is expressed as (red,green,blue). For each color in the vector, you can enter a value from 0.0 to 1.0. For example, (1.0, 0.0, 0.0) is bright red, and (0.5, 0.5, 0.5) is gray.
Rendering Mode	Optional. Integer value representing the rendering mode to use for the layer. The three available modes are as follows: <ul style="list-style-type: none">• Rendering Mode 1. Points size is defined in screen space (points stay a constant size relative to the computer screen). Points are rendered using polygons, so there is no upper limit on point size. This is the default rendering mode.

Parameter	Description
	<ul style="list-style-type: none"> Rendering Mode 2. Point size is defined in world space (points stay a constant size relative to the globe). Points are rendered using polygons, so there is no upper limit on point size. Rendering Mode 3. Point size is defined in screen space. Points are rendered using OpenGL smooth points.

The Zip Points.layer file is formatted as follows:

```
Layer = ElementPointLayer:
  Data Paths = vector: 1 items
    0 = Path: Maps\\Zip Points.txt
  Longitude Column = string: LONGITUDE
  Latitude Column = string: LATITUDE
  Name Column = string: NAME
  Key Column = string: ZIP_CODE
  Dimension = ref: wdata/model/dim/Zipcode
  Metric = ref: wdata/model/metric/Sessions
```

Defining Element Point Layers Using Dynamic Points

Navigation title:Defining Element Point Layers Using Dynamic Points

When creating an element point layer using dynamic points, the latitude and longitude data is embedded in each element of the dimension.

To define an element point layer using dynamic points, you must create or already have available the following:

- A **dimension**, defined in the Transformation.cfg file or a transformation dataset include file, in which each element contains the string “latitude,longitude” or “latitude,longitude,name.”

For steps to create a dimension, see the *Dataset Configuration Guide*.

- A **layer file** that specifies the related dimension.

For more information about the required format of the layer file, see [Element Point Layer File Format](#) on page 755.



Note: When using **Dynamic Points**, it is essential to ensure that the cardinality of the dimension specified in the layer file is reasonable. If every row of a dataset has a different latitude and longitude, the dimension quickly fills up and most rows fall into a Small Elements element. Because the Small Elements element does not have a latitude and longitude, it does not appear on the globe.

Element Point Layer File Format

Each element point layer file using dynamic points must be formatted using the following template:

```
Layer = ElementPointLayer:
  Dimension = ref: wdata/model/dim/Dimension Name
  Metric = ref: wdata/model/metric/Metric Name
  Dynamic Points = bool: true
  Scale = double: Scale
  Color = v3d: RGB Color Vector
  Rendering Mode = int: Mode Number
```

Element Point Layer Parameters: Dynamic Points

Parameter	Description
Dimension	<p>The name of the dimension (defined in a transformation configuration file), which must contain elements with the string “latitude,longitude” or “latitude,longitude,name” as shown in the following examples:</p> <ul style="list-style-type: none"> • 37.5181,-77.1903

Parameter	Description
	<ul style="list-style-type: none"> • 35.3317,-77.8126,Somewhere
Metric	The name of the metric that is evaluated over the dimension specified in the Dimension parameter.
Dynamic Points	Enables Dynamic Points. Set to true.
Scale	Optional. Value used to size the points in the layer. The default value is 100. Larger values make the points bigger, and smaller values make them smaller.
Color	Optional. The RGB color vector, which is expressed as (red,green,blue). For each color in the vector, you can enter a value from 0.0 to 1.0. For example, (1.0, 0.0, 0.0) is bright red, and (0.5, 0.5, 0.5) is gray.
Rendering Mode	<p>Optional. Integer value representing the rendering mode to use for the layer. The three available modes are as follows:</p> <ul style="list-style-type: none"> • Rendering Mode 1. Points size is defined in screen space (points stay a constant size relative to the computer screen). Points are rendered using polygons, so there is no upper limit on point size. This is the default rendering mode. • Rendering Mode 2. Point size is defined in world space (points stay a constant size relative to the globe). Points are rendered using polygons, so there is no upper limit on point size. • Rendering Mode 3. Point size is defined in screen space. Points are rendered using OpenGL smooth points.

The IP Coordinates.layer file is formatted as follows:

```
Layer = ElementPointLayer:
  Dimension = ref: wdata/model/dim/Coordinates
  Metric = ref: wdata/model/metric/Visitors
  Dynamic Points = bool: true
```

Making a New Element Point Layer Available

Navigation title: Making a New Element Point Layer Available

Steps to make an element point layer available to display on the globe visualization.

1. In the Profiles\profile name\Maps folder within the data workbench server installation directory, place the layer file and its related lookup file.
2. If you defined a new dimension for your element point layer and have not yet retransformed your dataset, retransform your dataset now.
3. Edit the order.txt file in the Profiles\profile name\Maps folder to reflect the order in which you want the layers to display. By default, layers appear in lexicographic order by their names.



Note: When editing the order.txt file, take care not to cover up map layers that you want to show.

For more information about using order.txt files, see the Configuring Interface and Analysis Features chapter of the *Data Workbench User Guide*.

4. In data workbench, select the desired profile by right-clicking the workspace title bar and clicking **Switch Profile > <profile name>**.
5. Right-click the workspace title bar and click **Work Online**. An X appears next to **Work Online**.
6. Open a workspace and on a globe visualization, right-click and select the new layer. An X appears next to the layer name.

Working with Vector Layers

Navigation title: Working with Vector Layers

In data workbench, a vector layer displays static vector data (line art) on the globe.

You can use vector layers to display outlines of geographic regions, such as states, countries, islands, or bodies of water.

Defining Vector Layers Referencing Vector Files

You can create a vector layer that references one or more vector (.vec) files, which contains the data that defines the vectors to be drawn on the globe.

To define a vector layer that references one or more .vec files, you must have the following:

- One or more .vec files that contain the data used to draw the vectors on the globe.



Note: To obtain .vec files to use with your vector layers, contact Adobe.

- A layer file that specifies the location of the .vec files. For more information about the required format of the layer file, see [Vector Layer File Format](#) on page 758.



Note: The `Boundaries.layer` file, provided with the **Geography** profile, is a vector layer that references the `mwnation.vec`, `mwstate.vec`, `mwcoast.vec`, `mwlake.vec`, and `mwisland.vec` files.

Vector Layer File Format

Each vector layer file referencing .vec files must be formatted using the following template:

```
Layer = VectorLayer:  
  Vec Files = vector< string >  
    0 = string: Maps\\\.vec file 1  
    1 = string: Maps\\\.vec file 2  
    . . .  
    n-1 = string: Maps\\\.vec file n  
  Color = v3d: color vector  
  Alpha = double: alpha  
  Width = double: width  
  Error Factor = double: error factor
```

Vector Layer Parameters: .vec Files

Parameter	Description
Vec Files	Path(s) to the .vec file(s) containing the vector data.
Color	The RGB color vector, which is expressed as (red,green,blue). For each color in the vector, you can enter a value from 0.0 to 1.0. For example, (1.0, 0.0, 0.0) is bright red, and (0.5, 0.5, 0.5) is gray.
Alpha	Controls the transparency of the vectors shown on the globe. The range is 0 to 1, with 0 being the most transparent.
Width	Optional. Sets the width of the data in pixels. The recommended range is 1 to 4.
Error Factor	Controls how accurately the vectors are drawn. For larger values, the vectors are drawn less accurately but faster. The default value is 5.

The `Boundaries.layer` file is formatted as follows:

```
Boundaries.layer file is formatted as follows:  
Layer = VectorLayer:  
  Vec Files = vector< string >  
    0 = string: Maps\\mwnation.vec  
    1 = string: Maps\\mwstate.vec  
    2 = string: Maps\\mwcoast.vec  
    3 = string: Maps\\mwlake.vec
```

```
4 = string: Maps\\mwisland.vec
Color = v3d: (.5,.5,1)
Alpha = double: .5
Error Factor = double: 4
```

Vector layers referencing tab separated values files

Navigation title:

When creating a vector layer that references a tab separated values (.tsv) file, the vector data is obtained by retrieving drawing instructions as well as longitude and latitude data from the .tsv file.

To define a vector layer that references a .tsv files, you must have the following:

- **A .tsv file** that contains the data used to draw the vectors on the globe, including longitude and latitude data. For more information about the required format of the .tsv file, see [Vector TSV file format](#) on page 759.
- **A layer file** that specifies the location of the .tsv file. For more information about the required format of the layer file, see [Vector layer file format](#) on page 759.

Vector TSV file format

The .tsv file must contain the following three tab separated columns:

- **Begin:** This column should indicate whether to begin a new line. Values in this column can be either 0 (do not begin a new line) or 1 (begin a new line).
- **Longitude:** This column should contain longitude values.
- **Latitude:** This column should contain latitude values.



Note: Any additional columns are ignored.

Following is a sample .tsv file that contains data for a vector layer:

```
Begin    Longitude      Latitude
1        -77.256903    38.98191
0        -77.262703    38.985309
0        -77.259803    38.989409
0        -77.255903    38.991009
0        -77.255217    38.988962
0        -77.256458    38.985688
0        -77.256903    38.98191
1        -77.23540928  38.97669798
0        -77.234803    38.97631
0        -77.232268    38.979502
0        -77.231601    38.979917
0        -77.229992    38.979858
```

Vector layer file format

Each vector layer file referencing .tsv files must be formatted using the following template:

```
Layer = VectorLayer:
  TSV Files = vector: n items
    0 = string: Maps\\File Name.tsv
    1 = string: Maps\\File Name.tsv
    .
    .
    n-1 = string: Maps\\File Name.tsv
  Color = v3d: color vector
  Alpha = double: alpha
  Width = double: width
  Error Factor = double: error factor
```

Vector layer parameters: .vec files

Parameter	Description
TSV Files	Path(s) to the .tsv file(s) containing the vector data. Example: Maps\\USVectorData.tsv
Color	The RGB color vector, which is expressed as (red,green,blue). For each color in the vector, you can enter a value from 0.0 to 1.0. For example, (1.0, 0.0, 0.0) is bright red, and (0.5, 0.5, 0.5) is gray.
Alpha	Controls the transparency of the vectors shown on the globe. The range is 0 to 1, with 0 being the most transparent.
Width	Optional. Sets the width of the data in pixels. The recommended range is 1 to 4.
Error Factor	Controls how accurately the vectors are drawn. For larger values, the vectors are drawn less accurately but faster. The default value is 5.

Making a New Vector Layer Available

Navigation title: Making a New Vector Layer Available

Steps to make a vector layer available to display on the globe visualization.

1. In the Profiles\profile name\Maps folder within the data workbench server installation directory, place the layer file and the .vec or .tsv files.
2. Edit the order.txt file in the Profiles\profile name\Maps folder to reflect the order in which you want the layers to display. By default, layers appear in lexicographic order by their names.



Note: When editing the order.txt file, take care not to cover up map layers that you want to show.

For more information about using order.txt files, see the Configuring Interface and Analysis Features chapter of the *Data Workbench User Guide*.

3. In data workbench, select the desired profile by right-clicking the workspace title bar and clicking **Switch Profile > <profile name>**.
4. Right-click the workspace title bar and click **Work Online**. An X appears next to **Work Online**.
5. Open a workspace and on a globe visualization, right-click and select the new layer. An X appears next to the layer name.

Working with Data Services

Navigation title: Working with Data Services

General information about data services, including installing and updating files.

About the Data Services

Information about data service options for mapping IP addresses.

If you are using an Adobe application, have licensed data workbench**Geography**, and want to map IP addresses on the globe or create visualizations showing geographically related dimensions of data (for example, area code, time zone, MSA) that are based on IP address, you can subscribe to one of two optional data services:

- IP Geo-intelligence data service (offered by Adobe, provided by Digital Envoy)
- IP Geo-location data service (offered by Adobe, provided by Quova)

Each data service provides you with a profile containing additional dimensions of data (such as Area Code, MSA, and Time Zone) and the supporting lookup files that contain the elements of those dimensions. You can use these additional dimensions in your analysis workspaces in data workbench.

Installing a Data Service on a Data Workbench Server

Navigation title:Installing a Data Service on a Data Workbench Server

Information about installing a data service on a data workbench server.

If you are using the IP Geo-intelligence data service or the IP Geo-location data service, you must install either the **IP Geo-intelligence** or **IP Geo-location** profile and the related lookup files on your data workbench server. You must complete the following procedures after you have installed the data workbench**Geography** profile. See [Installing Data Workbench Geography](#) on page 744. If you have not installed data workbench, follow the instructions in the *Data Workbench User Guide* before proceeding.



Note: To install the data service files, you must have access to the files on the data workbench server.

Adobe distributes the IP Geo-intelligence and IP Geo-location data services as .zip files. Each .zip file contains two folders: Lookups and Profiles. To install a data service on the data workbench server, you must perform the following steps:

- Install the data service profile. See [Installing the Data Service Profile](#) on page 761.
- Install the data service lookups. See [Installing the Data Service Lookup Files](#) on page 763.

You must install the data service profile and lookup files on the data workbench server machine on which you are processing and running your dataset profile. If you are running a data workbench server cluster, you must install the files on the master server. For information about dataset profiles, see the *Dataset Configuration Guide*.

Installing the Data Service Profile

Navigation title:Installing the Data Service Profile

The data service profiles (**IP Geo-intelligence** and **IP Geo-location**) are internal profiles that provide additional functionality to your Adobe application.

As with all other internal profiles provided by Adobe, these profiles should not be changed. All customization must occur in your dataset or role-specific profiles or other profiles that you create.

The data service profiles include the following dataset include files to be installed on a data workbench server:

- **Profiles\profile name\Dataset\Log Processing\Traffic\IP.cfg:** Lists the c-ip field to be passed from log processing to transformation.
- **Profiles\profile name\Dataset\Transformation\Geography\IPLookup.cfg:** Defines an IPLookup transformation that produces several fields of geographical data using the provided IP Geo-intelligence or IP Geo-location lookup file.

For information about transformation dataset include files, see the *Dataset Configuration Guide*.

In addition, each data service profile provides you with an element point layer file named **IP Coordinates.layer**. This layer file enables you to map locations in your dataset on the globe dynamically using IP addresses. After installation, the layer is stored in the **Profiles\data service name\Maps** folder within the data workbench server installation directory.

The **IP Coordinates.layer** file references the **Coordinates** dimension, which is defined in the **Coordinates.cfg** file provided with the **Geography** profile and located in the **Dataset\Transformation\Geography** folder. Each element of the **Coordinates** dimension defined in your dataset is mapped on the globe using the latitude and longitude information contained in that element. For more information about element point layers that use dynamic points, see [Defining Element Point Layers Using Dynamic Points](#) on page 756.



Note: If you installed the IP Geo-intelligence and IP Geo-location data service prior to version 5.1, your element point layer file references a lookup file instead of using dynamic points. Each layer file references the IP Geocodes lookup file and the IP Geocode dimension. The IP Geocodes lookup file contains a list of IP geocodes (geographical locations based on IP address) and the latitude and longitude for each. Each

element of an IP Geocode dimension defined in your dataset is mapped on the globe using the latitude and longitude listed for that IP geocode in the IP Geocodes lookup file.

The name of the layer file and the files that it references differ for each data service:

- The `IP Geocodes D.layer` file is installed with the IP Geo-intelligence (Digital Envoy) profile. This element point layer references the `IP Geocodes D yyyyymmdd.txt` lookup file (which you need to update periodically) and the IP Geocode D dimension.
- The `IP Geocodes Q.layer` file is installed with the IP Geo-location (Quova) profile. This element point layer references the `IP Geocodes Q yyyyymmdd.txt` lookup file (which you need to update periodically) and the IP Geocode Q dimension.

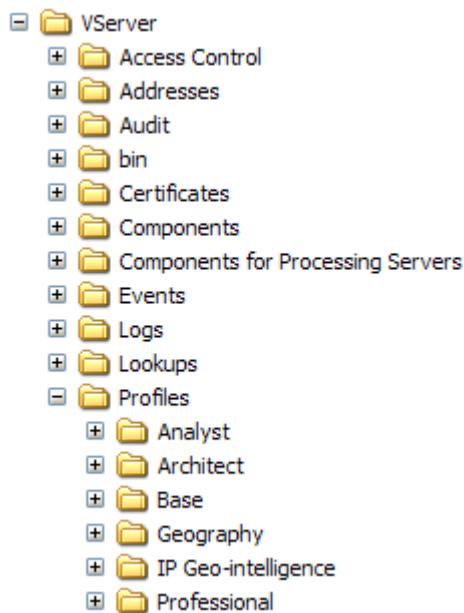
For more information about element point layers that use lookup files, see [Defining Element Point Layers Referencing Lookup Files](#) on page 754.

To install the IP Geo-intelligence or IP Geo-location profile



Note: The following installation instructions assume that you have installed data workbench and established a connection between data workbench and the data workbench Server on which you are installing data workbench**Geography**. If you have not done so, see the *Data Workbench User Guide*.

1. Open the Profiles folder from the `.zip` file that you received from Adobe.
2. Copy the IP Geo-intelligence or IP Geo-location folder to the Profiles folder in your data workbench server installation directory. You want to end up with a `...\Profiles\IP Geo-intelligence` folder or a `...\Profiles\IP Geo-location` on your data workbench server as shown in the following example. The names of the other folders within the `Profiles` folder may differ from the ones shown.



3. Use the following steps to update the `profile.cfg` file for each profile with which you want to use the **IP Geo-intelligence** or **IP Geo-location** profile.
 - a. Open the **Profile Manager**.
 - b. Right-click the check mark next to `profile.cfg` and click **Make Local**. A check mark for this file appears in the **User** column.

-
- c. Right-click the newly created check mark and click **Open > from the workbench**. The **profile.cfg** window appears.
 - d. In the **profile.cfg** window, right-click **Directories** and click **Add new > Directory**.
To add the new directory to the end of the list of directories, right-click the number or name of the last directory in the list and click **Add new > Directory**.
 - e. Type the name of the new directory: **IP Geo-intelligence** or **IP Geo-location**.
 - f. Right-click (**modified**) at the top of the window and click **Save**.
 - g. In the **Profile Manager**, right-click the check mark for **profile.cfg** in the **User** column, then click **Save to > <profile name>**.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe (including the **IP Geo-location** or **IP Geo-intelligence** profile), as your changes are overwritten when you install updates to these profiles.

Installing the Data Service Lookup Files

Navigation title:Installing the Data Service Lookup Files

Steps to install the IP Geo-intelligence or IP Geo-location lookup files.

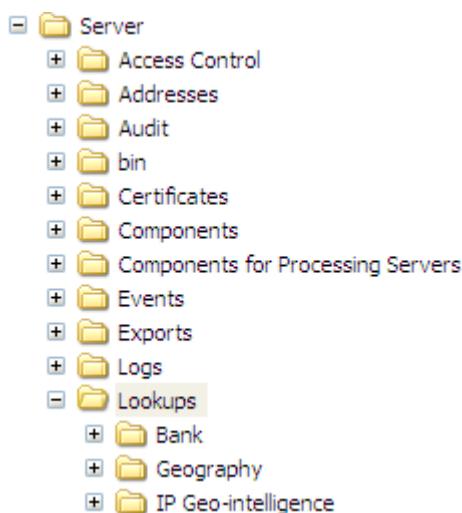
The lookup file (Lookups\profile name\data file name) provided with the data service profile is a binary file (.bin) that contains geographically related data based on IP address. You must replace this file periodically to ensure that you have the most recent geographic data. See [Updating Data Service Files](#) on page 764.

To install the IP Geo-intelligence or IP Geo-location lookup files



Note: All of your **IP Geo-location** or **IP Geo-intelligence** data files must fit in the available physical memory of your data workbench server.

1. Open the Lookups folder from the .zip file that you received from Adobe.
2. Copy the IP Geo-intelligence or IP Geo-location folder to the Lookups folder in your data workbench server installation directory (you want to end up with a ...\\Lookups\\IP Geo-intelligence or a ...\\Lookups\\IP Geo-location folder on your data workbench server as shown in the following example. The names of the other folders within the Lookups folder may differ from the ones shown.





Note: Periodically, Adobe sends you files containing updated **IP Geo-intelligence** or **IP Geo-location** lookup files. When you receive these files, you need to load them onto your data workbench server as directed by Adobe. For instructions, see the following section.

Updating Data Service Files

Navigation title: Updating Data Service Files

If you subscribe to either data service, you periodically must update the data service files provided by Adobe.

To do so, you must have access to the files on the data workbench server.

To load **IP Geo-location** or **IP Geo-intelligence** data files, you must complete the following procedures.

Replacing the Data File

1. In data workbench, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the **Servers Manager** workspace.
2. Within the **Servers Manager** window, right-click the icon of the data workbench server onto which you want to load the files and click **Server Files**.
3. In the **Server Files Manager**, right-click in the **Temp** column for **Lookups\IP Geo-location** or **Lookups\IP Geo-intelligence** and click **Open > <folder>**.
4. Copy the **.bin** data file provided by Adobe to the **Lookups\IP Geo-location** or **Lookups\IP Geo-intelligence** folder window.
5. Save the file to the data workbench server machine by right-clicking the **Temp** column for the data file and clicking **Save to > <server name>**.

If you are running a cluster, upload the files to the master data workbench server in the cluster.

Updating the IPLookup Transformation

1. In the **Profile Manager**, click **Dataset, Transformation**, and **Geography**.
2. Right-click the check mark next to **IP Lookup.cfg** and click **Make Local**. A check mark for this file appears in the **User** column.
3. Right-click the new check mark and click **Open > from the workbench**. A transformation configuration window appears.
4. In the window, click **Transformation**, then click **Transformations**.
5. Locate and click either **IPLookup Quova** or **IPLookup Digital Envoy**.
6. For the **File** parameter, update the name of the file to match the name of the new data (**.bin**) file provided by Adobe.
7. Save the transformation configuration file by right-clicking **(modified)** at the top of the configuration window and clicking **Save**.
8. Save the modified configuration file to each profile that uses the data service by right-clicking the check mark next to **IP Lookup.cfg** in the **User** column and clicking **Save to > <profile name>**. Retransformation of the data begins after synchronization of the dataset profile.

For information about retransformation of your dataset, see the Reprocessing and Retransformation chapter of the *Dataset Configuration Guide*.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe (including the **IP Geo-location** or **IP Geo-intelligence** profile), as your changes are overwritten when you install updates to these profiles.

If you installed the **IP Geo-intelligence** and **IP Geo-location** data service for version 5.1 or later, you have completed the data service update. However, if you installed the **IP Geo-intelligence** and **IP Geo-location** data service prior to version 5.1, you must complete the following additional procedures.

Replacing the Lookup File

You should complete the following steps only if you installed the **IP Geo-intelligence** and **IP Geo-location** data service prior to version 5.1.

1. In the **Server Files Manager**, click either **Profiles > IP Geo-intelligence** or **Profiles > IP Geo-location**, then click **Maps** to view its contents.
2. Right-click in the **Temp** column for **Maps** and click **Open > <folder>**.
3. Copy the new **.txt** file provided by Adobe to the **Maps** folder window.
4. Save the file to the data workbench server machine by right-clicking the check mark in the **Temp** column for the **.txt** file and clicking **Save to > <server name>**.



Note: If you are running a cluster, upload the files to the master data workbench server in the cluster.

Updating the Layer Files



Note: You should complete the following steps only if you installed the **IP Geo-intelligence** and **IP Geo-location** data service prior to version 5.1.

Complete these steps for every layer (**.layer**) file that references the **IP Geo-intelligence** or **IP Geo-location** lookup (**.txt**) file.

1. In the **Profiles\<data service name>\Maps** folder within the data workbench server installation directory, open the **.layer** file in a text editor such as Notepad.
2. In the **Data Paths** vector, update the name of the **.txt** lookup file to match the name of the new **.txt** file provided by Adobe, as shown highlighted in the following file sample:

```
Layer = ElementPointLayer:  
    Data Paths = vector< 1 items  
        0 = Path: Maps\\LookupFileName.txt  
    Longitude Column = string: LongitudeColumnName  
    Latitude Column = string: LatitudeColumnName  
    Name Column = string: LocationColumnName  
    Key Column = string: KeyColumnName  
    Dimension = ref< wdata/model/dim/DimensionName  
    Metric = ref< wdata/model/metric/MetricName
```

3. Save the updated layer file.
4. Repeat Steps 2 and 3 for every **.layer** file that references the **IP Geo-intelligence** or **IP Geo-location** **.txt** file.

Upgrading and Uninstalling Data Workbench Geography

Navigation title: Upgrading and Uninstalling Data Workbench Geography

Information about upgrading and uninstalling your data workbench**Geography** software.



Note: Be aware that when data workbench**Geography** is upgraded or uninstalled, the dataset automatically reprocesses.

Upgrading Data Workbench Geography

Navigation title:Upgrading Data Workbench Geography

Steps to upgrade the **Geography** profile and update the **Geography** lookups on your data workbench server.



Note: If the profile with which you are using data workbench**Geography** is running on a data workbench server cluster, upgrade the **Geography** profile on the master data workbench server in the cluster.

To upgrade the Geography profile on the data workbench server

1. Open the Profiles folder from the new .zip file provided to you by Adobe.
2. Copy the Geography folder to the Profiles folder in your data workbench server installation directory to replace the existing Geography folder.

To update the Geography lookups on the data workbench server

1. Open the Lookups folder from the new .zip file provided to you by Adobe.
2. Copy the Geography folder to the Lookups folder in your data workbench server installation directory to replace the existing Geography folder.

Uninstalling Data Workbench Geography

Navigation title:Uninstalling Data Workbench Geography

Steps to uninstall data workbench**Geography**.



Note: If the profile with which you are using data workbench**Geography** is running on a data workbench server cluster, uninstall the **Geography** profile from the master data workbench server in the cluster.

1. Use the following steps to update the `profile.cfg` file for each profile with which you were using data workbench**Geography**.
 1. Open the **Profile Manager**.
 2. Right-click the check mark next to `profile.cfg` and click **Make Local**. A check mark for this file appears in the **User** column.
 3. Right-click the newly created check mark and click **Open > from the workbench**. The `profile.cfg` window appears.
 4. In the `profile.cfg` window, delete the **Geography** profile entry from the **Directories** vector.
 5. If you have been using a data service, delete the **IP Geo-intelligence** or **IP Geo-location** profile entry from the **Directories** vector.
 6. Right-click (**modified**) at the top of the window and click **Save**.
 7. In the **Profile Manager**, right-click the check mark for `profile.cfg` in the **User** column, then click **Save to > <profile name>**.
2. Delete the Geography folder from the Profiles folder in your data workbench server installation directory.
3. If you have been using a data service, delete the IP Geo-intelligence or IP Geo-location folder from the Profiles folder in your data workbench server installation directory.
4. Delete the Geography folder from the Lookups folder in your data workbench server installation directory.
5. If you have been using a data service, delete the IP Geo-intelligence or IP Geo-location folder from the Lookups folder in your data workbench server installation directory.
6. If you created new terrain images, delete the `Terrain Images.cfg` file from the Components folder in your data workbench server installation directory.

Data Workbench Controlled Experiments

Controlled experiments are tests that enable you to compare results obtained from an experimental sample group with those from a standard control group.

Site enables you to implement, measure, and analyze controlled experiments and their results as they relate to different aspects of your website. Doing so enables you to test hypotheses regarding the improvement of website performance before spending significant time or money fully implementing the proposed changes.



Note:

Site experiments can be analyzed only in datasets where the only method of visitor identification in use is the **Sensor** set persistent cookie method. Sensors running on J2EE servers (JBoss, Tomcat, WebLogic, and WebSphere) do not support controlled experimentation. For more information, see the following section.

Using Site, you can implement A/B, A/B/A, and multivariate controlled experiments to gather enough test data to provide statistically accurate data for a detailed evaluation of your hypothesis, without impacting current website performance.

How Does Site Identify Visitors?

Navigation title:How Does Site Identify Visitors?

A typical configuration of Site uses cookies to uniquely identify visitors to your website and track their behavior over time.

The first time that a particular browser (considered a visitor) makes a request of your website, **Sensor** works with your web server to set a persistent cookie (by default `cs(cookie)(v1st)`), which is interpreted internally within the system as `x-trackingid`. This cookie is set only once, on the very first request made to your website by that visitor. It then is collected from that visitor each time that browser makes a request (either page or embedded object request) of your website in the future.

Accepting a persistent cookie is at the browser's discretion. If a user does choose to block persistent cookies, their page view requests are still logged, but the measurement data from those requests are not correlated to a particular visitor or their sessions on the website unless you implement an alternate method of visitor identification, such as using the Hash transformation on the IP and UserAgent fields.



Note: Site experiments can be analyzed only in datasets where the only method of visitor identification in use is the **Sensor** set persistent cookie method. Sensors running on J2EE servers (JBoss, Tomcat, WebLogic, and WebSphere) do not support controlled experimentation.

During a controlled experiment, users who do not accept cookies could be placed in different experiment groups from one click to the next. This becomes an issue only if you perform your test analysis with the Broken Session Filter turned off in Insight, which Adobe does not recommend.

For more information about the Broken Session Filter, see the *Insight User Guide*.

If a visitor clears the cookie during an experiment, the visitor is assigned a new cookie and potentially could be assigned to a different group. Because Adobe identifies the visitor as new, the experiment is not invalidated.

How Do Controlled Experiments Work?

Navigation title:How Do Controlled Experiments Work?

In an experiment, you can define any number of test groups in addition to the control group.

When an experiment is running, all visitors to your website become part of the experiment, either as part of a test group or of the control group, as soon as they access any page involved in the experiment. Visitors are allocated to your experiment groups randomly, in proportions defined during the experiment configuration.

Controlled experiments are implemented using the **Sensor** software that is installed on each of the content servers in your web cluster. As the content servers receive requests, **Sensor** randomly selects visitors for your test groups and redirects their page requests to the experimental content. When **Sensor** selects a visitor to view the test content, the address bar continues to list the originally requested URI but the visitor is routed to the test URI. Because this process takes place internally in the server application, users are not aware of when they are being tested, which is an important consideration for unbiased experimentation.

Sensor passes the test URIs, not the original URI displayed to the user, to the log files for use in analysis.

The results of the experiments can be analyzed easily using Insight to determine whether the experimental hypothesis that you were testing is correct.



Note:

Adobe strongly recommends that controlled experiments be coordinated and performed with input from those individuals in your organization who are responsible for configuring and maintaining your analysis datasets.

What Should I Test?

Navigation title:What Should I Test?

Test results must be clear and meaningful so that you can feel confident making large dollar decisions based on those results.

Although you can test various page layouts with **Sensor** and Site, Adobe suggests that you focus on testing high-value, strategic business initiatives, or new or redesigned website functionality that address the goals that you have set for your website as well as for your business. You can test for such issues as best price guarantees, personalization functionality, market offers (for example, packages or bundles), creative design, and application processes.

The following concepts are most important when developing your controlled experiment:

- **Understand the right changes to make.** This requires some research into how your website functions and the business processes underlying the front-end website. You want to make changes that provide the most impact and can be tested easily.
- **Small changes can have significant impact.** Not all of the changes that you test need to be drastic to have a significant impact on your business. Always be open to making small, but very important changes.

Supported Methodologies

Many types of experiments with many different goals can be performed using Site. The following list provides a few examples:

- Altering pages, content, and website processes to improve conversion rates.
- Changing marketing campaigns, promotions, cross-sells, and up-sells to increase revenue.
- Varying page load times to understand customer quality of service and the actual value of infrastructure performance.

To reach these goals, Site supports the following types of methodologies for controlled experimentation and testing:

- **Page Replacement:** Replace static URL X with static URL Y. This methodology is of limited use in a dynamic environment.

-
- **Dynamic URI Replacement:** This is a variant of Page Replacement that replaces static page X with dynamic page Y to render dynamic content.
 - **Object Replacement:** Replace fixed object X with fixed object Y.
 - **Content Replacement:** Replace content set X (multiple objects, pages, table, and so on) with content set Y.
 - **Experiment Variable Replacement:** Replace JavaScript object /writeCookie_X.js with JavaScript object /writeCookie_Y.js to write a cookie that can be used by a back-end system to serve particular content.



Note:

Controlled experiments are based on URI replacement, not query string replacement. The URI within a particular URL is highlighted in the following example:

```
http://www.omniture.com/index.asp?id=1
```

For example, in your controlled experiment you could specify that the control group URI index.asp be replaced with the test group URI index2.asp to determine which page design would result in more value.

What Are the Requirements?

Navigation title:What Are the Requirements?

To perform a controlled experiment on your website using Site, you must meet the following requirements:

- **Sensor** must be installed and working properly on each web or application server that supports the website content that you are testing.



Note: Sensors running on J2EE servers (JBoss, Tomcat, WebLogic, and WebSphere) do not support controlled experimentation.

- You must have a process in place for pushing content to all of your web or application server(s), such as a content management system.

Enabling Controlled Experimentation

Navigation title:Enabling Controlled Experimentation

To enable controlled experimentation, someone with administrator access to your web or application servers must modify the ExpFile parameter in the **Sensor** configuration file (usually named using txlogd.conf) on each web or application server in your web cluster on which a **Sensor** is installed.

In addition, two other parameters in this file can be modified to implement a testing tool (ExpCookieURL parameter) or to remap large sections of your website (ExpPartialMatch parameter). This chapter provides more information about these parameters.

To edit the txlogd.conf file

If you have administrator access, complete the following steps. If you do not have administrator access, contact your system architect to request the changes, providing them with the following steps.

1. Navigate to the **Sensor** installation folder on a web or application server in your web cluster on which a **Sensor** is installed.
2. Open the **Sensor** configuration file (usually named using txlogd.conf) using a text editor and edit the file as indicated in *Modifying the ExpFile Parameter* on page 770.
(And optionally in *Modifying the ExpCookieURL Paramter (Optional)* on page 770 and *Modifying the ExpPartialMatch Parameter (Optional)* on page 770.)
3. Save and close the file.

-
4. Repeat this procedure for each web or application server in your web cluster on which a **Sensor** is installed.

Modifying the ExpFile Parameter

Navigation title:Modifying the ExpFile Parameter

The ExpFile parameter points to the location of the experiment configuration file, which defines your experiment.

Setting this parameter enables you to run experiments. For steps to create the experiment configuration file, see [Configuring and Deploying the Experiment](#) on page 774.

Following is an example of the ExpFile parameter:

```
ExpFile /home/experiment.txt
```

This tab delimited text file (.txt) can be located anywhere in the **Sensor** folder and can have any convenient name.

Make sure you record the location of the experiments directory and the name of the configuration file that you specify because you need to save your experiment configuration file (to be described later in this guide) using this name and in this directory.



Note: If you do not set this parameter identically on each machine in your web cluster on which a **Sensor** is installed, controlled experimentation does not work.

This entry can be preconfigured and remain in the **Sensor** configuration file on an ongoing basis with no adverse effect. If the experiment configuration file name specified is not found by **Sensor** or it is blank (that is, it exists but has no content), **Sensor** does not conduct the experiment, logs an error event on the HTTP server, and continues to operate normally in all other respects.

Modifying the ExpCookieURL Paramter (Optional)

Navigation title:Modifying the ExpCookieURL Paramter (Optional)

The ExpCookieURL parameter can be used to test that your controlled experiment is working properly.

This parameter defines the URL of a virtual page that when requested places you into a specified experiment and group and then redirects you to the root of your website. From that point through the end of the experiment, you are part of the specified experiment and group.

The default page for this parameter is `setcookie.htm`, but you can use any valid virtual URL.



Note: This must be a virtual URL and must not be a real page or piece of existing content. There should be no file on the web server at the specified path with the specified name.

Following is an example of the ExpCookieURL parameter:

```
ExpCookieURL /setcookie.htm
```

Modifying the ExpPartialMatch Parameter (Optional)

Navigation title:Modifying the ExpPartialMatch Parameter (Optional)

If you want to enable your controlled experiments to remap your entire website or an entire subdirectory of your website to another location, you can set the ExpPartialMatch parameter in the `txlogd.conf` file to “on.” The default is “off.”

Following is an example of the ExpPartialMatch parameter:

```
ExpPartialMatch off
```

Be very careful when setting this parameter to “on” because it can result in an inadvertent remapping of your entire website.

Creating a Controlled Experiment

Navigation title:Creating a Controlled Experiment

The overall objective of running a controlled experiment on a website is to determine whether and to what degree a defined change or set of changes produces an effect on the users of the website.

Defining the Objective

Navigation title:Defining the Objective

Define the primary function of the objective, and target audience.

When defining your objective, consider the purpose of the website or website process you are analyzing: *What is its primary function? Who is its target audience?*

Common website objectives include converting more visitors into customers or increasing the average amount of revenue gained for all visitors within an experiment over the duration of that experiment.

Common website process objectives include improving the steps or pages in a process that cause visitors to abandon the process, removing unnecessary and confusing options in the process that tend to stop visitors from reaching the end of the process, or consolidating or expanding the process to eliminate or add steps or pages.

Make sure to think carefully about what it is specifically that you want to understand about your website. Carefully planning your experiment makes the results much more meaningful to your business.

Objective: To increase the number of visitors to our website who request a demo of our product using the “Request a Demo” graphical link, as shown in the following image:



Forming a Hypothesis

Navigation title:Forming a Hypothesis

A hypothesis is an assumption, which can be taken as true for the purpose of argument, that provides a tentative explanation that can be tested by further investigation.

Try to think about your hypothesis in terms of alternative pages, images, or processes that could be substituted easily for existing ones.

Your hypothesis must be able to produce a result with statistical significance. This can be achieved by increasing the percentage of visitors included in the test, or by running the test for a longer period of time.

At this point, you also should define your visitor-based success metrics, either as part of the hypothesis or as an additional matrix.

Hypothesis: Moving the “Request a Demo” graphical link closer to the top of the page results in a Visitor Conversion increase of at least 1.5%.

In our example hypothesis, we have defined the success criterion for this experiment as an increase in Visitor Conversion of at least 1.5%.





Determining the Number of Visitors Needed

Navigation title:Determining the Number of Visitors Needed

To create a statistically significant experiment, you must determine how long you need to run the experiment to include enough visitors to effectively evaluate the results of the changes to your website.

If you need help determining the minimum length of your experiment, you can use the experiment design spreadsheet provided by Adobe as a tool to help you design your experiment. This file, named VS_Controlled_Experiment_Design.xls by default, functions not only as a worksheet but also as a record of your decisions about the experiment. For more information about this file, see *Experiment Design Spreadsheet* on page 783.



Note: The experiment design spreadsheet can provide useful statistical inferences only when the metric in question is defined as a percentage of visitors that meet some criteria. That is, it is useful only when trying to test a visitor-based metric hypothesis.

Creating the Test Content

Navigation title:Creating the Test Content

Before you configure the experiment, you should create the alternate content that you want to use in the experiment.

The control group is sent to the original URI, while the test group is sent to the new, alternate URI.

To avoid confusion, do not reuse test group file names. For example, if you run an experiment using a test group file named test2.asp, do not use test2.asp as the name for the test file in your next experiment.

For the hypothesis that moving the “Request a Demo” graphical link on your home page impacts Visitor Conversion, we create the alternate home page containing the “Request a Demo” graphical link in the new position. The following section describes how you then specify that the control group URI index.asp be replaced with the test group URI index2.asp for a certain percentage of visitors.

Configuring and Deploying the Experiment

Navigation title: Configuring and Deploying the Experiment

After you have defined your objective, hypothesis, and experiment details as well as created your test content, you must configure **Sensor** to deploy the controlled experiment.

Configuring the Experiment Configuration File

To configure the experiment, you must complete the experiment configuration spreadsheet provided by Adobe (named `TestExperiment.xls` by default). This file configures **Sensor** to perform the experiment and is the Excel version of the text file that you specified in [Modifying the ExpFile Parameter](#) on page 770.

This file can contain information about multiple experiments, which can run at the same or at different times and use different groups and percentages, but these experiments are not correlated in any way.

Users are placed into a group for each experiment listed in the file that is configured to be running at this time.



Note: Each experiment is independent of all other experiments. Changes you make to one experiment do not affect any other experiment, and although visitors may be in multiple experiments, the results do not relate to one another. If you think a correlation exists between the changes in multiple experiments, you must create a new experiment that tests these changes together.

To configure your experiment

You should complete this file before the experiment begins and not modify the information while the experiment is running.



Note: Any experiment is promptly invalid if the definition of the experiment changes after the experiment has begun.

1. If you have administrator access to your web or application servers, navigate to the **Sensor** installation folder on any **Sensor** machine in your web cluster to access the `TestExperiment.xls` file. If you do not have administrator access, contact your Adobe account manager to request the `TestExperiment.xls` file.
2. Open the `TestExperiment.xls` file (you can rename this file if desired) and complete the following fields:

Field	Description
Experiment	A descriptive name for the experiment. Each experiment name must be unique and cannot contain spaces. Experiment names are used when displaying the results of experiments in Insight. The names appear as the first half of the element names in the controlled experiment dimension. The second half of the element name is the group name from the Group field in this file. Each group is named in the following format using the experiment name followed by the group name: <i>ExperimentName.Group Name</i> For example: <code>New_Homepage.Control</code>
Start	The date and time that you want the experiment to begin. If you do not enter values, the experiment begins immediately after the file is deployed. Format: MM/DD/YYYY H:MM • If you leave the start and stop times empty, the experiment runs indefinitely. • You can predefine start and stop times well in advance; therefore, you can configure all of your experiments for the next year at once if desired.

Field	Description
	<ul style="list-style-type: none"> Start and stop times are based on the system time of the web server. If that clock changes for any reason, your experiment may start or stop unexpectedly. If you would like to add an experiment as a configuration file entry but do not want the experiment to run in the near future, you can comment out the experiment information using the number sign “#” or define start and stop times in the past.
Stop	<p>The date and time that you want the experiment to end. When the stop date and time occurs, Sensor will stop sending the cookie values identified as a test group to the test URIs and will send all cookies to the control group URIs.</p> <p>Format: MM/DD/YYYY H:MM</p> <p>See the notes for the Start field.</p>
Group	<p>A descriptive name for each group of visitors in the experiment. Group names cannot contain spaces.</p> <p>Group names are used when displaying the results of experiments in Insight. For more information, see the Experiment field description.</p> <p>A control group can be implicitly or explicitly defined based on the value entered in the Percentage field.</p> <p> Note: To meet the number of visitors needed during the defined time period for the experiment to be statistically valid, you may need to either decrease the confidence level or increase the time period. For example, if your time frame is five days, your confidence level is 98%, and your number of visitors needed exceeds the number expected for that time period, you need to either increase the time period or decrease the confidence level until the number of visitors expected exceeds the number needed to run a statistically valid experiment.</p>
Percentage	<p>The percentage of website visitors to include in each defined group. These values can be expressed as either percentages or decimal values. In addition, both values must be either greater than or less than one.</p> <p>For example:</p> <p>33.3% and 66.7%</p> <p>.99 and .01</p> <p>If the sum for all groups is less than 100, the undefined excess defaults to a control group.</p>
Original URL	<p>The URI of the content to be remapped, followed by \$. This value is case-sensitive.</p> <p>Format: index.asp\$</p> <p>Original URIs can be specified using a dollar sign (\$) at the end of the URI to denote that an exact match of the file name is required. For example, the expression /product/product_view.asp\$ matches that exact page only, while /product matches any page in the /product directory and could be used to remap that entire sub-tree. Original URL entries that do not specify the \$ character at the end of the file name are ignored by the experiment unless the ExpPartialMatch parameter has been set to “on.” For more information about this parameter, see Modifying the ExpPartialMatch Parameter (Optional) on page 770.</p> <p>The controlled experiment functionality ignores any query strings appended to the URI stem. For example, the page /product/product_view.asp?productId=53982 is not a valid URI, but the page /product/product_view.asp is a valid URI.</p>
Remapped URL	<p>The URI of the alternate content.</p>

Field	Description
	Format: index2.asp See the notes for the Original URL field.

The following is an example of a completed `TextExperiment.xls` spreadsheet:

	A	B	C	D	E	F	Re
1	#Experiment	Start	Stop	Group	Percentage	Original URL	Re
2	New_Homepage	6/1/06 0:00	6/30/06 23:59				
3				Control	50%		
4						/index.asp\$	/in
5							
6				index2	50%		
7						/index.asp\$	/in



Note: Do not modify the column positions in the spreadsheet.

This example indicates that the “New_Homepage” experiment starts on June 1, 2006, ends on June 30, 2006, and contains a control group with 50% of the visitors and a test group with 50% of the visitors, who see different content for one URI.



Note: Although the sample file above has an explicit control group defined, it is not necessary to explicitly define a control group — the experiment automatically creates the control group. If the sum of the percentages for all groups in an experiment is less than 100%, an implicit control group is assigned to users that do not fall into one of the explicit groups.

- To insert comments to provide additional information about specific experiments, begin the cell with a number sign (#) and follow with your comments. Comments can be inserted anywhere in the file.
- After you have completed the variables in the experiment configuration spreadsheet, save the changes, then save the file in tab-delimited text format (*.txt) using the name that you specified in the ExpFile parameter in the **Sensor** configuration file. See [Modifying the ExpFile Parameter](#) on page 770.

The following is an example of an experiment configuration text file:

```
TestExperiment.txt - Notepad
File Edit Format View Help
#Experiment      Start     Stop      Group      Percentage      Original URL      Remapped URL
New_Homepage    6/1/06 0:00  6/30/06 23:59
                           Control   50%          /index.asp$        /index.asp$
                                         index2   50%          /index.asp$        /index2.asp$
```



Note: Because of the tabs required in this file, do not edit the experiment configuration text file by hand. If you need to make changes to the file, make the changes in the experiment configuration Excel file and re-save the file as a tab-delimited text file.

If you have defined Start and Stop times, there is no reason to ever delete an experiment from the experiment configuration file. Keeping all of your experiments listed in the experiment configuration file is actually a good way to keep a record of how you defined each of your experiments.

Deploying the Configuration File and Test Content

You must deploy the experiment configuration file to each machine in your web cluster that is running a **Sensor** and serving the pages involved in the experiment. You can do so using either a manual procedure or your existing content management system.

To deploy your test content

- On each application or web server running a **Sensor** that is serving pages involved in the experiment, use your existing publishing process to deploy the test content to the appropriate location.

For example, if you want to publish the test group page `index2.asp` to the test folder for your website (`mysite.com`), you would publish the file to `www.mysite.com/test`.



Note: Do not link to any of your test files directly from a page on your website. Doing so invalidates your test results and your index scores.

To deploy your experiment

- On each application or web server running a **Sensor** that is serving pages involved in the experiment, place the experiment configuration text file in the directory that you specified in the `ExpFile` parameter in the **Sensor** configuration file. See [Modifying the ExpFile Parameter](#) on page 770.

Sensor randomly selects website visitors for each group based on the percentages that you defined in the file and serves the test or control group content to them as appropriate.

Validating the Experiment

Navigation title:Validating the Experiment

After you have deployed your experiment, you should validate that the experiment is working properly.

As discussed in [Modifying the ExpCookieURL Paramter \(Optional\)](#) on page 770, the page specified in the `ExpCookieURL` parameter in the **Sensor** configuration file can be used to place yourself in a specific experiment group.

The default virtual page is `/setcookie.htm`, but you must use the value that you set in the `ExpCookieURL` parameter.

Requesting the Test Page

To test a specific experiment group for your website, your browser must be configured to accept cookies and you must not already have a cookie for this website.

Each time you want to test a new group, make sure to clear your cookies for the website.

To place yourself into a specific group within a specific experiment, request the test page with a query string in the following form:

`http://<sitename>/?Experiment Name=Group Name`

For example:

`http://www.omniture.com/setcookie.htm?New_Homepage=index2`

When the virtual URL request is sent to the server, **Sensor** identifies you as a member of the specified group within the specified experiment and then redirects you to the root of the website. You now can navigate to the

appropriate location on the website to validate whether the correct content displays for that experiment and group.

If you were to type the following into your browser, the browser would display the home page of the website and place you into the index2 group within the New_Homepage experiment:

http://www.omniture.com/setcookie.htm?New_Homepage=index2

When visitors in the index2 group request the home page, the “Request a Demo” graphical link displays higher on the page, as in the following graphic:



Viewing the Results

Navigation title:Viewing the Results

You can view the results of an experiment in a visualization within Insight.

Modifying Log Processing.cfg

Navigation title:Modifying Log Processing.cfg

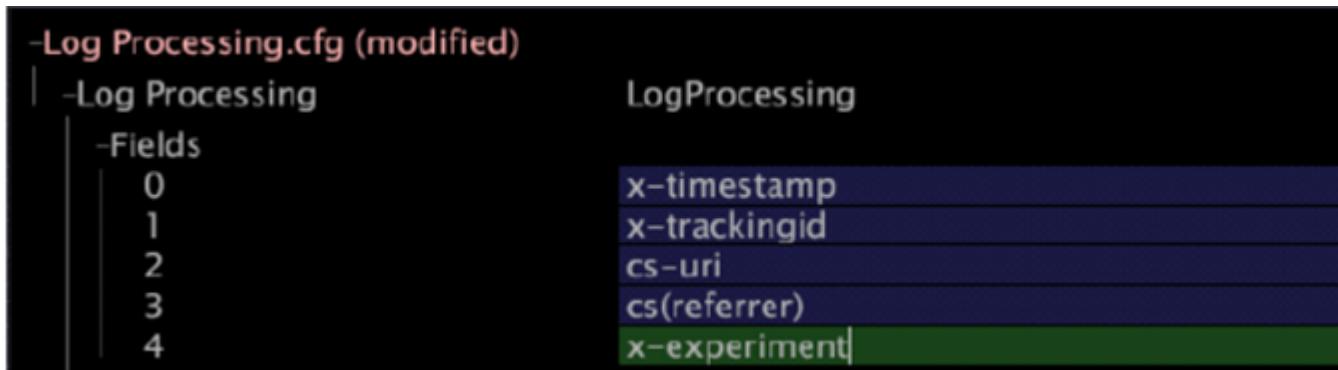
You must add the x-experiment field to the Log_Processing.cfg file, which is used to create an extended dimension.

See [Modifying Transformation.cfg](#) on page 779.

To modify Log Processing.cfg

1. In Insight, open the **Profile Manager** by right-clicking within a workspace and clicking **Admin > Profile Manager**, or by opening the Profile Management workspace on the **Admin** tab.
2. In the **Profile Manager**, click **Dataset** to show its contents.
3. Right-click the check mark next to Log_Processing.cfg and click **Make Local**. A check mark for this file appears in the **User** column.

4. Right-click the newly created check mark and click **Open > in Insight**. The Log Processing.cfg window appears.
5. Click **Fields** to show its contents.
6. Right-click the last field in the current list and click **Add new > Field**.
7. Type x-experiment in the newly created field, as shown in the following example:



8. Right-click **(modified)** at the top of the window and click **Save**.
9. In the **Profile Manager**, right-click the check mark for Log Processing.cfg in the **User** column, then click **Save to > <profile name>** to save the locally made changes to the working profile.



Note: The dataset begins reprocessing immediately.

For more information about Log Processing and Fields see the *Dataset Configuration Guide*.

Modifying Transformation.cfg

Navigation title:Modifying Transformation.cfg

Now that the x-experiment field is available, you must create an extended dimension to include the x-experiment field in your dataset, which allows you to view your results in Insight.

To do so, you must add a new dimension to the Transformation.cfg file.

If you plan to run multiple experiments, you also must add a new Split transformation to the Transformation.cfg file. This Split transformation separates the different experiment and group names so that the information is easier to interpret. To avoid reprocessing your data again if you were to need to add additional experiments at a later date, Adobe recommends that you add the Split transformation even if you are not currently planning to run multiple experiments.

The following procedure includes the creation of both the new Split transformation and the extended dimension. If you do not want to add the Split transformation, simply skip steps 5–7.

To modify Transformation.cfg

1. In Insight, open the **Profile Manager** by right-clicking within a workspace and clicking **Admin > Profile Manager**, or by opening the Profile Management workspace on the **Admin** tab.
2. In the **Profile Manager**, click **Dataset** to show its contents.
3. Right-click the check mark next to Transformation.cfg and click **Make Local**. A check mark for this file appears in the **User** column.
4. Right-click the newly created check mark and click **Open > in Insight**. The Transformation.cfg window appears.
5. Click **Transformation** to show its contents.

-
6. Right-click **Transformations** and click **Add new > Split**.
 7. Complete the new split on comma transformation as shown in the following example:

-Transformation.cfg (modified)	
-Transformation	Transformations
+Extended Dimensions	
+Internal Domains	
+Page View Condition	AndCondition
-Transformations	
-0	Split
Name	Split the Experiment Dimension
Comments	Comment
Condition	AndCondition
Delimiter	,
Input	x-experiment
Output	x-experimentsplit



Note: You can enter any value in the Name field.

8. Right-click **Extended Dimensions** and click **Add new > ManyToMany**.
9. Complete the new dimension as shown in the following example:

-Transformation.cfg (modified)	
-Transformation	Transformations
-Extended Dimensions	
-0	ManyToMany
Name	Controlled Experiment Groups
Case Sensitive	true
Comments	Comment
Condition	AndCondition
Hidden	false
Input	x-experimentsplit
Maximum Elements	0
Parent	visitors



Note:

- You can enter any value in the Name field.
- If you did not include the Split transformation, you must type “x-experiment” in the **Input** field.

10. Right-click **(modified)** at the top of the window and click **Save**.
11. In the **Profile Manager**, right-click the check mark for Transformation.cfg in the **User** column, then click **Save to > <profile name>** to save the locally made changes to the working profile.



Note: The dataset begins retransforming immediately.

For more information about Transformation.cfg and extended dimensions, see the *Dataset Configuration Guide*.

Viewing the Experiment Results

Navigation title:Viewing the Experiment Results

After you have added the new field to Log Processing.cfg and created the new Split transformation and extended dimension, you can view the new extended dimension that you created as soon as the Fast Input stage of data reprocessing has finished.

This dimension, by default, displays the number of sessions for each of your experiment groups.

To view the experiment dimension

- Within any workspace in Insight, open a table with the experiment dimension that you created.

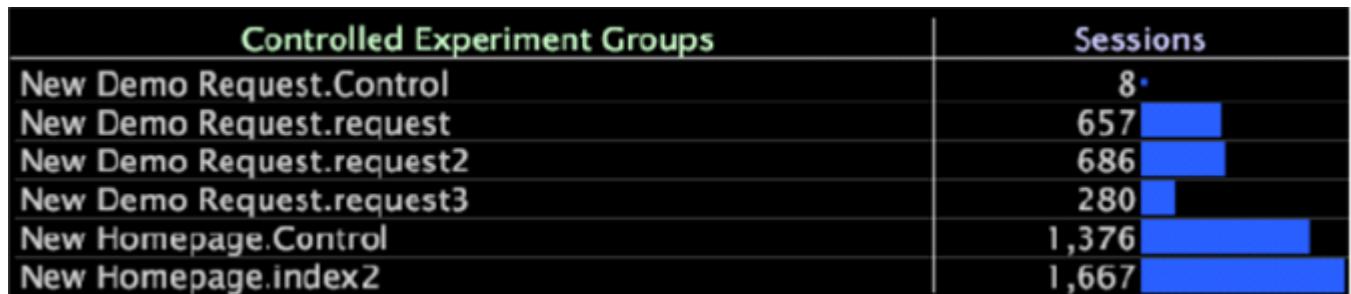
The experiment dimension elements, which represent each experiment you are currently running and each group within each experiment, display with the current number of sessions for each group. Each group is named in the following format using the experiment name followed by the group name:

Experiment Name.Group Name

For example: New_Homepage.Control

The following table shows the Controlled Experiment Groups dimension that was created in Transformation.cfg and each of the experiments and their groups.

The New_Homepage experiment is shown at the bottom of the table with its two groups: Control and index2.



You now can use the experiment dimension and any relevant metrics to explore and interpret the experiment results, as well as create useful reports detailing those results.

Evaluating the Experiment

Navigation title:Evaluating the Experiment

After running the experiment until the required minimum number of visitors have participated in the experiment, you can be assured of sufficient statistical confidence to evaluate the results of your experiment.

Using Insight, compare whichever metrics or key performance indicators were defined as part of the hypothesis to determine whether the experiment was a success (that is, the hypothesis was validated with the specified confidence.)

In our example experiment, our hypothesis is proven correct if the Visitor Conversion increases by at least 1.5%, which is the success criterion we defined earlier.

The following workspace example shows that the Conversion for the index2 test group was actually 1.8% higher than for the control group, proving our hypothesis.



- [Summarizing the Experiment Results](#) on page 782
- [Taking Action Based on the Results](#) on page 783
- [Monitoring Your Actions](#) on page 783

Summarizing the Experiment Results

Using Insight, you can create detailed reports to summarize and illustrate the results of your experiment.

You then can use your reports, as shown in the following example, to make recommendations based on the results, which are backed up by the visual information you have provided in your reports:



Taking Action Based on the Results

After the results are clear, you are ready to act on those results by making production-level changes to the tested pages, applying these same changes to other areas of your website, and making sure to completely document the test, its results, and the changes that you have made.

Monitoring Your Actions

After the controlled experiment is complete and you have implemented the appropriate changes, make sure to continue to monitor the changes that you made by, for example, viewing validation metrics, creating control charts, and providing dashboard metrics.

Always be prepared to re-test your hypothesis if you think the changes that you tested and made are not bearing out the original results.

Experiment Design Spreadsheet

Navigation title:Experiment Design Spreadsheet

This file functions not only as a worksheet but also as a record of your decisions about the experiment.

If you need help designing your experiment, you can use the experiment design spreadsheet (named VS Controlled Experiment Design.xls by default) provided by Adobe.

The experiment design spreadsheet can provide useful statistical inferences only when the metric in question is defined as a percentage of visitors that meet some criteria. That is, it is useful only when testing a visitor-based metric hypothesis.

To design your experiment using the experiment design file

1. If you have administrator access to your web or application servers, navigate to the **Sensor** installation folder on any **Sensor** machine in your web cluster. If you do not have administrator access, contact your Adobe account manager to request the file.
2. Open the VS Controlled Experiment Design.xls file. (You can rename this file if desired.)

The spreadsheet on the following page is an example of how you would complete the spreadsheet when preparing to test the example hypothesis used throughout this guide.

<p>This experiment design worksheet should be completed before the experiment begins, and never modified thereafter. Any experiment is promptly invalid if the design of the experiment is changed after the experiment has begun.</p>																																
Experiment Title:	New Homepage - Moving the "Request a Demo" graphical link																															
Experiment Description:	To determine whether moving the "Request a Demo" graphical link closer to the top of the page will increase A) the relative number of clicks on the link and B) the relative number of successful demo requests the company receives.																															
Metric Being Studied:	Visitor Conversion	Examples and Explanations																														
Metric Definition:	Visitors[URI='conversionpage.asp']/Visitors	Visitor Conversion Must be like Visitors[X]/Visitors																														
Intended Start Time:	Jun 1, 2006 12:00 AM	e.g. 9/1/2006 12:00:00 AM																														
Intended End Time:	Jun 30, 2006 11:59 PM	e.g. 11/1/2006 12:00:00 AM																														
Applicable Selections:	Dimension N/A	Element Set or Range N/A	Segmentation of Dataset 1, If Any Segmentation of Dataset 2, If Any Segmentation of Dataset 3, If Any																													
Experiment URIs:	Control http://www.visualsciences.com/index.asp	Test http://www.visualsciences.com/index2.asp	Control URI Test URI Control URI Test URI																													
<table border="1"> <thead> <tr> <th colspan="2">Expected Metrics for Applicable Selections:</th> </tr> </thead> <tbody> <tr> <td>Average Visitors per Day</td> <td>1,000</td> <td>Examples and Explanations</td> </tr> <tr> <td>Visitor Conversion</td> <td>10%</td> <td>Expected Visitors Per Day During Test Period</td> </tr> <tr> <td>Expected Visitors per Period</td> <td>30,000</td> <td>Expected Value of Metric Being Studied</td> </tr> </tbody> </table>		Expected Metrics for Applicable Selections:		Average Visitors per Day	1,000	Examples and Explanations	Visitor Conversion	10%	Expected Visitors Per Day During Test Period	Expected Visitors per Period	30,000	Expected Value of Metric Being Studied	Calculated Expected Visitors per Test Period																			
Expected Metrics for Applicable Selections:																																
Average Visitors per Day	1,000	Examples and Explanations																														
Visitor Conversion	10%	Expected Visitors Per Day During Test Period																														
Expected Visitors per Period	30,000	Expected Value of Metric Being Studied																														
<p>Experiment Will Determine If the Visitor Conversion for the Test Group Is...</p> <table border="1"> <tbody> <tr> <td>Greater Than The Value For The Control Group?</td> <td>TRUE</td> <td>TRUE or FALSE</td> </tr> <tr> <td>Less Than The Value For The Control Group?</td> <td>TRUE</td> <td>TRUE or FALSE</td> </tr> <tr> <td>By at Least (Detection Level)</td> <td>1.50%</td> <td>(e.g. 9.25% vs. 10.75%)</td> </tr> <tr> <td>With a Confidence Level of at Least</td> <td>95.00%</td> <td>(5.00% False Positives)</td> </tr> <tr> <td>and a Power Level of</td> <td>95.00%</td> <td>(5.00% False Negatives)</td> </tr> <tr> <td>Calculated Z Score for Type I Error</td> <td>1.96</td> <td>False Positives Intermediate Calculation</td> </tr> <tr> <td>Calculated Z Score for Type II Error</td> <td>1.64</td> <td>False Negatives Intermediate Calculation</td> </tr> <tr> <td>Minimum Visitors Required (Test+Control Groups)</td> <td>34.9%</td> <td>Percentage of Available Visitors Given Segmentation</td> </tr> <tr> <td>Minimum Visitors Required (Test+Control Groups)</td> <td>10,475</td> <td>Number of Required Visitors Given Segmentation</td> </tr> <tr> <td>Minimum Experiment Time (Days)</td> <td>10.5</td> <td>Minimum Number of Days Experiment Must Run</td> </tr> </tbody> </table>		Greater Than The Value For The Control Group?	TRUE	TRUE or FALSE	Less Than The Value For The Control Group?	TRUE	TRUE or FALSE	By at Least (Detection Level)	1.50%	(e.g. 9.25% vs. 10.75%)	With a Confidence Level of at Least	95.00%	(5.00% False Positives)	and a Power Level of	95.00%	(5.00% False Negatives)	Calculated Z Score for Type I Error	1.96	False Positives Intermediate Calculation	Calculated Z Score for Type II Error	1.64	False Negatives Intermediate Calculation	Minimum Visitors Required (Test+Control Groups)	34.9%	Percentage of Available Visitors Given Segmentation	Minimum Visitors Required (Test+Control Groups)	10,475	Number of Required Visitors Given Segmentation	Minimum Experiment Time (Days)	10.5	Minimum Number of Days Experiment Must Run	
Greater Than The Value For The Control Group?	TRUE	TRUE or FALSE																														
Less Than The Value For The Control Group?	TRUE	TRUE or FALSE																														
By at Least (Detection Level)	1.50%	(e.g. 9.25% vs. 10.75%)																														
With a Confidence Level of at Least	95.00%	(5.00% False Positives)																														
and a Power Level of	95.00%	(5.00% False Negatives)																														
Calculated Z Score for Type I Error	1.96	False Positives Intermediate Calculation																														
Calculated Z Score for Type II Error	1.64	False Negatives Intermediate Calculation																														
Minimum Visitors Required (Test+Control Groups)	34.9%	Percentage of Available Visitors Given Segmentation																														
Minimum Visitors Required (Test+Control Groups)	10,475	Number of Required Visitors Given Segmentation																														
Minimum Experiment Time (Days)	10.5	Minimum Number of Days Experiment Must Run																														

	% of Visitors	Visitors	
Test group	50%	15,000	Set % Test Group Calculates Number of Visitors
Control group	50%	15,000	Set % Control Group Calculates Number of Visitors
Total (Usually 100%)	100%	30,000	Visitors Available For Control and Test Groups
Test Group Accuracy (at Target Confidence Level)		0.48%	Based On CI How Accurately Fall Into Test Group
Control Group Accuracy (at Target Confidence Level)		0.48%	Base On CI How Accurately Fall Into Control Group
Z Score (at Target Accuracy)		4.33	Intermediate Calculation for Reference
Actual Confidence Level (at Target Interval)		100.00%	Experimentally Achieved Confidence Level
Actual Interval (at Target Confidence Level)		0.68%	Experimentally Achieved Detection Level
Other Design Notes:			
<i>For Use with Percentage of Visitors Metrics Only</i>			

3. Enter text or values for all of the fields in blue in this file, which are described in the following table. The calculated fields are defined in the second table.

Fields to Define

In this field...	Specify
Experiment Title	A descriptive name for your experiment.
Experiment Description	A textual description of the experiment.
Metric Being Studied	The name of the metric on which the experiment is based. Example: Visitor Conversion
Metric Definition	The definition of the metric on which the experiment is based. Format: Visitors[X]/Visitors Example: <code>Visitors[URI='conversionpage.asp']/Visitors</code>
Intended Start Time	The date and time you want the experiment to begin.
Intended End Time	The date and time you want the experiment to end.
Applicable Selections	(Optional) The dimension name and element set or range by which you want to further segment the dataset.
Experiment URIs	The URIs involved in your hypothesis. You define the current URIs for the control group and the alternate URIs that you have created or will create for the test group(s).
Expected Metrics for Application Selections	Heading for the metric values that you expect for your website.
Average Visitors per Day	The average number of visitors to your website per day.
Visitor Conversion	The average visitor conversion rate for your website.
Experiment Will Determine if the metric name for the Test Groups is ...	Heading for how the metric values should be compared.
Greater Than The Value For the Control Group?	Set this field to True if you want the ability to conclude that the test group's metric increased during the experiment. Set this field to False to reduce the number of visitors needed to draw conclusions. Adobe recommends that you set it to True.
Less Than The Value For the Control Group?	Set this field to True if you want the ability to conclude that the test group's metric decreased during the experiment. Adobe recommends that you set it to True.
By at Least (Detection Level)	The percentage by which you want the metric for the test group to be higher or lower than that for the control group.

In this field...	Specify
With a Confidence Level of at Least	The desired confidence level for the test group values. The confidence level determines the number of false positives to measure the probability that the stated expectation is true.
and a Power Level of	The desired power level for the test group values. The power level determines the number of false negatives.
% of Visitors	Heading for the percent of visitors values.
Test Group	Percent of visitors you want to include in the test group. You can play with this number until the value in the Total (Usually 100%) field in the Visitors section is equal to or greater than the value in the Minimum Visitors Required (Test+Control Groups) field, both of which are described in the following table.
Control Group	Percent of visitors you want to include in the control group.
Other Design Notes	Any notes that you want to save for future reference.

The remaining fields are calculated based on the values that you entered and are described in the following table.

Calculated Fields

Field	Description
Expected Metrics for Application Selections	Heading for the metric values that you expect for your website.
Expected Visitors per Period	This field is normally automatically calculated by the spreadsheet. It relies on the assumption that on most days the website receives many more new visitors than return visitors. If this is not the case, this cell's calculation should be overwritten with the actual number of visitors expected during the experiment.
Calculated Z Score for Type I Error	The Z score for a false positive result. This is an intermediate statistical calculation.
Calculated Z Score for Type II Error	The Z score for a false negative result. This is an intermediate statistical calculation.
Minimum Visitors Required (Test+Control Groups)	Minimum number of visitors needed in your experiment to meet your specified confidence level, power level, and Z score, expressed as a percentage of the value in the Expected Visitors per Period field.
Minimum Visitors Required (Test+Control Groups)	Minimum number of visitors needed in your experiment to meet your specified confidence level, power level, and Z score. This value must be less than or equal to the value in the Total (Usually 100%) field in the Visitors section.
Minimum Experiment Time (Days)	Minimum number of days you need to run your experiment to meet your specified confidence level, power level, and Z score. This calculated number is subject to the same issues as discussed in the Expected Visitors per Period field. In the case of a website with many returning visitors, the Minimum Experiment Time (Days) field is the expected number of days it takes to see a number of unique visitors equal to the value in the Minimum Visitors Required field.
Visitors	Heading for the visitors values.
Test Group	Number of visitors needed in the test group.
Control Group	Number of visitors needed in the control group.
Total (Usually 100%)	Total number of visitors needed for the experiment. This value must be equal to or greater than the value in the Minimum Visitors Required (Test+Control Groups) field.
Test Group Accuracy (at Target Confidence Level)	Percentage indicating that there is a chance equal to the specified confidence level that the measured value of the metric calculated for the test group will be within this percentage of its true value.
Control Group Accuracy (at Target Confidence Level)	Percentage indicating that there is a chance equal to the specified confidence level that the measured value of the metric calculated for the control group will be within this percentage of its true value.
Z Score (at Target Accuracy)	Number of standard deviations a given value is from the test mean.
Actual Confidence Level (at Target Interval)	The confidence level achieved for the experiment. The confidence level measures the probability of the stated expectation to be true.

Field	Description
Actual Interval (at Target Confidence Level)	The confidence interval achieved for the experiment, which provides an estimated range of values that is likely to include an unknown population parameter. This range is calculated from a given set of sample data.

You need to look at the value in the Minimum Visitors Required (Test+Control Groups) field . . .

Minimum Visitors Required (Test+Control Groups)	34.9%
Minimum Visitors Required (Test+Control Groups)	10,475
Minimum Experiment Time (Days)	10.5

and compare it to the value in the Total field in the **Visitors** column.

	% of Visitors	Visitors
Test group	50%	15,000
Control group	50%	15,000
Total (Usually 100%)	100%	30,000

For your experiment to be statistically valid, the value in the Total (Usually 100%) field must be equal to or greater than the value in the Minimum Visitors Required (Test+Control Groups) field.

Given the inputs provided, what the example worksheet shows you is that 10,475 visitors need to participate in this experiment to achieve the entered 95% confidence rate (which is the minimum suggested confidence for any controlled experiment, although you can increase this number). The experiment as currently designed includes 30,000 visitors, which is well over the minimum number of visitors required.

If you keep the number of days the same, you could increase the confidence level as long as the total number of visitors continues to meet or exceed the required minimum.

- Save the file for your records and then use the information from the file to configure the experiment using the experiment configuration spreadsheet. For more information about this spreadsheet, see *Configuring and Deploying the Experiment* on page 774.

Server Administration and Installation

The Server Installation and Administration Guide is written for administrators responsible for installing and administering Insight servers. This guide provides you with specific guidelines and instructions for installing Adobe server products and services such as Insight Server, Insight Server clusters, **Transform**, and more.

Server System Requirements

Identify minimum requirements and recommendations for Data Workbench server components before planning and implementing your system.

DPU Requirements

The server Data Processing Unit (DPU) is the main data processing component of Data Workbench.

It listens for network connections from Data Workbench, reads raw source data from the File Server Unit (FSU) and uses substantial computational and storage resources.

Licensed Capacity

Please refer to the Services Description in the *Adobe Data Workbench (Insight) Service Agreement* for license capacity information.



Note: For *MS System Center Endpoint Protection* in Windows 2012 Servers, these executables need to be added to the **Excluded Processes**:

- `InsightServer64.exe`
- `ReportServer.exe`
- `ExportIntegration.exe`

DPU System Recommendations and Requirements

Adobe provides recommendations regarding a Data Workbench design that meets your business needs. However, the following guidelines are useful when selecting the operating system (OS) and hardware, because the optimized nature of the DPU software places specific requirements on the OS/hardware platform.

If a single dataset is limited by the capacity or speed of a single DPU, you can cluster them. For example, suppose you have three licensed copies of the DPU software that are used together to more quickly run a larger dataset. As the data is divided between the machines evenly, the licensed capacity of the dataset is multiplied by three. In addition, the processing speed per row becomes three times faster than a single DPU.

To achieve the best performance from your DPU investment, Adobe recommends the following high-performance components described in the following table:

Server DPU Requirements

	Required	Recommended
Operating System	Microsoft Windows Server 2008 x64	Microsoft Windows Server 2012 x64 Microsoft Windows Server 2016 x64
CPU	See recommendations.	Latest-generation 4-core+ processors from Intel or AMD are recommended. For optimal performance, 8-cores; for a trade-off between speed and cost, 4-cores are recommended.
RAM	8 GB	12 GB
Working Data Storage	1TB+ of total logical temp storage. Low latency access to the disk sub-system	For temporary storage Adobe recommends either: <ul style="list-style-type: none">• (4 to 8) * (750GB or higher) SATA HDDs (3.5" spindle.)• (6 to 10) * (300GB or higher) SATA HDDs (2.5" spindle.) These should be configured in a JBOD array. Alternatively, when gross disk capacity exceeds 2TB, an array of 2-disk RAID1 volumes can be used. For example, configure six disks as a 3*(2*750GB RAID 1 pair.)

	Required	Recommended
System Data Storage	Additionally, Adobe requires high-availability storage of a modest size (20GB) for the OS, DPU software, and other system software.	
Clustering Hardware	See recommendations.	Use a homogenous set of servers. In a DPU cluster, the slowest server reduces the performance of the whole dataset.
Clustering Network Performance	A switched-gigabit Ethernet connection or greater.	

Alternative Disk Subsystems

When considering alternative disk subsystems for temp storage, consider the following factors and guidelines:

- The DPU is unusually demanding of a high performance disk system, so setting up an inadequate disk subsystem can cause performance bottlenecks.
- The DPU software does its own performance-oriented data striping across a set of JBOD disks. Do not use RAID to increase speed.
- Adobe recommends that the DPU has 400+ MB/s aggregate sustained bandwidth to the disks.
- Average read sizes are very high (2MB+). For this reason 15K or 10K SAS disks often perform a little better (or worse) than SATA disks at a significant cost and capacity penalty.
- Avoid using a SAN architecture. Experience shows that the cost to get a SAN to perform at the levels required is usually extreme.
- The local disk subsystem is used as scratch space—no data is permanently lost from a HDD failure, so consider avoiding costly, slower, high-availability systems.

Speed Considerations

Adobe cannot provide a warranty or representation concerning the speed at which data is processed by a configured Data Workbench, because a variety of factors impact the data processing speed, including but not limited to the following:

- Number of rows of data
- Number of dimensions (columns) of data
- Number and complexity of custom processing steps
- Use of clustering
- Speed of hardware

File Server Unit Requirements

The server's File Serving Unit (FSU) is the main data storage and management component of Data Workbench.

The FSU acts as a file server for raw source data to the DPU, and, when appropriate, coordinates the clustering of DPUs. Each FSU is licensed to supply source data to up to five (5) DPUs.

FSU Components	Recommendations
Operating System, CPU, RAM	These requirements are the same as those of the DPU. However, for the FSU, Adobe recommends using the minimum requirements rather than following the recommendations.
Disk System	<p>Adobe recommends:</p> <ul style="list-style-type: none"> • (12 or more) * (750GB or greater) SATA HDDs in a RAID 5/6 configuration. • High-performance SAN connection supporting 100MB/s+ sustained bandwidth.

FSU Components	Recommendations	
work with you to determine your exact requirements.	As the FSU holds the raw source data, any loss would be unrecoverable, and Adobe suggests backing up this data on a regular basis.	
Network Performance	Adobe requires switched-gigabit Ethernet connections between FSUs and DPUs working together.	

Sensor Requirements

Data Workbench **Sensor** collects event data from web, application, and data collection servers to be transmitted to any server.

Sensor's instrumentation ensures consistently accurate measurement of events that occur in your Internet channel. **Sensor** supports many combinations of Web server software and operating system.

Sensor System Recommendations

The following table describes system recommendations for **Sensor**:

Features	Recommended
Disk Storage	512 MB minimum.
RAM	32 MB of RAM must be available to Sensor on the HTTP or other server computer that is its host.
Network Performance	1 Mbps or greater network connection to a repeater server or data workbench server. Sensor typically consumes far less bandwidth than one (1) Mbps. Your Adobe consultants will help you estimate the actual amount of bandwidth that would be required on a routine basis.
Network Ports and Firewalls	Sensor connects to the data workbench server using HTTPS (typically port 443, though this is configurable) or HTTP (typically port 80, though this is configurable). The appropriate port on any firewall that resides between a Sensor and the target data workbench server or repeater server should be opened only between the respective Sensor hosting computer and the data workbench server or repeater server before beginning the Sensor installation process. Sensor makes a uni-directional HTTPS or HTTP connection to a data workbench server or repeater server.
Network Management Systems	Existing network management systems should monitor the health of the underlying computer hardware (for example, disk space, network service) and network connectivity as well as the Windows Event Log or UNIX syslog.
Server Time Synchronization	Ensure that the computer system time is continuously synchronized across every computer that hosts a Sensor . The Web server applications and computers that are monitored by Sensor must have synchronized system times for the event data collected from them to be accurate. Please refer to your operating system's documentation for steps to synchronize system times on an ongoing basis with NTP or other such time synchronization facility.
DNS Name Usage	Adobe recommends that Sensors use a DNS name (instead of an IP address) to resolve the network address of a data workbench server or repeater server. When a Sensor uses a DNS name, the host web server's DNS or local hosts file needs to be configured to resolve the name of the data workbench server or repeater server.

Support Server Software

The following table lists the most common combinations that **Sensor** supports:

Support Server Software

Web Server Software	Operating System
Apache Server / IBM HTTP Server 2.2	Microsoft Windows Server 2003 or later; RedHat Enterprise Linux 6.x or later; Sun Solaris 8.x or later; IBM AIX 5.1x or later.
Apache Server 2.4	RedHat Enterprise Linux 6.x or later
Microsoft IIS	Microsoft Windows Server 2003 or later
Java Application Servers (Tomcat, JBoss, iPlanet, Weblogic)	Microsoft Windows Server 2003 or later; RedHat Enterprise Linux 6.x or later; Sun Solaris 8.x or later; IBM AIX 5.1x or later.

For other server and operating system combinations, please consult Adobe regarding availability. Not all features of **Sensor** are available with all combinations of web/application server and operating system. For more information about particular **Sensor** releases, please contact Adobe Support.

Report Server Requirements

Data workbench report server is the component that allows the output of scheduled reporting.

The reports that are output can either be in the form of .PNG images or .XLS spreadsheets placed in a file system, or as emails. Its hardware requirements are identical to the *[data workbench client](#)*.

The following requirements exist for report server:

- Access to file system for output of data (network share, or local drive).
- Access to configured SMTP server.
- Microsoft Excel 2003 or above installed on report server. See *[Considerations for server-side Automation of Office](#)* for additional information.

Network Management

Adobe recommends that existing network management systems monitor the hardware and network that the Data Workbench platform relies on.

In addition, Adobe recommends monitoring the Windows event logs of the FSUs and DPUs, which are written to when an error occurs.



Note: Any networked storage system hosting log files needs to provide at least 10MB per DPU of sustained bandwidth.

Data Availability

It is a normal and required practice for a server DPU to process and re-process data into new or refreshed dataset.

This may occur because of configuration changes, data source changes, hardware changes, inappropriate configuration, hardware failure, software failure, power failure, and so forth. When such processing or re-processing occurs, all dataset and system data is required to be immediately available to the DPU and FSU components. Failure to adhere to this requirement can lead to significant and unnecessary system down time.

DPU and FSU Network Issues

Considerations to keep in mind when working with DPU and FSU networks.

-
- For networked log file distribution, any networked storage system hosting log files needs to provide at least 10MB per DPU of sustained bandwidth.
 - The DPU, FSU, and Data Workbench inter-communicate bi-directionally via HTTP or HTTPS on port 80 or 443 (by default; ports can be alternatively configured).
 - Data Workbench **Sensor(s)** must be able to connect (one-way) to the servers.
 - To allow the DPU to send alert messages via SMTP, it must be able to contact the configured SMTP server.
 - Adobe recommends that FSUs and DPUs be given network names such as FSU01.CLIENT.COM to avoid reconfiguration if the case of an IP-address change.

Installing Data Workbench Servers

Navigation title:

Insight Server is licensed per server computer.

In many implementations of Adobe applications, only one installation of Insight Server is required; however, additional Insight Servers may be added to the first to make an Insight Server cluster.

This section describes how to install a single Insight Server Data Processing Unit (DPU) or File Server Unit (FSU). For steps to install and configure an Insight Server cluster, see *Insight Server Clusters* on page 813.

About Insight Server License Units

Navigation title:About Insight Server License Units

Insight Server is available under two license types.

The following are the two license types:

- *Data Processing Unit (DPU)* on page 792
- *File Server Unit (FSU)* on page 792

Data Processing Unit (DPU)

This type of Insight Server can process, store, and serve data from an Adobe dataset. It optionally can store the log files that contain the source data from which the dataset is constructed, or it can receive that data from an Insight Server File Server Unit (FSU). A DPU is the type of Insight Server with which Insight and Report clients interact directly.

If you are installing an Insight Server DPU, see *Installation Procedures for an Insight Server DPU* on page 792.

File Server Unit (FSU)

This type of server is configured to receive and store event data from one or more **Sensor** or event data replication instances (**Repeater** functionality provided with a special use license) and stream the data to one or more Insight Server Data Processing Units (DPUs) for constructing Adobe datasets. DPUs communicate with an FSU using a protocol that optimizes the transfer of event data to the DPU and is significantly faster than maintaining log files on ordinary file servers. The use of an FSU also reduces hardware costs by enabling log data to be stored on lower cost storage hardware and reduces administrative complexity by allowing multiple **Sensors** to point to a single Insight Server.

If you are installing an Insight Server FSU, see *Installation Procedures for an Insight Server FSU* on page 813.

Installation Procedures for an Insight Server DPU

Navigation title:Installation Procedures for an Insight Server DPU

Detailed instructions for installing an Insight Server DPU and configuring it for administrative use.

To install and configure an Insight Server DPU, you must complete the following tasks in order:

1. Install the Insight Server program files. See *Installing the Insight Server Program Files* on page 793.

-
2. Download and install the Insight Server digital certificate. See [Downloading and Installing the Digital Certificates](#) on page 793.
 3. Check the port settings in the `Communications.cfg` file. See [Checking the Port Settings](#) on page 802.
 4. Modify the `Access Control.cfg` file to allow administrative access to Insight Server from Insight. See [Updating the Access Control File](#) on page 803.
 5. Modify the `server.address` file to define the server's network location. See [Defining the Server's Network Location](#) on page 804.
 6. (Optional) Modify the `Disk Files.cfg` file to specify where processed data is stored. See [Configuring the Location of the Dataset \(`temp.db`\)](#) on page 808.
 7. Install the profiles and lookup files. See [Installing Profiles and Lookup Files](#) on page 809.
 8. Set the Microsoft Windows memory utilization parameters.
 9. Register Insight Server as a Windows service. See [Registering Insight Server as a Windows Service](#) on page 811.

Installing the Insight Server Program Files

Navigation title:Installing the Insight Server Program Files

To install the Insight Server software, you download the program files from the location provided by Adobe. Typically, the software is provided in a compressed file from an FTP site.

1. On your Windows machine, create a directory in which to install Insight Server.

Example: C:\Adobe\Server

2. Go to the download site provided by Adobe and locate the installation file for Insight Server.
3. Extract the contents of the installation file to the directory you created above.

For a list of the directories and files installed with Insight Server, see [Insight Server Directory Structure](#) on page 870.

Downloading and Installing the Digital Certificates

Navigation title:Downloading and Installing the Digital Certificates

General information about digital certificates, and procedures to download and install them.

- [Understanding Digital Certificates](#) on page 793
- [Node-locked Certificates](#) on page 794
- [Current Certificates](#) on page 794
- [Using Digital Certificates on Machines Without Internet Access](#) on page 794
- [Digital Certificate Installation Procedures](#) on page 795

Understanding Digital Certificates

Adobe uses X.509 digital certificates to identify and authenticate the client and server components that make up an implementation.

When you install a server component (Insight Server or **Repeater**), you must install the digital certificate that Adobe has issued for the component. If you need to migrate your Adobe application to another machine, you must obtain a new certificate from Adobe. To do so, contact Adobe Customer Care.

The common name that appears on this certificate identifies the server by a specified domain name (for example, `vs001a.mycompany.com`). When a server client connects to this server, the server presents this certificate as proof that it is, indeed, the server that the client requested.

Similarly, when you install a server client (for example, Insight or Report) you must install the digital certificate that authorizes a named individual (for example, Jane Smith) to use the installed client application. If you need

to migrate your Adobe application to another machine or another named user, you must obtain a new certificate from Adobe. To do so, contact Adobe Customer Care.

The client application presents this digital certificate to gain access to a server component. An administrator of the server component can restrict access to server resources based on the common name or organizational unit values that appear in the client's certificate.

The X.509 digital certificates installed with Adobe applications also enable its client and server components to exchange information over Secure Sockets Layer (SSL). SSL secures transmissions over HTTP using a public-and-private key encryption system. Adobe's implementation of SSL supports 1024-bit RSA keys and uses a 128-bit RC4 encryption algorithm.

In addition to security, the digital certificates that you install also function as license keys that enable you to run the installed Adobe software. To function properly, a digital certificate must be node-locked and current, or the application does not start.

String Encryption

See [String Encryption](#) on page 795 for encrypting passwords.

Node-locked Certificates

A node-locked certificate is a digital certificate that has been registered to the machine on which it is installed. Node locking permanently associates a certificate with a specific node identifier (a value that uniquely identifies a particular machine). To node lock your certificate, your machine must have Internet access to the Adobe License Server or to a proxy server that has access to the License Server.

If you are installing on a machine that cannot access the Internet, you must obtain and install a special pre-locked certificate as described in [Using Digital Certificates on Machines Without Internet Access](#) on page 794.

If you are installing on a machine that can access the Internet, your digital certificate is node-locked automatically the first time that you start your Adobe product. After being node-locked, the certificate cannot be used on any other machine. If you need to migrate your Adobe product to another machine, you must obtain a new, unlocked certificate from Adobe.

Current Certificates

Besides being node-locked, a digital certificate must be current. To remain current, your certificate must be revalidated on a regular basis (generally every 30 days, but can vary depending on your agreement with Adobe). If your machine has Internet access, the revalidation process is completely transparent. Your Adobe product automatically connects to the License Server and revalidates the certificate when necessary. If your machine does not have Internet access, you need to manually install updated certificates as described in the following section.

Using Digital Certificates on Machines Without Internet Access

If you are installing on a machine that cannot access the Internet, you must request a pre-locked certificate for your installation of Insight Server. A pre-locked certificate is a digital certificate that Adobe manually locks to the node identifier for the machine.

To request a pre-locked certificate, you need to send the node identifier and your certificate number to Adobe Customer Care. To obtain the node identifier for your machine, contact Adobe Customer Care to request the Adobe Node Identifier utility. You also can obtain the node identifier from the alert that the Adobe software issues when it attempts to connect to the License Server and cannot.

When you receive the pre-locked certificate, install it as described in the last two steps of [Digital Certificate Installation Procedures](#) on page 795. When the certificate needs to be revalidated, you must download a new, validated certificate from the License Server and reinstall it on your machine.

Digital Certificate Installation Procedures

To download and install the digital certificate

1. Open your web browser to <https://aap.adobe.com>.



Note: Your browser might prompt you to present a digital certificate at this point. If it does, simply click **Cancel** to dismiss the dialog box.

2. On the login screen, enter the **Account Name** and the **Password** that you received from Adobe, then click **login**.
3. Locate the certificate that has been issued for your Insight Server, then click the icon associated with that certificate.



Note: Make a note of the common name that is assigned to this certificate. You use this name in a later step.

4. When prompted to save the certificate, click **Save**. (Note that the name of the file matches the common name associated with the certificate.)
5. Download the file to the **Certificates** folder in the directory where you installed Insight Server. This folder already contains a certificate file named `trust_ca_cert.pem`. This certificate file must always be present.
6. Rename the downloaded certificate file to:

```
server_cert.pem
```

String Encryption

Encrypt passwords and other strings when communicating between the client and server.

When communicating between the Data Workbench client (workstation) and server, you can save a Value parameter (such as a password) with the Type of *EncryptedString*. This hides the parameter and saves the string to the *Windows Credential Store* on the server with the corresponding key returned. This primarily stores credentials used in exports but can be used to encrypt any parameter.

- A new folder was added at `Server\EncryptStrings`.

This is where you set the configuration file to encrypt strings.

- A new configuration file was added at `Server\Component\EncryptedStrings.cfg`.

```
component = EncryptionComponent:  
Path = Path: EncryptStrings\\*.cfg
```

This file polls the `Server\EncryptStrings` folder for encryption configuration files.

To encrypt a string:

1. Create an `EncryptedStrings.cfg` configuration file for a string with these fields set:

```
Names = vector: 1 items  
0 = NameEncryptValuePair:  
EncryptValue = EncryptedString: // left empty as input then output will  
be filled by server  
Name = string: // Name for identifier  
Value = string: // Value to be encrypted
```

- *Value* - This field contains the plain text string that needs to be encrypted.

This is server-side encryption only. The *Value* setting is encrypted only on the server computer.

- *Name* - This field contains a value identifying the encrypted string.

-
- *EncryptValue* - This field will be left empty in the input configuration file. The encrypted value will be returned in this field.

You can add multiple **NameEncryptValuePair** values for different fields for encryption.



Note: All empty Value fields will be removed.

2. Save the **EncryptedStrings.cfg** file to the Server\EncryptStrings folder.

Output file

An output file will be generated with the same name as the input file with a *<filename>.encrypted* extension. For example, if the input file is named *sample.cfg* then the output file will be named *sample.cfg.encrypted*.

Using Custom Certificates in Data Workbench

Instructions for using custom certificates.

A certificate used by either the Data Workbench client or server needs to be signed by a trusted CA (Certificate Authority). Data Workbench customers receive certificates that are signed by the Visual Sciences CA. These certificates are trusted by the Data Workbench software, since the *trust_ca_cert.pem* (provided along with the Insight software and stored in the **Certificates** directory of both servers and clients) contains a *Root CA Certificate* for the Visual Sciences CA. These certificates are used for both licensing of the software and authentication when clients and servers communicate with each other using SSL. Only certificates issued by the Visual Sciences CA can be used for licensing, but other certificates may be used for communication and authentication. Certificates issued by CAs other than Visual Sciences are referred to below as *custom certificates*.

Important note: For servers and clients, Data Workbench software uses the certificate files installed in the client or server's **Certificates** directory or certificates explicitly identified in its configuration. However, you can also use the [Windows Certificate Store](#) on page 798 for clients.

The following instructions describe the procedures to be followed to use custom certificates for communication between Data Workbench clients and servers. Not every detail is a hard requirement and different variations in the process can be employed. However, the procedures below have been tested to work.

Setting up Custom Client Certificates

1. Add the certificate of the issuing CA to the *trust_cert_ca.pem*, which is installed in the **Certificates** directory of the client and that of every server in every cluster that is to be accessed using this custom certificate.
2. Obtain a custom certificate for each server in the cluster with the following conditions:
 - a. Certificate is formatted as a *.pem* certificate.
 - b. Certificate contains its key and is unencrypted (i.e., it has no password/pass phrase).

A certificate contains its key with one of the following lines:

```
BEGIN PRIVATE KEY  
BEGIN RSA PRIVATE KEY
```

One way to remove the password phrase from a *.pem* certificate:

```
openssl rsa -in password-protected-cert.pem -out no-password-cert.pem  
openssl x509 -in password-protected-cert.pem >> no-password.pem
```

- c. Certificate has the CN, O, OU, etc. as required for this client in the servers' *Access Control.cfg* file.
- d. Certificate was issued with a *purpose* *** of *client* (or both *server and client*).

To verify that a certificate has a purpose code of server and/or client, the following commands can be used:

```
openssl verify -CAfile trust_ca_cert.pem -purpose sslserver -x509_strict  
custom_communications_cert.pem  
openssl verify -CAfile trust_ca_cert.pem -purpose sslclient -x509_strict  
custom_communications_cert.pem
```

For a server certificates, both commands should yield:

```
custom_communications_cert.pem: OK
```

For a client certificate, only the second command is required to yield OK.

3. Place the certificate in the client's **Certificates** directory.
4. In `Insight.cfg` under the `serverInfo` for each cluster that you want to use this certificate, make sure the `custom client cert` is named, such as:

```
Servers = vector: 1 items  
0 = serverInfo:  
    SSL Client Certificate = string: my_custom_client_cert.pem
```

Setting up Custom Server Certificates

This section assumes that you have a cluster that is up and running, using Visual Sciences issued certificates, and the configuration follows common practices (such as the *Components for Processing Servers* directory on the master gets synchronized to the *Components* directories of all DPUs).

1. Add the certificate of the issuing CA to the `trust_cert_ca.pem` which is installed on every server in the cluster and every client that needs to communicate with this cluster.
2. Obtain a custom certificate for each server in the cluster with these requirements:
 1. Custom certificate is formatted as a `.pem` certificate.
 2. Certificate contains its key and is unencrypted (i.e., it has no password/pass phrase).

A certificate contains its key if it has a line such as:

```
BEGIN PRIVATE KEY  
BEGIN RSA PRIVATE KEY
```

One way to remove the password phrase from a `.pem` certificate:

```
openssl rsa -in password-protected-cert.pem -out no-password-cert.pem  
openssl x509 -in password-protected-cert.pem >> no-password.pem
```

3. Certificate has the same CN as the `server_cert.pem` currently installed on the server.
4. Certificate was issued with a purpose of `server` and `client`.

To verify that a certificate has a purpose code of server and/or client, the following commands can be used:

```
openssl verify -CAfile trust_ca_cert.pem -purpose sslserver -x509_strict  
custom_communications_cert.pem  
openssl verify -CAfile trust_ca_cert.pem -purpose sslclient -x509_strict  
custom_communications_cert.pem
```

For a server certificates, both commands should yield:

```
custom_communications_cert.pem: OK
```

For a client certificate, only the second command is required to yield OK.

-
3. Install each server's custom certificate in the **Certificates** directory of the server as `custom_communications_cert.pem`.
 4. Using a text editor, add the following line to **Communications.cfg** file in both the *Components* and *Components for Processing Servers* directories, directly below the first line (`component = CommServer`):
`Certificate = string: Certificates\\custom_communications_cert.pem`
 5. Restart all servers.

About Certificate Failure Warning

When the Insight server or client is looking for a **license** certificate in the **Certificates** directory, it tries to validate all the certificates (except `trust_ca_cert.pem`), against a hard coded copy of the Insight CA certificate, which fails on any custom certificate present in the directory. The server issues this warning:

```
Certificate failed to verify. Error 20 at 0 depth. Desc: unable to get local  
issuer certificate. Cert details:
```

This warning can be safely ignored.

Windows Certificate Store

The Windows Certificate Store allows you to store the client's certificate and private key in the Windows Certificate Store for SSL communication with servers.

The Windows Certificate Store for the Client is a new feature that allows you to store the SSL communication certificate and private key in the Windows Certificate Store rather than in `Insight/Certificates/<CertName>.pem` file. Using the Windows Certificate Store may be preferable if you use the certificate store for other applications and wish to do certificate management in one place, or for users that enjoy the additional Windows audit logging that the Windows Certificate store provides.



Note: Licensing with the license server is still maintained using the existing `<Common Name>.pem` file, and that the certificate obtained from the certificate store will only be used for communication to the servers that you specify.

Prerequisites

1. You must have access to the `certmgr.msc` file with the ability to import a certificate and key into the **Personal** store. (This should be true by default for most Windows users.)
2. The user doing the configuration must have a copy of the **OpenSSL** command-line tool.
3. The server and client must already be configured to use a custom SSL certificate, as described in [Using Custom Certificates in Data Workbench](#) on page 796, giving instructions to store the client certificate in the Windows certificate store rather than storing it in the **Certificates** directory.

Configuring the Windows Certificate Store

The Windows Certificate Store for Clients is enabled following these steps:

Step 1: Import the user's SSL certificate and private key into the Windows Certificate Store.

In [Using Custom Certificates in Data Workbench](#) on page 796 you are directed to put the SSL certificate and key in the following directory:

```
<DWB Install folder>\Certificates\
```

The certificate's name is `<Common Name>.pem` (such as `Analytics Server 1.pem`(not the `trust_ca_cert.pem` file.))

Before the certificate and private key can be imported, they must be converted from `.pem` format to a `.pfx` format, such as `pkcs12.pfx`).

-
1. Open a command prompt or terminal and navigate to the directory:

```
<CommonName>.pem c: cd \<DWB Install folder>\Certificates
```

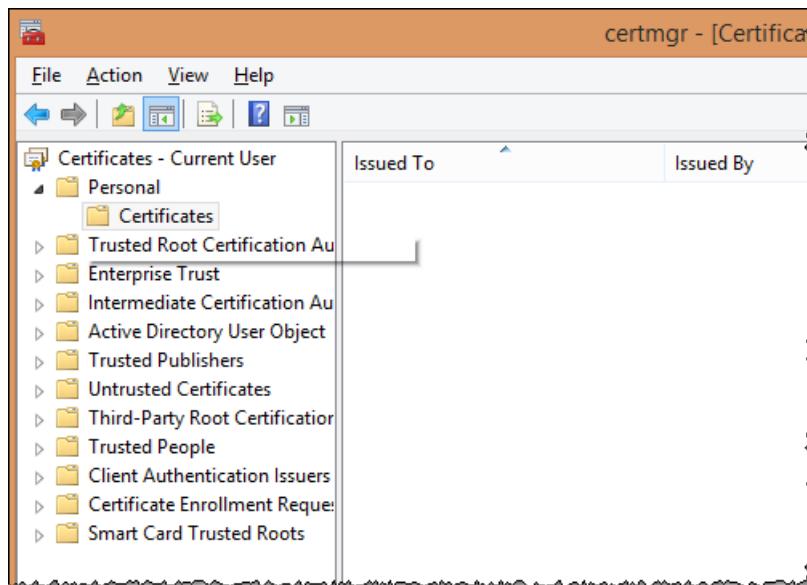
2. Run openssl with the following arguments (with the actual .pem file name):

```
openssl pkcs12 -in "<Common Name>.pem" -export -out "<Common Name>.pfx"
```

If prompted, hit **Enter** to skip entering an export password.

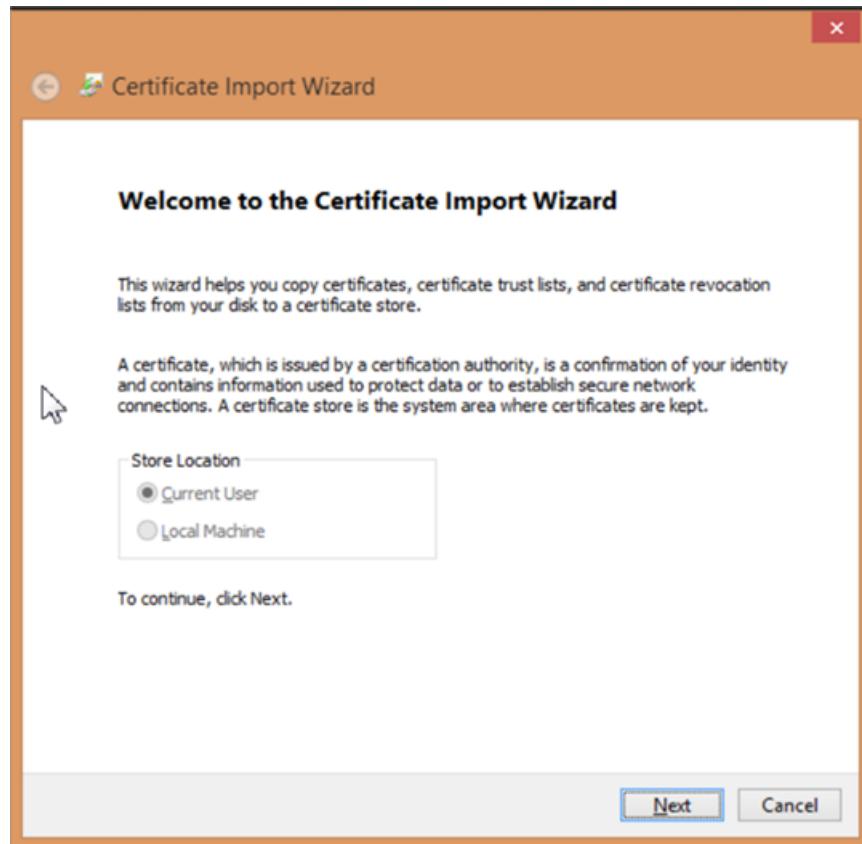
3. Run certmgr.msc from the run prompt, start menu, or command line.

4. Open the **Personal** certificate store for the current user.

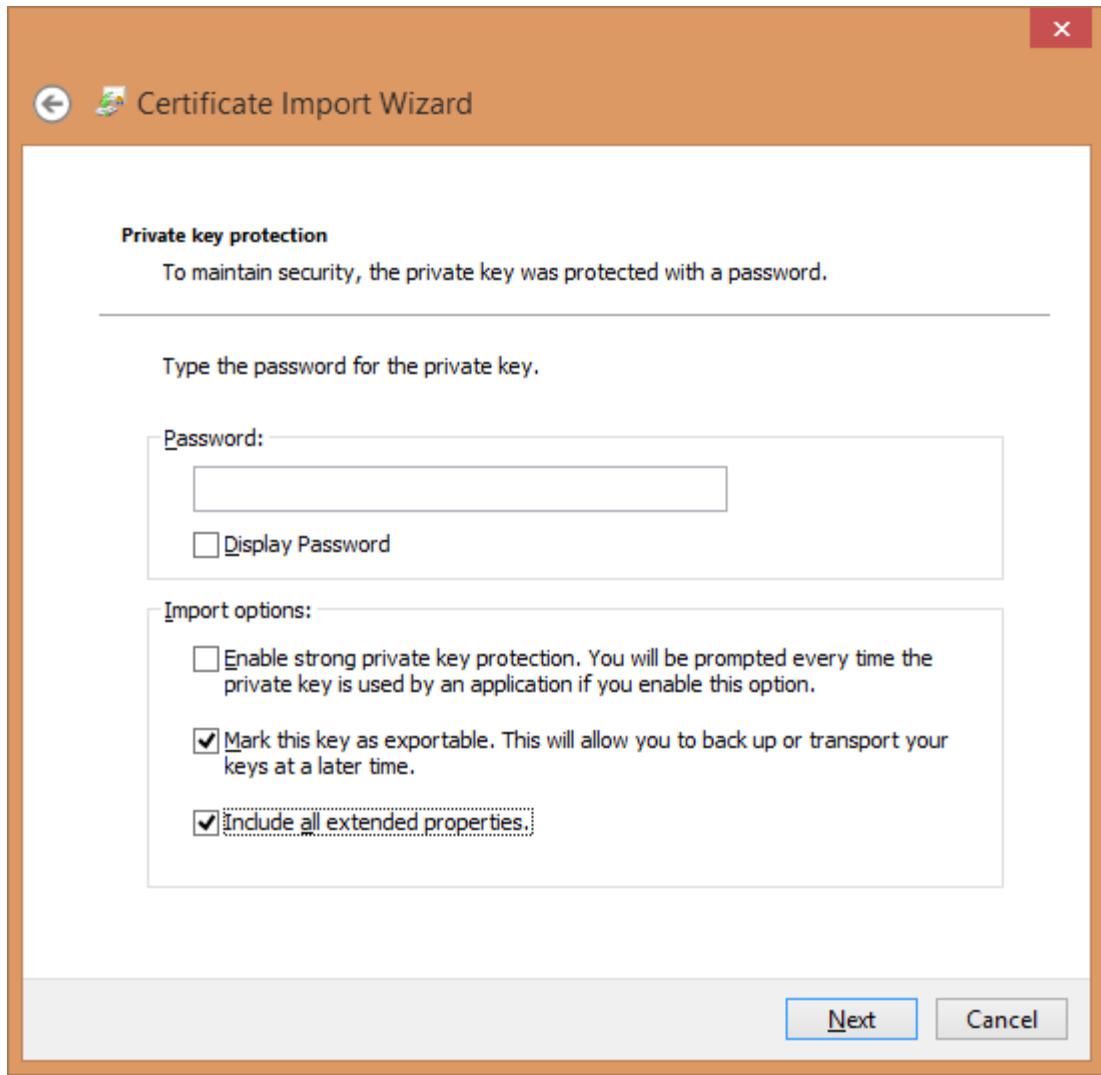


5. Right-click **Certificates** and click **All Tasks > Import**.

Make sure the **Current User** option is selected, then click **Next**.



6. Click **Browse** and select the <CommonName>.pfx file you created previously. You will have to change the file extension dropdown box from an X.509 Certificate to either **Personal Information Exchange** or to **All Files** in order to see it.
Select the file and click **Open**, and then **Next**.
7. Do not enter a password, and make sure that only the options **Mark this key as exportable** and **Include all extended properties** are selected.



Click **Next**.

8. Make sure that **Place all certificates in the following store** is selected, and that the certificate store listed is **Personal**. (If you're an advanced user, you can select another store at this point, but you'll have to change the configuration later.)
9. Click **Next** and then click **Finish**. You should see a dialog box telling you that the import was successful and see your certificate in the Certificates folder of the store.



Note: Pay particular attention to the **Issued To** and **Issued By** fields. You'll need these in the next step.

Step 2: Edit the Insight.cfg file.

The `Insight .cfg` file must be edited in order to direct Data Workbench to use the Windows Certificate Store feature. Each server entry in this file must have some additional parameters specified. If the parameters are omitted, the workstation will default to using the existing certificate configuration. If the parameters are specified but have incorrect values, the workstation will enter an error state and you will have to refer to the log file for error information.

-
1. Open the **Insight.cfg** file (located in the **Insight** installation directory).
 2. Scroll down to the server entry that you wish to configure. If you wish to use the Windows Certificate store for every server, you have to make these modifications to every entry in the vector of **serverInfo** objects.
 3. Add these parameters to their **Insight.cfg** file. You can do this from the workstation, or manually by adding the following parameters to the **serverInfo** object. (Be sure to use spaces instead of tab characters, and don't make other typographical or syntax errors in this file.)

```
SSL Use CryptoAPI = bool: true  
SSL CryptoAPI Cert Name = string: <Common Name>  
SSL CryptoAPI Cert Issuer Name = string: Visual Sciences, LLC  
SSL CryptoAPI Cert Store Name = string: My
```

The boolean enables or disables the feature. The certificate name matches **Issuer To** in the certificate manager. The certificate issuer name matches **Issued By**, and the **Store Name** must match the certificate store name.



Note: The name "Personal" in the Certificate Manager (certmgr.msc) actually refers to the certificate store named **My**. Consequently, if you import your SSL communication certificate and key (.PFX) into the **Personal** certificate store as recommended, you must set the **SSL CryptoAPI Cert Store Name** string to "My". Setting this parameter to "Personal" will not work. This is a peculiarity of the Windows certificate store.

A full list of the predefined system stores can be obtained here:

[https://msdn.microsoft.com/en-us/library/windows/desktop/aa388136\(v=vs.85\).aspx](https://msdn.microsoft.com/en-us/library/windows/desktop/aa388136(v=vs.85).aspx). Your system might have additional certificate stores. If you wish to use a store other than "Personal" (such as **My**), you must obtain the canonical name of the certificate store and provide it in the **Insight.cfg** file. (The system store name "My" is inconsistently referred to as "My" and "MY" by the Windows documentation. The parameter does not appear to be case sensitive.)

4. After you have added these parameters and verified that the values match the list in the Windows Certificate Manager, save the **Insight.cfg** file.

You can now start the workstation (or disconnect/reconnect to the server). Data Workbench should load your certificate and key from the certificate store and connect normally.

Log output

When a certificate is not found or is invalid, this error message is thrown to the **HTTP.log** file.

```
ERROR Fatal error: the cert could not be found!
```



Note: The L4 logging framework can be enabled by setting up the **L4.cfg** file (see your account manager to set this up).

Checking the Port Settings

Navigation title: Checking the Port Settings

By default, Insight Server listens on ports 80 (for HTTP) and 443 (for HTTPS).

If these ports are already allocated by another process on the machine where you have installed Insight Server, use the following procedure to change Insight Server's port assignments.

To change the port assignments

1. Navigate to the **Components** folder in the directory where you installed Insight Server.
Example: C:\Adobe\Server\Components
2. Open the **Communications.cfg** file in a text editor such as Notepad.

-
3. Locate the Port and the SSL Port entries, as shown below:

```
component = CommServer:  
Access Control File = Path: Access Control\\Access Control.cfg  
Access Log Directory = string: P:\\Audit\\  
IP Interface = string:  
Port = int: 80  
SSL Port = int: 443  
Servers = vector: 17 items  
    0 = InitServer:  
        Client Type = string: Sensor  
        URI = string: /SensorInit.vsp  
    . . .
```

4. If these are not the ports that you want Insight Server to use, change the port assignments, then save and close the file.

Updating the Access Control File

Navigation title: Updating the Access Control File

The Access Control.cfg file manages access to certain features in Insight Server.

It defines entities called AccessGroups. An AccessGroup identifies a group of users that have permission to use certain features of the server.

Before you can connect to Insight Server with Insight, you must update the Administrators AccessGroup to include one of the Insight licenses that Adobe has issued to your organization. This AccessGroup identifies users that are permitted to perform administrative functions through Insight.

The following procedure describes how to add a license to the Administrators AccessGroup. To complete this task, you must determine which Insight license has administrative privileges for your organization. (For initial set-up and configuration, granting administrative privileges to a single license is sufficient. You can grant administrative privileges to additional licenses later.) You also need to know the “common name” assigned to this license. To obtain this value, you can examine the license certificates for your account at <https://aap.adobe.com>.

The purpose of this procedure is simply to identify a licensed copy of Insight that you can use to initially set up and configure Insight Server. Once you identify this license, you can perform all subsequent server configuration (including additional AccessGroup configuration) using the licensed copy of Insight. For additional information about controlling access to the server using AccessGroups, see *Configuring Access Control* on page 827.

To update the access control file

1. Navigate to the Access Control folder in the directory where you installed Insight Server.

Example: C:\\Adobe\\Server\\Access Control

2. Open the Access Control.cfg file in a text editor such as Notepad.
3. Locate the CN entry in the Administrators AccessGroup and replace the existing value of this entry with the common name that identifies the Insight that you will use to initially set up and administer Insight Server. The following file fragment illustrates where you insert the common name in the Access Control.cfg file.

```
Access Control Groups = vector: 5 items  
0 = AccessGroup:  
    Members = vector: 2 items  
        0 = string: IP:127.0.0.1  
        1 = string: CN: CommonName  
    Name = string: Administrators  
    Read-Only Access = vector: 0 items  
    Read-Write Access = vector: 1 items  
        0 = string: /
```

```
1 = AccessGroup:
```

```
. . .
```

If you are using credentials-based authentication, a few extra entries will be available for configuration. These entries are:

- O (Organization ID): This entry represents the ID of the organization. For example, 1 = string: O:46F582D4582596B40A45491@ExampleOrg. This ID can be found in the Admin Console.
- PLC – This entry allows access to the users provisioned for a particular product configuration. It can be used in format Organization_Id-PLC. For example, 1 = string: PLC:46F582D4582596B40A45491@ExampleOrg-DataworkbenchAdminUsers. The users provisioned for Data Workbench using the PLC DataworkbenchAdminUsers will get access on their servers.
- Email – This entry allows access to any individual user. Its value should be the email address of the provisioned user. For example, 1 = string: Email:kim@exampleorg.com.



Note:

- The entries are case sensitive. You must ensure that the values specified for O, PLC, and Email are exactly the same as the ones shown in the Admin Console.
- Type the common name exactly as it appears on the certificate.
- Do not use the Tab key to generate whitespace in the Access Control.cfg file (or in any other configuration file for an Adobe component). Use only spaces to create whitespace. A tab character prevents the system from reading the file correctly.

4. Save and close the file.

Defining the Server's Network Location

Navigation title:Defining the Server's Network Location

Insight Server's clients (Report and Insight) use common names to refer to Insight Servers.

For example, when you connect Insight or Report to an Insight Server, you identify the server by its common name (for example, Server.MyCompany.com). Internally, the client resolves the common name to a numeric IP address before sending a request to the server.

To resolve common names to IP addresses, Insight Server's clients use a local lookup file called the address file. The address file lists the common names of Insight Servers installed at your organization and identifies their numeric IP addresses. A client automatically receives a copy of the address file when it opens a profile on the Insight Server.

When a client issues a request to an Insight Server, it attempts to resolve the server's common name through the address file. If the address file identifies an IP address for the requested server, the client routes the request to the specified address. If the address is not defined (for example, the address file does not define the server's common name), the request fails. Optionally, you can configure clients to resolve addresses through the operating environment's normal Domain Naming Service (DNS) mechanism if a name is not defined in the client's address file. For instructions, see the Parent parameter in the [Network Locations](#) on page 804 in the following section.

Network Locations

Navigation title:Network Locations

Conceptually, the address file serves the same purpose as the ETC\HOSTS file on a networked machine.

However, unlike the HOSTS file, which describes a single collection of names, the address file contains multiple collections of names called network locations.

A network location is a named collection of address definitions. Each address definition in the collection associates a common name with an IP address.

In the address file, a network location is defined in a structure called a NetworkLocation. The NetworkLocation in the following example defines a network location called “MyCorporate Intranet.” It contains an address definition that maps the common name VS01.myCompany.com to the IP address “10.2.1.70.”

```
0 = NetworkLocation:  
    Addresses = vector: 1 items  
        0 = AddressDefinition:  
            Address = string: 10.2.1.70  
            Name = string: VS01.myCompany.com  
        Name = string: MyCorporateIntranet  
        Parent = string:
```

As shown in the example above, the NetworkLocation structure consists of three main parameters:

NetworkLocation Parameters

Parameter	Description
Addresses	Defines zero or more AddressDefinitions. Each AddressDefinition associates a common name with an IP address.
Name	Assigns a name to the NetworkLocation. The name assigned to a NetworkLocation must be unique within the address file.
Parent	<p>Specifies the name of another NetworkLocation whose members are included in this NetworkLocation. This parameter enables one NetworkLocation to extend another.</p> <p>You can set the Parent parameter to “DNS,” to extend a NetworkLocation to the client’s normal DNS system.</p> <p>Example: Parent = string: DNS</p> <p>When DNS is the parent, clients attempt to resolve a common name using the client machine’s DNS system when they cannot resolve the name through the NetworkLocation.</p>

Multiple IP Addresses for an Insight Server

Navigation title:Multiple IP Addresses for an Insight Server

If clients can reach an Insight Server through multiple networks (for example, through the corporate intranet and through the Internet), the address file must define a separate network location for each of the server’s IP addresses.

For example, if server VS01.myCompany.com has an IP address of 10.2.1.70 on an internal network, and an IP address of 65.196.125.167 on the Internet, the address file would include a network location for each of the addresses as illustrated in the example below:

```
0 = NetworkLocation:  
    Addresses = vector: 1 items  
        0 = AddressDefinition:  
            Address = string: 10.2.1.70  
            Name = string: VS01.myCompany.com  
        Name = string: MyCorporateIntranet  
        Parent = string:  
1 = NetworkLocation:  
    Addresses = vector: 1 items  
        0 = AddressDefinition:  
            Address = string: 65.196.125.167  
            Name = string: VS01.myCompany.com  
        Name = string: Internet  
        Parent = string:
```

When users connect to an Insight Server, they use the NetworkLocation parameter (in the client user interface) to specify the network location through which they want the server’s common name resolved. For example, given an address file with the two NetworkLocations shown above, a user would set the NetworkLocation parameter to “MyCorporate Intranet” to connect to Insight Server through the internal network and to “Internet” to connect to the server through the Internet.

The Address File Installed on Insight Server

Navigation title:The Address File Installed on Insight Server

The address file installed on Insight Server contains four pre-defined network locations.

The procedure in the next section describes how to configure this file.

```
Locations = vector: 4 items
0 = NetworkLocation:
    Addresses = vector: 1 items
        0 = AddressDefinition:
            Address = string:
            Name = string:
            Name = string:
            Parent = string:
    1 = NetworkLocation:
        Addresses = vector: 0 items
        Name = string: Insight
        Parent = string:
    2 = NetworkLocation:
        Addresses = vector: 0 items
        Name = string: Insight Server
        Parent = string:
    3 = NetworkLocation:
        Addresses = vector: 0 items
        Name = string: Report Server
        Parent = string:
```

- NetworkLocation 0 is an empty, unnamed network location that you edit to associate the common name of your Insight Server to its IP address. If the server has multiple IP addresses, you create additional NetworkLocations.
- NetworkLocation 1 is the Insight network location. If you do not explicitly set the NetworkLocation parameter, Insight resolves common names through this network location.
- NetworkLocation 2 is the Insight Server network location. When Insight Servers operate in a cluster, they use this network location to resolve common names for inter-server communication.
- NetworkLocation 3 is the Report Server network location. If you do not explicitly set the NetworkLocation parameter, Report resolves common names through this network location.

To Configure the Address File

The following procedure describes how to configure the address file to define a network location (or network locations) for your Insight Server.

1. Navigate to the Addresses folder in the directory where you installed Insight Server (for example, C:\Adobe\Server\Addresses).
2. Locate the server.address file and rename this file to reflect the server's common name. For example, if the common name were server.mycompany.com, you would rename the file server.mycompany.com.address.
3. Open the renamed file in a text editor such as Notepad.
4. Edit NetworkLocation 0 to specify the common name and IP address of the Insight Server as shown below. If your server has multiple IP addresses, use NetworkLocation 0 to specify the server's IP address on the local, non-routable network (for example, its location on the internal network).

```
Locations = vector: 3 items
0 = NetworkLocation:
    Addresses = vector: 1 items
        0 = AddressDefinition:
            Address = string: IPAddress
            Name = string: CommonName
            Name = string: NetworkLocationName
            Parent = string:
```

For this value...	Specify
<i>IP Address</i>	The numeric IP address of the Insight Server machine. Example: 192.168.124.176
<i>Common Name</i>	The common name assigned to the digital certificate for Insight Server. Example: <code>server.mycompany.com</code> Note: Be sure to type this name exactly as it appears on the certificate.
<i>Network Location Name</i>	The name you want to assign to the collection of common names and IP addresses that this NetworkLocation represents. The name must be unique within the address file. Example: Corporate Intranet

5. If your Insight Server has additional IP addresses, create an additional NetworkLocation for each address. (An easy way to do this is to make a copy of the NetworkLocation you created above and update the IP address in the copy.)

You can add the new NetworkLocation to the end of the address file or insert it between existing NetworkLocation definitions. (The position of a NetworkLocation within the address file is not significant; however, the Insight, Insight Server, and Report Server NetworkLocations are typically placed at the end of the file.)

After you have added the necessary NetworkLocations, do the following to renumber the items in the file:

1. Update the item count for the Locations structure to match the total number of NetworkLocation definitions in the file. For example, if your file contains four NetworkLocation definitions, the Locations line would look as follows:

```
Locations = vector: 4 items
```

2. Update the NetworkLocation item numbers so that NetworkLocations are numbered consecutively (starting from 0).

For an example of an address file that defines an Insight Server with two IP addresses, see the example in this section.

6. In the Insight and Report Server network locations, edit the Parent parameter as shown below to specify the name of the NetworkLocation that Insight and Report use as their default network locations. (For an example of what the Parent parameter looks like when it is configured, see the example in this section.)

```
1 = NetworkLocation:
 Addresses = vector: 0 items
 Name = string: Insight
 Parent = string: ClientDefaultNetworkLocation

3 = NetworkLocation:
 Addresses = vector: 0 items
 Name = string: Report Server
 Parent = string: ClientDefaultNetworkLocation
```

If your Insight Server has a single IP address and, therefore, has only one NetworkLocation, point the Parent parameter to that NetworkLocation. If your Insight Server has multiple IP addresses, point the Parent parameter to the NetworkLocation that defines the address to which your Insight and Report clients connect most frequently.

7. In the Insight Server network location, edit the Parent parameter as shown below to point to the NetworkLocation that the server uses to resolve common names of other Insight Servers when it operates in

a cluster. (Although this network location is not used unless an Insight Server operates in a cluster, it is a good practice, even in a single server configuration, to point the Parent parameter to a network location that identifies the server's internal IP address.)

```
2 = NetworkLocation:  
    Addresses = vector: 0 items  
    Name = string: Insight Server  
    Parent = string: ServerDefaultNetworkLocation
```

The following example shows a completed address file. This file defines five network locations.

- NetworkLocation items 0 and 1 define network locations named “MyCorporateIntranet” and “Internet.” These network locations define two different IP addresses for a server named VS01.myCompany.com.
- NetworkLocation item 2 is the Insight network location. This is the default network location used by Insight. In this example, the Insight network location inherits its AddressDefinitions from the “Internet” NetworkLocation.
- NetworkLocation item 3 is the Insight Server network location. This is the default network location Insight Server uses when it communicates with other servers in a cluster. In this example, the Insight Server network location inherits its AddressDefinitions from the “MyCorporate Intranet” NetworkLocation.
- NetworkLocation item 4 is the Report Server network location. This is the default network location used by Report. In this example, the Report Server network location inherits its AddressDefinitions from the “Internet” NetworkLocation.

```
Locations = vector: 5 items  
0 = NetworkLocation:  
    Addresses = vector: 1 items  
        0 = AddressDefinition:  
            Address = string: 10.2.1.70  
            Name = string: VS01.myCompany.com  
        Name = string: MyCorporateIntranet  
        Parent = string:  
  
1 = NetworkLocation:  
    Addresses = vector: 1 items  
        0 = AddressDefinition:  
            Address = string: 65.196.125.167  
            Name = string: VS01.myCompany.com  
        Name = string: Internet  
        Parent = string:  
  
2 = NetworkLocation:  
    Addresses = vector: 0 items  
    Name = string: Insight  
    Parent = string: Internet  
  
3 = NetworkLocation:  
    Addresses = vector: 0 items  
    Name = string: Insight Server  
    Parent = string: MyCorporateIntranet  
  
4 = NetworkLocation:  
    Addresses = vector: 0 items  
    Name = string: Report Server  
    Parent = string: Internet
```

Configuring the Location of the Dataset (temp.db)

Navigation title: Configuring the Location of the Dataset (temp.db)

By default, Insight Server writes its dataset (temp.db) to the same drive as the Insight Server program files.

For example, if you install Insight Server on drive C, it writes the dataset to drive C.

If you want Insight Server to maintain the dataset on a different drive, or if the amount of data you expect to collect requires the use of multiple drives, you must update the `Disk Files.cfg` file to specify where you want Insight Server to write the `temp.db` file.

To configure the location of `temp.db`

1. Navigate to the `Components` folder in the directory where you installed Insight Server.

Example: `C:\Adobe\Server\Components`

2. Open the `Disk Files.cfg` file in a text editor such as Notepad.

By default, this file contains a single entry in the Disk Files structure as shown below.

```
component = DiskSpaceManagerComponent:  
  Disk Files = vector: 1 items  
    0 = string: Temp\\temp.db  
  Detect Disk Corruption = bool: true
```

3. To change the location of `temp.db`, modify the Disk Files definition. The following example illustrates how you would edit the configuration to spread the `temp.db` file across drives C, D, and E:

```
component = DiskSpaceManagerComponent:  
  Disk Files = vector: 3 items  
    0 = string: C:\\Temp\\temp.db  
    1 = string: D:\\Temp\\temp.db  
    2 = string: E:\\Temp\\temp.db  
  Detect Disk Corruption = bool: true
```



Note: Note the use of the double backslashes in the file names above. In Insight Server configuration files, the backslash character is an escape character. It is used to express special control sequences (for example, \t for a tab character) in text. To represent an actual backslash character, you must type the backslash twice (for example, \\) to override the escape function. This applies only when editing configuration files in a text editor such as Notepad.

Installing Profiles and Lookup Files

Navigation title:Installing Profiles and Lookup Files

The profiles and lookup files that Adobe has developed for your particular application are internal profiles that provide the metrics, dimensions, and workspaces that enable the analysis of your dataset.

As with all other internal profiles provided by Adobe, these profiles should not be changed. All customization must occur in your dataset or role-specific profiles or other profiles that you create.

Adobe distributes the profile and lookup files for your application as a `.zip` file. Each zip file is named for the application whose profile and lookup files that it contains. (For example, `Site52.zip` contains the profile files for Site v5.2.) The `.zip` file contains two folders (`Lookups` and `Profiles`).



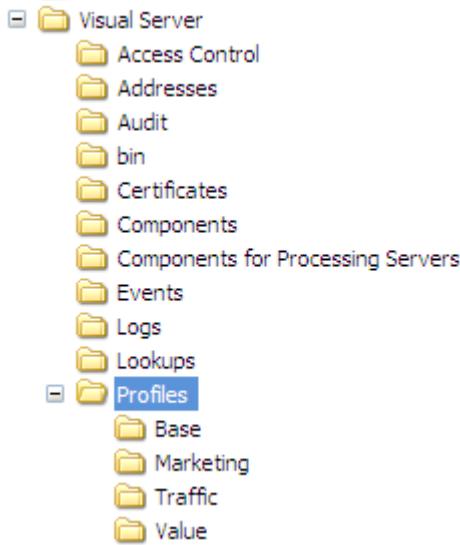
Note: If you do not already have the installation file containing the profiles and lookup files for your application, download them from the Adobe FTP site before you begin.

You must install the profile and its lookup files on the Insight Server machine on which you process and run your dataset profile. If you are running an Insight Server cluster, you must install the files on the master server. For information about dataset profiles, see the *Dataset Configuration Guide*.

To install profiles for your Adobe application

1. Open the `Profiles` folder from the `.zip` file provided to you by Adobe.

-
2. Copy all of the folders within the `Profiles` folder in the `.zip` file to the `Profiles` folder in your Insight Server installation directory. You want to end up with `...\\Profiles\\<internal profile name>` folders on your Insight Server as shown in the following example. Your actual profile names may differ.



3. Navigate to the `Profiles\\<dataset profile name>` folder in the directory where you installed Insight Server and locate the `profile.cfg` file in this directory.



Note: If you are installing profiles for the first time, you can use the provided Sample profile as your dataset profile. You can find the `profile.cfg` file (it might be named something like `profile.cfg.offline`) for the Sample profile within the `Profiles\\Sample` folder in your Insight Server installation directory.

4. Open the `profile.cfg` file using a text editor such as Notepad and do the following:

- Add entries for the internal profiles in the `Directories` vector. The profile names correspond to the names of the directories that you copied to the `Profiles` folder on the Insight Server machine.
- Update the number of directories as appropriate.
- Add the common name of the server to the `Common Name` line in this file, as highlighted below:

```
Profile = profileInfo:  
Directories = vector: n+1 items  
  0 = string: Base\\\\  
  1 = string: internal profile name 1\\\\  
  2 = string: internal profile name 2\\\\  
  . . .  
  n = string: internal profile name n\\\\  
Processing Servers = vector: 1 items  
  0 = ProfileServerInfo:  
    Common Name = string: serverCommonName  
    Server = string:
```

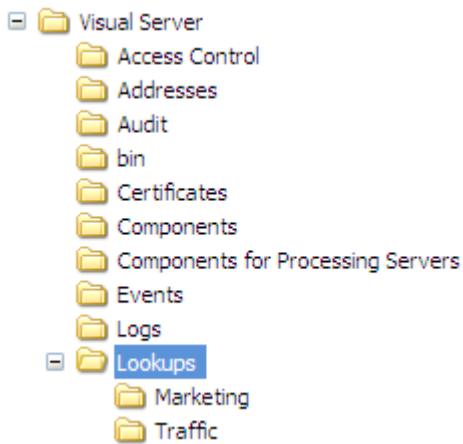


Note: The `serverCommonName` that you specify for the `Common Name` in the `profile.cfg` file corresponds to the server common name for the Insight Server machine on which you are processing and running the dataset profile. For instructions to update `profile.cfg` so that the dataset profile runs on an Insight Server cluster, see [Insight Server Clusters](#) on page 813.

5. Save the file. Be sure to save the file as `profile.cfg` if it was named differently.

To install the lookup files for your Adobe application

1. Open the Lookups folder from the .zip file provided to you by Adobe.
2. Copy all of the folders within the Lookups folder in the .zip file to the Lookups folder in your Insight Server installation directory. You want to end up with ...\\Lookups\\<internal profile name> folders on your Insight Server as shown in the following example. Your actual profile names may differ.



Registering Insight Server as a Windows Service

Navigation title: Registering Insight Server as a Windows Service

Procedure to start Insight Server and simultaneously register it as a Microsoft Windows Service.



Note: When you start Insight Server for the first time, it automatically connects to the Adobe License Server to register your digital certificate. To complete the registration process successfully, your machine must be connected to the Internet when you execute the following steps.

To start Insight Server and register it as a Windows Service

1. Open a command prompt and navigate to the bin sub-directory in the folder where you installed Insight Server.
Example: C:\\Adobe\\Server\\bin
2. At the command prompt, execute the following command to start Insight Server and simultaneously register it to run as a service under Microsoft Windows:
`InsightServer64.exe /regserver`
3. To confirm that Insight Server is running correctly, click **Start > Control Panel > Administrative Tools > Services**. This command sequence can vary depending on which version of Windows you are using.
 - a. In the service list, locate the entry for **Adobe Insight Server** and confirm that its status is Started and its startup type is Automatic.
 - b. Close the Services control panel.
4. To check whether Insight Server experienced any errors during start-up, click **Start > Control Panel > Administrative Tools > Event Viewer**. This command sequence can vary depending on which version of Windows you are using.
 - a. In the left pane of the **Event Viewer** window, select the **Application** log.
 - b. In the right pane, look for events with “Adobe” in the **Source** column.
 - c. If you find an error from “Adobe,” double-click the error to display the **Event Properties** window. This window provides detailed information about the error.

-
- When you finish examining the **Applications** log, close the Event Viewer.

You have completed the installation of Insight Server. Insight Server is designed to run continuously. If you restart the machine, Insight Server restarts automatically. If you need to start and stop Insight Server manually, you can do so using the Services control panel in Windows. As described in the following section, you can optionally configure Insight Server service to automatically restart periodically.

Configuring the Service to Restart Automatically

Insight Server is designed to continue running uninterrupted. It is typically stopped or started only when performing infrequent tasks such as software upgrades or certificate changes or in the event of certain system errors. It is not necessary to stop or restart the Insight Server service during normal system operation; however, you can configure the service to restart periodically (daily, weekly, or monthly) to, for example, clear the event messages.

To configure the Insight Server service to restart automatically

- Navigate to the Components folder in the directory where you installed Insight Server.

Example: C:\Adobe\Server\Components

- Use a text editor such as Notepad to create a new file called `ScheduledRestart.cfg`.
- Enter the following text in the `ScheduledRestart.cfg` file:

```
component = ScheduledRestart:  
Start Time = string: Month DD, YYYY HH:MM:SS TZone  
Restart Every = string: frequency
```

For this value...	Specify
<code>Month DD, YYYY HH:MM:SS TZone</code>	The time when you want Insight Server to be restarted for the first time. Example: August 13, 2013 22:30:00 EST  Note: You must specify a time zone. The time zone does not default to system time if not specified. If you wish to implement Daylight Saving Time or a similar clock-shifting policy, you must save the <code>.dst</code> file containing the appropriate rules in the <code>Base\Dataset\Timezone</code> directory on the Insight Server machine. For a list of supported time zone abbreviations and information about implementing Daylight Saving Time, see Time Zone Codes on page 874.
<code>frequency</code>	One of the following values: <ul style="list-style-type: none">• month• week• day To indicate the frequency at which you want Insight Server to be restarted after the initial time specified in Start Time. For example, if you wanted Insight Server to restart once a week, you would set this value to "week."

- Save the `ScheduledRestart.cfg` file.

Verify that the `ScheduledRestart.cfg` file is in the Components folder in the directory where you installed Insight Server.

Installation Procedures for an Insight Server FSU

Navigation title:Installation Procedures for an Insight Server FSU

The instructions for installing an Insight Server FSU and configuring it for administrative use are very similar to those for installing and configuring an Insight Server DPU.

For *MS System Center Endpoint Protection* in Windows 2012 Servers, these executables need to be added to the **Excluded Processes**:

- `InsightServer64.exe`
- `ReportServer.exe`
- `ExportIntegration.exe`

To install and configure an Insight Server FSU, you must complete the following tasks:

1. Install the Insight Server program files.
2. Install the Insight Server digital certificate.
3. Check the port settings in the `Communications.cfg` file.
4. Modify the `Access Control.cfg` file to allow administrative access to the Server from the Client.
5. Modify the `server.address` file to define the server's network location.
6. Set Windows memory utilization parameters as described.
7. Register Insight Server as a Windows service as described.

For instructions to configure data sources, permissions, and communications for an Insight Server FSU, see the *Dataset Configuration Guide*.

Next Steps

Navigation title:Next Steps

Instructions to complete installation of your Adobe application.

- Install at least one copy of Insight as described in the *Insight User Guide*.
- If you want to run Insight Server in a clustered configuration, set up the cluster as described in *Insight Server Clusters* on page 813.
- On the servers from which you want to collect data, install **Sensors** as described in the *Insight Sensor Guide*.
- If you have licensed **Transform**, see *Transform* on page 845 for instructions to install it on an Insight Server FSU.
- If you want to set up the Replication Service for Insight Server, see *Installing the Replication Service* on page 847 for instructions to install the service on an Insight Server FSU.
- If you have licensed **Repeater**, see *Repeater Functionality* on page 850 for installation instructions.

Insight Server Clusters

Navigation title:Insight Server Clusters

General information about Insight Server clusters and installing Insight Server clusters.

About Insight Server Clusters

Navigation title:About Insight Server Clusters

An Insight Server cluster is required when the amount of data you want to process and make accessible to your users of Insight and Report exceeds the capacity of a single Insight Server.

By setting up an Insight Server cluster, you can distribute a single analysis dataset across multiple machines in a cluster to harness the processing power of multiple Insight Servers.

The first step in the implementation of an Insight Server cluster is to allocate the Insight Server machines in your cluster. The first Insight Server machine that you set up is your master Insight Server (sometimes referred to as your primary Insight Server).



Note: If you are using an Insight Server File Server Unit (FSU), Adobe recommends that you configure the FSU as your master Insight Server. For information about configuring an FSU, see the *Dataset Configuration Guide*.

The master Insight Server manages the communication between the other Insight Servers in the cluster (called processing servers or, sometimes, query servers) and instances of Insight and Report. For a given dataset, log file processing occurs on the (one or more) designated Insight Servers (master or processing) as specified in the Insight Server configuration files. When working in a clustered environment, Insight installations are configured to access the master Insight Server, but queries can be handled by any of the Insight Servers within the cluster.



Note: The **PAServer.cfg** file. If you want to submit Predictive Analytics clustering jobs to Insight Servers, then you will need to configure the **PAServer.cfg** file for handling server-side clustering submissions. The custom profile should inherit the **PAServer.cfg** from the Predictive Analytics profile (**Server\Profiles\Predictive Analytics\Dataset**). Set a *Master Server* in this file and save the **PAServer.cfg** to the implementation site.

```
PAServer = PAServerConfig:  
    Master Server = serverInfo:  
        Address = string:  
        Port = int: 80  
        Use SSL = bool: false
```



Important: The instructions in this chapter do not apply to the creation of an Insight Server cluster consisting of more than five (5) Insight Servers. Please contact Adobe to obtain system requirements and profile configuration recommendations for clusters larger than five Insight Servers.

Installing an Insight Server Cluster

Navigation title:Installing an Insight Server Cluster

Steps to install an Insight Server cluster.

To set up a clustered configuration of Insight Server, you perform the following general steps:

- Install the master Insight Server.
- Update cluster-related configuration files on the master Insight Server.
- Install the processing Insight Servers and activate synchronization.

After you complete these installation steps, you can configure a profile to run on a cluster. For instructions, see the [Configuring a Profile to Run on a Cluster](#) on page 821.

Installation Procedures for a Cluster

Information about procedures for installing and configuring clusters, including installing the master server, configuring the master server for clustering, and more.

Installing the Master Insight Server

Navigation title:Installing the Master Insight Server

To use a cluster, you must designate one Insight Server in the cluster to act as the master Insight Server.

A client component such as Insight or Report connects to the master Insight Server at the beginning of a session. At subsequent points during the session, the client might connect to other Insight Servers in the cluster to perform a query. These subsequent connections between the client and the other Insight Servers in the cluster are brokered by the master Insight Server and are transparent to the client.

Besides brokering connections between a client and other Insight Servers in the cluster, the master Insight Server acts as the central administrative point for the entire cluster. When you administer a cluster, you update the master Insight Server. The administrative changes that you make on the master Insight Server are retrieved by the other Insight Servers in the cluster.

To install the master Insight Server

-
1. Determine which machine will act as the master Insight Server.
 2. Install and configure Insight Server (typically, an Insight Server FSU) on this machine as described in [Server Administration and Installation](#) on page 787.
 3. Install Insight and configure a connection to the master Insight Server as described in the *Insight User Guide*.

Configuring the Master Insight Server for Clustering

Navigation title:Configuring the Master Insight Server for Clustering

Information about configuring the cluster on the Master Insight Server, updating the access control file for a cluster, and more.

To configure the cluster, perform the following steps on the master Insight Server:

- Add the processing Insight Servers' common names and addresses to the address file.
- Add all of the Insight Servers to the Cluster Servers group in the `Access Control.cfg` file.
- Update the `Synchronize.cfg` file in the Components for Processing Servers directory to point to the master Insight Server.
- If necessary, modify the `Disk Files.cfg` file in the Components for Processing Servers directory to specify the location of the `temp.db` file on the processing Insight Servers.

To complete these steps, you need to know the common names (as specified on the digital certificates for the individual Insight Server) and the IP addresses of each Insight Server in the cluster. If you do not already have this information, obtain it before you proceed.



Note: The procedures described in this section require Insight. If you have not installed Insight, follow the instructions in the *Insight User Guide* before proceeding.

Adding the Processing Insight Servers to the Address File

Use the following procedure to add the processing Insight Servers' common names and IP addresses to the address file on the master Insight Server. (Although the address file is maintained and administered on the master Insight Server, it is used by all the Insight Servers in the cluster.)



Note: The following assumes that the address file has already been configured for the master Insight Server. If you have not already added the master Insight Server's IP address(es) to the address file, complete the procedure described in [Defining the Server's Network Location](#) on page 804 before you begin.

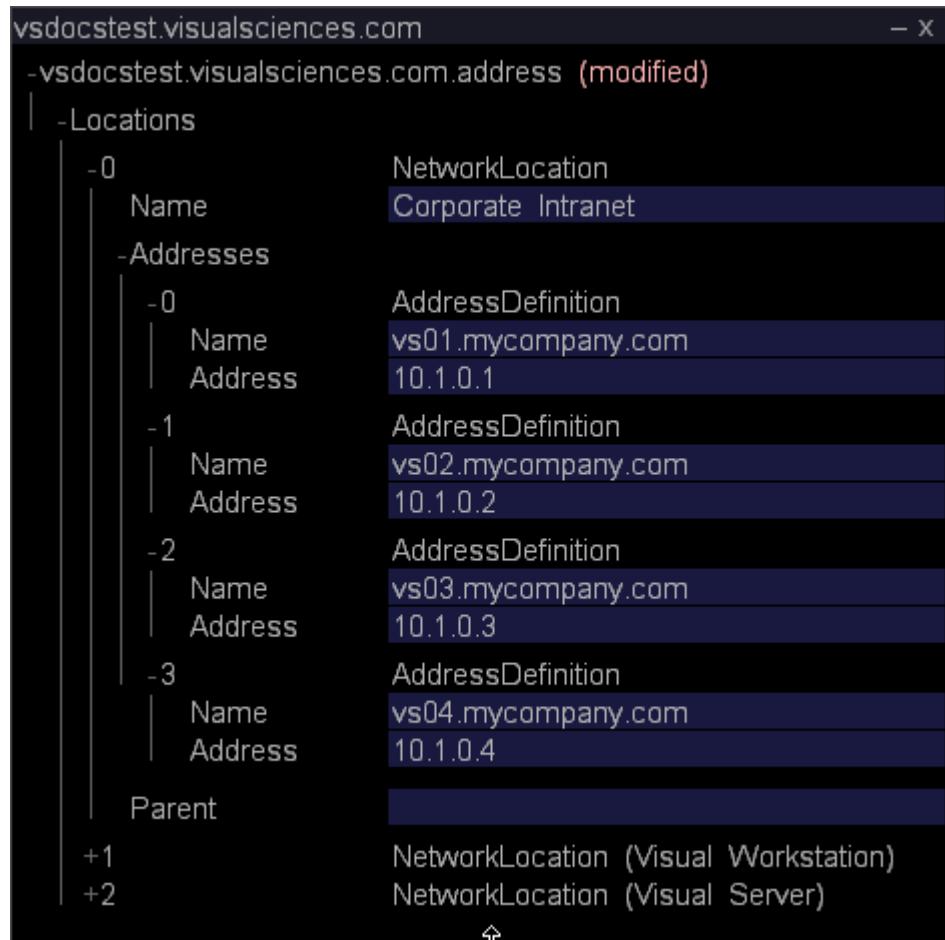
To add the processing Insight Servers to the address file

1. Start Insight and load the Configuration profile (if it is not already open) by right-clicking the title bar and clicking **Switch Profile > Configuration**.
2. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
3. Right-click the icon of the master **Insight Server** and click **Server Files**.
4. In the **Server Files Manager**, open the Addresses directory and do the following to open the Insight Server's address file:
 - a. Right-click the check mark in the *server name* column and click **Make Local**.
 - b. Right-click the check mark in the **Temp** column and click **Open > in Insight**.
5. Expand the contents of the **Locations** structure, then expand NetworkLocation 0, Addresses, and AddressDefinition.
6. Do the following to add an AddressDefinition to NetworkLocation 0 for each processing Insight Server in the cluster:
 - a. Right-click **AddressDefinition** and click **Add New > Address Definition**.

-
- b. In the Name parameter, specify the processing Insight Server's common name.
 - c. In the Address parameter, specify the processing Insight Server's IP address.

You can use an asterisk as a wildcard in the Address field, such as 10.10.116.*, to simplify clustering. See [Understanding Access Levels](#) on page 828.

The following example defines a cluster containing two Insight Servers:



The screenshot shows a configuration interface for a processing Insight Server named "vsdocstest.visualsciences.com". The tree structure is as follows:

- vsdocstest.visualsciences.com.address (modified)
 - Locations
 - 0 NetworkLocation
 - Name Corporate Intranet
 - Addresses
 - 0 AddressDefinition
 - Name vs01.mycompany.com
 - Address 10.1.0.1
 - 1 AddressDefinition
 - Name vs02.mycompany.com
 - Address 10.1.0.2
 - 2 AddressDefinition
 - Name vs03.mycompany.com
 - Address 10.1.0.3
 - 3 AddressDefinition
 - Name vs04.mycompany.com
 - Address 10.1.0.4
 - Parent
 - +1 NetworkLocation (Visual Workstation)
 - +2 NetworkLocation (Visual Server)

- 7. If the servers are connected to multiple networks, repeat Step 6 to add the processing Insight Servers to the NetworkLocations for those networks.

The following example shows a cluster of four Insight Servers attached to two networks ("Corporate Intranet" and "Internet").

vsdocstest.visualsciences.com	
-vsdocstest.visualsciences.com.address (modified)	
-Locations	
-0	NetworkLocation
Name	Corporate Intranet
-Addresses	
-0	AddressDefinition
Name	vs01.mycompany.com
Address	10.1.0.1
-1	AddressDefinition
Name	vs02.mycompany.com
Address	10.1.0.2
-2	AddressDefinition
Name	vs03.mycompany.com
Address	10.1.0.3
-3	AddressDefinition
Name	vs04.mycompany.com
Address	10.1.0.4
Parent	
-1	NetworkLocation
Name	Internet
-Addresses	
-0	AddressDefinition
Name	vs01.mycompany.com
Address	65.196.125.167
-1	AddressDefinition
Name	vs02.mycompany.com
Address	65.196.125.168
-2	AddressDefinition
Name	vs03.mycompany.com
Address	65.196.125.169
-3	AddressDefinition
Name	vs04.mycompany.com
Address	65.196.125.170
Parent	
+2	NetworkLocation (Visual Workstation)
+3	NetworkLocation (Visual Server)

8. Save your changes to the server by doing the following:

-
- a. Right-click (**modified**) at the top of the window and click **Save**.
 - b. In the **Server Files Manager**, right-click the check mark for the file in the **Temp** column and select **Save to > <server name>**.

Updating the Access Control File for a Cluster

To use Insight Servers in a cluster, each Insight Server in the cluster (including the master Insight Server) must belong to the Cluster Servers access control group. The Cluster Servers group identifies the servers (by IP address) that are allowed to participate in the cluster. Although this file is maintained and administered on the master Insight Server, it is used by all of the Insight Servers in the cluster.

To edit the access control file

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of the master Insight Server and click **Server Files**.
3. In the **Server Files Manager**, open the Access Control directory.
4. Do the following to open the **Access Control.cfg** file:
 - a. Right-click the check mark in the *server name* column and click **Make Local**.
 - b. Right-click the check mark in the **Temp** column and click **Open > in Insight**.
5. Expand the Access Control Groups structure, then expand AccessGroup (Cluster Servers).
6. For each Insight Server in the cluster (including the master Insight Server), do the following:
 - a. Right-click **Members** and click **Add New > New Member**.
 - b. Specify the Insight Server's IP address (its numeric IP address, not its name). If the Insight Servers are connected to multiple networks, this AccessGroup should contain only the internal addresses that the Insight Servers use for inter-server communication within the cluster.

The following shows the AccessGroup (Cluster Servers) for a cluster of four Insight Servers.



7. Save your changes to the server by doing the following:

- a. Right-click (**modified**) at the top of the window and click **Save**.
- b. In the **Server Files Manager**, right-click the check mark for the file in the **Temp** column and click **Save to > <server name>**.

Configuring the Synchronization File

You can use the following procedure to configure the central copy of the `Synchronize.cfg` file. The central copy of this file is maintained on the master Insight Server. The processing Insight Servers in the cluster initiate communication with the master Insight Server to retrieve an updated copy of this file.

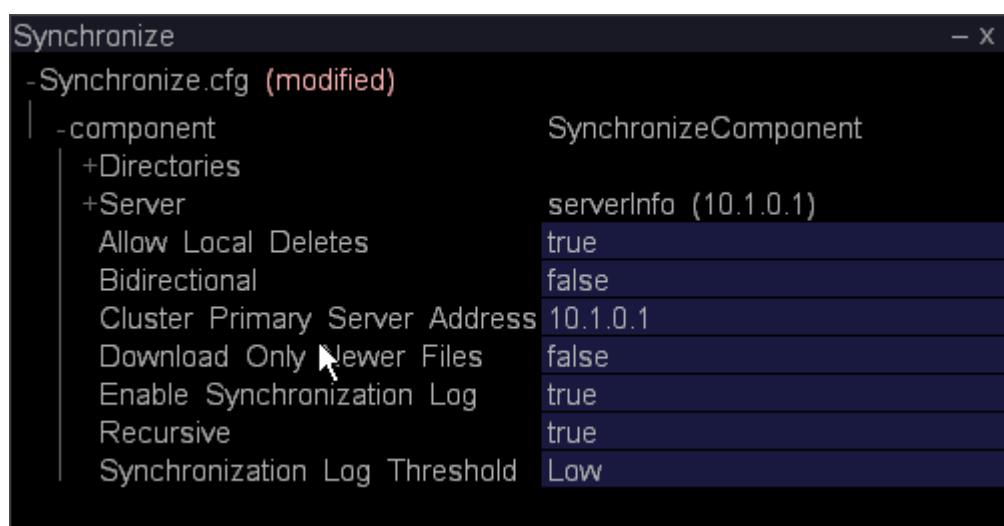
The `Synchronize.cfg` file specifies the location of the master Insight Server. It also identifies the set of administrative files that each of the processing Insight Servers in the cluster retrieves from the master Insight Server. The processing Insight Servers automatically download these files from the master Insight Server when they start. They also dynamically retrieve updated copies of these files from the master Insight Server when the files change.



Note: Although you configure the `Synchronize.cfg` file on the master Insight Server, the master Insight Server itself does not use this file. You update this file on the master Insight Server so that it is properly configured when the processing Insight Servers retrieve the file.

To update the `Synchronize.cfg` file on the master Insight Server

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of the master Insight Server and click **Server Files**.
3. In **Server Files Manager**, open the **Components** for Processing Servers directory.
4. Do the following to open `Synchronize.cfg`:
 - a. Right-click the check mark in the *server name* column and click **Make Local**.
 - b. Right-click the **Temp** check mark and click **Open > in Insight**.
5. Expand the component structure.
6. In the Cluster Primary Server Address parameter, specify the IP address of the master (primary) **Insight Server**.



To create a log that records each time synchronization occurs between the master Insight Server and the processing Insight Servers, make sure that the Enable Synchronization Log parameter is set to “true.”

7. Save your changes to the server by doing the following:
 - a. Right-click (**modified**) at the top of the window and click **Save**.
 - b. In **Server Files Manager**, right-click the check mark for the file in the **Temp** column and click **Save to > <server name>**.

Configuring the Location of the Dataset (temp.db)

Perform the following procedure if you want the processing Insight Servers to maintain `temp.db` (the dataset) in a directory or drive different than the default location or if you want to distribute `temp.db` across multiple drives.



Note: Because the processing Insight Servers all share the same `Disk Files.cfg`, they all must support the file location(s) you specify in this file. For example, if you assign `temp.db` to the E: drive, every processing Insight Server in the cluster must have an E: drive.

To configure the location of temp.db

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of the master Insight Server and click **Server Files**.
3. In the **Server Files Manager**, open the **Components for Processing Servers** directory.
4. Do the following to open `Disk Files.cfg`:
 - a. Right-click the check mark in the *server name* column and click **Make Local**.
 - b. Right-click the check mark in the **Temp** column and click **Open > in Insight**.
5. Expand the `DiskSpaceManagerComponent` structure, then expand the `Disk Files` list.
6. Edit entry 0 to change the location of the `temp.db` file.
7. If you want to distribute `temp.db` across multiple drives, use the steps below to create an additional entry for each additional drive.
 - a. Right-click **Disk Files** and click **Add New > Disk File**.
 - b. In the new entry, specify the location where you want `temp.db` written.

The following shows `temp.db` written across four drives.

-Disk Files.cfg	
-component	DiskSpaceManagerComponent
-Disk Files	
0	E:\Temp\temp.db
1	F:\Temp\temp.db
2	G:\Temp\temp.db
3	H:\Temp\temp.db
Detect Disk Corruption	true
Disk Cache Size (MB)	128

8. Save your changes to the server by doing the following:
 - a. Right-click (**modified**) at the top of the window and click **Save**.
 - b. In **Server Files Manager**, right-click the check mark for the file in the **Temp** column and click **Save to > <server name>**.

Installing and Configuring the Processing Insight Servers

Navigation title:Installing and Configuring the Processing Insight Servers

A processing Insight Server is identical to a master Insight Server except for the contents of its Components directory.

The Components directory on a processing Insight Server contains a special set of files that are specifically configured for processing Insight Servers. These files are derived from the Components for Processing Servers directory on the master Insight Server.

When you install a processing Insight Server, you do the following to set up its Components directory:

1. Delete the original Components directory on the processing Insight Server.
2. Rename the Components for Processing Servers directory to Components.
3. Modify the Synchronize.cfg in the Components directory to point to the master Insight Server.

To install and configure a processing Insight Server

1. Install the Insight Server program files as described in [Installing the Insight Server Program Files](#) on page 793. Be sure to duplicate the directory structure that was used on the master Insight Server. For example, if Insight Server is installed in C:\Adobe\Server on the master Insight Server, you must install it in C:\Adobe\Server on the processing Insight Servers as well.
2. Install the digital certificate that Adobe has issued for this particular processing Insight Server. See [Downloading and Installing the Digital Certificates](#) on page 793 for instructions.
3. Using Windows Explorer, do the following on the processing Insight Server:
 1. Delete the **Components** folder.
 2. Rename the Components for Processing Servers folder to Components.
4. Using a text editor such as Notepad, open the Synchronize.cfg file in the Components directory on the processing Insight Server.
5. Add the IP address of the master (primary) Insight Server to the second line of this file as shown in the following file fragment. Do not edit anything else in this file.

```
component = SynchronizeComponent:
    Cluster Primary Server Address = string: PrimaryIPAddress
    Directories = vector: 7 items
        0 = SynchronizeDir:
            Local Path = string: Profiles\\
            Remote URI = string: /Profiles/
        1 = SynchronizeDir:
            Local Path = string: Lookups\\
            Remote URI = string: /Lookups/
        . . .
```

6. Save the file.
7. Start the Insight Server as described in [Registering Insight Server as a Windows Service](#) on page 811.

You have now completed the installation of your Insight Server cluster. Next, configure the dataset profile to run on the cluster as described in the following section.

Configuring a Profile to Run on a Cluster

Navigation title:Configuring a Profile to Run on a Cluster

When you configure a dataset profile to run on an Insight Server cluster, all of the machines in the cluster share all of the dataset configuration files for that profile.

Therefore, the entries for the parameters in these files must be applicable to all Insight Servers in the cluster. For example, the locations of the log files to be read, the lookup files to be used by Insight Server, and the location of the data output by Insight Server must be the same on all machines in the cluster.

You perform all configuration tasks on the cluster's master Insight Server, which is the Insight Server you use to edit your configuration files. All saved configuration file changes made on the master Insight Server are synchronized automatically to the files on the processing Insight Servers in the cluster.

To run a dataset profile on an Insight Server cluster, you must perform the following processes in the order listed:

1. [Determining Which Insight Servers Process Event Data](#) on page 822
2. [Specifying the Processing Servers in Profile.cfg](#) on page 822
3. (If necessary) [Modifying the Dataset Configuration Files for the Profile](#) on page 823

Determining Which Insight Servers Process Event Data

It is not required that all Insight Servers in the cluster process event data. You can designate one Insight Server in the cluster as a File Server Unit that stores the source files (VSL and log files) and serves the files to all of the Data Processing Units (processing servers) in the cluster. This setup provides the benefit of a single event data repository and leverages the processing power of all the processing servers in the cluster. The processing servers divide the data files among them and guarantee that the same file is not processed more than once.

For more information about designating an Insight Server to run as a File Server Unit, see the Log Processing Configuration File chapter of the *Dataset Configuration Guide*.

If you decide to store source data files on each of the processing servers rather than on a single File Server Unit, you must divide the files equally among the processing servers. Do not store all of the dataset's source files on each of the processing servers. If multiple copies of the same file are available to multiple processing servers, the data is read multiple times (once by each machine) and skews your data.

For help determining which Insight Servers should process log files, please contact Adobe Consulting.

Specifying the Processing Servers in Profile.cfg

In the `profile.cfg` file, specify the processing servers that process the data for the profile.

To access the `profile.cfg` file

You access the profile configuration file using the **Profile Manager** in Insight.

1. While working in your dataset profile, open the **Profile Manager** by right-clicking within a workspace and clicking **Admin > Profile > Profile Manager**, or by opening the Profile Management workspace on the **Admin** tab.
2. In the **Profile Manager**, right-click the check mark next to `profile.cfg` and click **Make Local**. A check mark for this file appears in the **User** column.
3. Right-click the newly created check mark and click **Open > in Insight**. The profile configuration window appears.

To add the Processing Servers

1. In the `profile.cfg` file, click **Profile**, then click **Processing Servers** to display its contents.
2. Right-click **Processing Servers** and click **Add new > Processing Server**.
3. In the Common Name parameter, type the common name for the first processing server in the cluster. For example: `server1.mycompany.com`
4. Repeat Steps 2 and 3 until you have added the common names of all of the processing servers in the cluster.



Note: If the master Insight Server processes data, you must add it as well.

5. Right-click (**modified**) at the top of the window and click **Save**.
6. Right-click the check mark in the **User** column next to `profile.cfg`. Click **Save to > <dataset profile name>**.

Modifying the Dataset Configuration Files for the Profile

To modify the dataset configuration files

If you need to make changes to the dataset configuration files (`Log Processing.cfg`, `Transformation.cfg`, dataset include files, `Log Processing Mode.cfg`, and so forth), do so only on the master Insight Server.

1. Access the files you want to modify:

For instructions to access the files, see the *Dataset Configuration Guide*.

2. Make your changes. See the *Dataset Configuration Guide* for details regarding the parameters within the configuration file(s).
3. Save the file.
 - a. Right-click (**modified**) at the top of the window and click **Save**.
 - b. Right-click the check mark in the **User** column next to the file name.
 - c. Click **Save to** and select the desired profile.



Note: Insight users who access a dataset profile running on a cluster identify only the master Insight Server in the Insight configuration file (`insight.cfg`). From the perspective of the Insight user, the profile is accessible on only one Insight Server (the master Insight Server); however, query requests from analysts can be directed to any of the Insight Servers in the cluster.

An Insight Server cluster permits the centralized storage of `.vs1` log files (from **Sensor**) on a single Insight Server machine called a File Server Unit (FSU). For information about installing an FSU, see [Installation Procedures for an Insight Server FSU](#) on page 813. For information about configuring an FSU, see the *Dataset Configuration Guide*.

Adding Insight Servers to an Existing Cluster

Navigation title: Adding Insight Servers to an Existing Cluster

In some cases, you may find it necessary to add an Insight Server machine to an existing Insight Server cluster.

When you add an Insight Server DPU or Insight Server FSU to a cluster, you must update the configuration files on the master Insight Server to include the address information for the new machine(s) and set up the new DPU or FSU.



Note: The procedures described in this section require Insight. If you have not installed Insight, follow the instructions in the *Insight User Guide* before proceeding.

Adding an Insight Server DPU to an Existing Cluster

Navigation title: Adding an Insight Server DPU to an Existing Cluster

Typically, you want to add an Insight Server DPU to an existing cluster when the amount of data you want to process and make accessible to your users of Insight and Report exceeds the capacity of your cluster's current configuration.

Updating the Configuration Files on the Master Server

In Insight, open up the **Server Files Manager** for your master Insight Server (usually an Insight Server FSU) and do the following for each DPU that you want to add to the cluster:

1. Edit the address file on the master Insight Server to include the name and address of the new DPU as described in [Adding the Processing Insight Servers to the Address File](#) on page 815. Add the name and address of the new DPU to the group in which your cluster's current Insight Servers are listed.
2. Edit the access control file on the master Insight Server to include the IP address of the new DPU as described in [Updating the Access Control File for a Cluster](#) on page 818.

Installing the New Insight Server DPU

This task applies only to Windows 32 bit.

1. Copy the following directories from one of the current DPUs in your cluster to the new DPU:
 - Insight Server installation directory/bin/
 - Insight Server installation directory/Certificates/
 - Insight Server installation directory/Components/
2. Download and install the digital certificate for the new DPU as described in [Downloading and Installing the Digital Certificates](#) on page 793.
3. Set the Windows memory utilization parameters on the new DPU.
4. Register Insight Server as a Windows service on the new DPU machine as described in [Registering Insight Server as a Windows Service](#) on page 811.
5. Check the Trace logs to make sure that the DPU is synchronizing to the master Insight Server.
6. Repeat steps 1 through 6 for each additional DPU that you are adding to the cluster.

Adding the New Insight Server DPU to the Dataset Profile's Processing Servers

If the new DPU processes the same dataset as the other DPUs in the cluster, add the common name of the new DPU to the `profile.cfg` file on master Insight Server as described in [Specifying the Processing Servers in Profile.cfg](#) on page 822.

Adding an Insight Server FSU to an Existing Cluster

Navigation title:Adding an Insight Server FSU to an Existing Cluster

You may want to add an Insight Server FSU to an existing cluster if you want to store source data on an additional file server or if you want to set up a back up for your master Insight Server.

To add an Insight Server FSU to an existing cluster, you must perform the following procedures:

1. [Updating the Configuration Files on the Master Server](#) on page 824
2. [Installing the New Insight Server FSU](#) on page 824
3. [Configuring the New Insight Server FSU](#) on page 825

Updating the Configuration Files on the Master Server

In Insight, open up the **Server Files Manager** for your master Insight Server (usually an Insight Server FSU) and do the following for each FSU that you want to add to the cluster:

1. Edit the address file on the master Insight Server to include the name and address of the new FSU as described in [Adding the Processing Insight Servers to the Address File](#) on page 815. Add the name and address of the new FSU to the group in which your cluster's current Insight Servers are listed.
2. Edit the access control file on the master Insight Server to include the IP address of the new FSU as described in [Updating the Access Control File for a Cluster](#) on page 818.

Installing the New Insight Server FSU

1. On your current FSU, make a zip file of the Insight Server installation directory and copy the file to the new FSU.
2. Unzip the file to the location where you wish to place the Insight Server software.
3. Download and install the digital certificate for the new FSU as described in [Downloading and Installing the Digital Certificates](#) on page 793.
4. Set the Windows memory utilization parameters on the new FSU.
5. Change the name of the `.address` file to reflect the name of the FSU as described in [Defining the Server's Network Location](#) on page 804.

-
6. If the drive structure on the new FSU is different than that on the primary FSU, you need to edit the `Disk Files.cfg` file.
 - a. Open the `Disk Files.cfg` file on the new FSU.
 - b. Update the settings to match the drives of the primary FSU as described in [Monitoring Dataset Data Space](#) on page 835.
 - c. Save the file locally and to the server.
 7. Register Insight Server as a Windows service on the new FSU machine as described in [Registering Insight Server as a Windows Service](#) on page 811.
 8. Repeat steps 1 through 6 for each additional FSU that you are adding to the cluster.

Configuring the New Insight Server FSU

The following procedures provide instructions for specific configuration tasks. Follow the instructions that are appropriate to your implementation of the new FSU.

To configure the FSU for source data storage

If the new FSU stores additional source data for the dataset running on the cluster, you must complete the file server configuration process as described in Configuring an Insight Server File Server Unit in the Log Processing Configuration File chapter of the *Dataset Configuration Guide*.

To make to new FSU the backup for the master Insight Server FSU

If you wish to make the new FSU the backup for the master Insight Server (which serves as an FSU for the cluster), you must modify the synchronization file on the new (backup) FSU so that it synchronizes with the master FSU.

1. On the backup Insight Server FSU, use the **Server Files Manager** to copy the `Synchronize.cfg` file in the `Components for Processing Servers` folder to the `Components` folder.
2. Open the `Synchronize.cfg` file (in the `Components` folder) in Insight.
3. Find the `SynchronizeDir` that specifies the location of the `Components` directory. There may be several synchronize directories listed under Directories, so you may need to view the contents for many of them (by clicking the server number) to find the desired server.
4. Edit the `SynchronizeDir` entry and add a second `SynchronizeDir` entry as shown in the example below.

-Synchronize.cfg (modified)	
-component	SynchronizeComponent
-Directories	
+0	SynchronizeDir
-1	SynchronizeDir
Local Path	Lookups\
Remote URI	/Lookups/
-2	SynchronizeDir
Local Path	Components\
Remote URI	/Components/
-3	SynchronizeDir
Local Path	Components for Processing Servers\
Remote URI	/Components for Processing Servers/
+4	SynchronizeDir
+5	SynchronizeDir
+6	SynchronizeDir
+Server	serverInfo
Allow Local Deletes	true
Bidirectional	false
Cluster Primary Server Address	
Download Only Newer Files	false
Enable Synchronization Log	true
Recursive	true
Synchronization Log Threshold	Low

- Save the modified file with a new name such as FSU_Synchronize.cfg so that you do not confuse it with the Synchronize.cfg files on the DPUs in the cluster.
- Use the **Server Files Manager** to save the local copy of the renamed file to the server. The backup FSU downloads the files in the identified directories from the master Insight Server FSU and dynamically retrieves updated copies of these files from the master Insight Server FSU when they change.

Administering Data Workbench Server

Navigation title:

Information about administrative tasks such as configuring access control, monitoring memory usage, and more.



Note: Many of the administrative procedures require the Insight client application. If you have not installed the client, follow the instructions in the [client install guide](#) before proceeding.

For information about the administrative interfaces available in Insight that enable you to monitor and manage your Adobe implementation, see [Administrative Interfaces](#).

Re-validating the Digital Certificate

Navigation title:Re-validating the Digital Certificate

After installation, the digital certificate issued by Adobe acts as a key that enables you to run your Adobe product.

Recommended Frequency: As needed

To function properly, a digital certificate must be current.

To remain current, your digital certificate must be re-validated on a regular basis (generally, every 30 days, but this can vary depending on your agreement with Adobe). If your machine has Internet access, the revalidation

process is completely transparent. Your Adobe product connects automatically to the Adobe License Server and re-validates the certificate when necessary. If your machine does not have Internet access, you must download a new, validated certificate from the Adobe License Server and install it on your machine using the steps provided in [Downloading and Installing the Digital Certificates](#) on page 793.

Confirming that the Service Is Running

Navigation title:Confirming that the Service Is Running

To ensure that your implementation is functioning properly, you should confirm that the Insight Server service is running.

Recommended Frequency: Every 5-10 minutes

To manually confirm that the Insight Server service is running

- Click **Start > Control Panel > Administrative Tools > Services**.

Insight Server is listed with a status of “Started.” If the service is not running when it should be, restart the service and check the **Event Viewer (Start > Control Panel > Administrative Tools > Event Viewer)** for errors. If you need help determining the cause of the problem, please contact Adobe Support.

Configuring Access Control

Navigation title:Configuring Access Control

The Access Control configuration file, `Access_Control.cfg`, defines the access control groups by which Insight Server grants permissions to files based on the attributes (OU, CN, and so forth) of the incoming connection’s certificate.

Recommended Frequency: As needed

Insight Server provides configurable access control groups to manage the security of connections to Insight Server. Access control groups identify users that are permitted to perform administrative functions through Insight.

If you need to provide access to new users or new machines, such as when adding a machine to an Insight Server cluster, you simply edit the Access Control configuration file.

Understanding Access Control Groups

Navigation title:Understanding Access Control Groups

Five pre-built access control groups are available, but you can create and manage additional groups as required.

You may define the members of each access control group, as well as the directories to which each group has Read-Only or Read-Write access.

The pre-defined groups are defined as follows:

Access Control Groups

Group	Description
Administrators	Allows access to all directories and files. The default IP address is 127.0.0.1 (local host).
Sensors	Allows access to only those files used by Sensor to transmit data.
Users	Allows read-only access to the elements required for an Insight user to perform basic analysis tasks.
Power Users	Allows read-only access to the elements required for an Insight user to perform basic analysis tasks, plus read and write access to the Profiles folder for modifying profiles.
Cluster Servers	Allows access to Insight Servers that are designated as cluster servers.
Report Servers	Allows access to Report machines that connect to the Insight Server.

Members of an access control group are defined using their IP addresses or SSL certificate information.

If an SSL certificate is not available, an IP address may be used to define a group member. The typical installation of Insight includes an SSL certificate, while the use of certificates for **Sensor(s)** is optional. For Insight Server, Cluster Servers are defined using IP addresses instead of SSL certificates.

The following codes can be used to define group members:

Access Type Code Definitions

Access Types Code	Definition
O	Organization
CN	Common Name
L	Locality
ST	State or Province
C	Country
OU	Organizational Unit
Email	Email address

Understanding Access Levels

Navigation title: Understanding Access Levels

Access levels describe which URIs on the machine a group of users is permitted to read or modify.

Follow these guidelines to define access levels as desired for your organization's users:

- Specific URIs with no trailing slash character restrict access to that URI only. For example, /Components/Communications.cfg provides access to the Communications.cfg file only.
- A trailing slash (/), specifying a directory, provides group members access to any URI that begins with that string. For example, /Profiles/ provides access to the entire Profiles directory.
- A trailing dollar sign symbol (\$) restricts access to the exact URI only, even if it is a directory. For example, /Profiles/\$ provides access to read the main Profiles directory, but not to read any files within that directory.

For access to specific files, you do not need to use a trailing \$.

For example, /Components/Communications.cfg and /Components/Communications.cfg\$ provide the same access.

- A percent symbol (%) can be used with CN (Common Name) to permit access. For example, /Users/%CN%/ allows access to the User directory matching the SSL certificate common name of the Insight user. Note that this syntax can be used only once in a URI.

The URIs in the pre-defined access control groups have been configured as follows:

Access Control Group URIs

Group Name	Read-Only Access	Read-Write Access	Description
Administrators		/	Read and write access to all Insight Server directories.
Sensors		/SensorInit.vsp /Submit.vsp	Read and write access to the two files that the Sensors use to communicate with the Insight Server.
Users	/Profiles/ /Status/ /Software/ /Addresses/ /Users/\$	/Users/%CN%/ /	Read and write access to the User directory matching the SSL certificate common name of the Insight user.

Group Name	Read-Only Access	Read-Write Access	Description
Power Users	/Profiles/\$ /Status/ /Software/ /Addresses/ /Users/\$	/Profiles/ /Users/%CN%/ /Cluster/	Power Users are allowed the same access as Users, with the added ability to write to the Profiles directory. These users may edit profiles and enable changes to be updated automatically for other Insight users, such as when distributing newly defined workspaces.
Cluster Servers	/Components for Processing Servers/ /Addresses/ /Profiles/ /Lookups/ /Access Control/ /Bin/ /Logs/		Read and write access to the Cluster directory.
Report Servers	/Profiles/\$ /Status/ /Software/ /Addresses/ /Users/\$	/Profiles/ /Users/%CN%/ /ReportStatus.vsp	Report machines are allowed the same access as Power Users, with the added ability to write to the ReportStatus.vsp file.

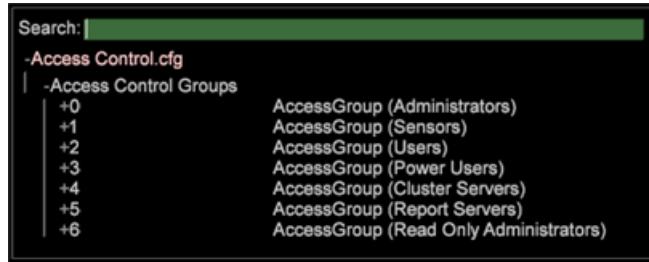
To configure Access Control

When defining access control groups, you need to include all System Administrators, users, cluster servers, and Report Server users that require access to this Insight Server computer. You can grant access using IP address or SSL certificate information, such as the common name or organization.



Note: When the Access Control.cfg file is changed on Insight Server, all existing connections are terminated and forced to reconnect. Connections are checked against the permissions in the updated Access Control.cfg file. In the Servers Manager interface, the Insight Server icon turns red temporarily and then green again because connection is terminated and forced to reconnect along with all others.

1. On the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of the Insight Server you want to configure and click **Files**.
3. In the **Server Files Manager**, click **Access Control** to view its contents. The Access Control.cfg file is located within this directory.
4. Right-click the check mark in the *server name* column for Access Control.cfg and click **Make Local**. A check mark appears in the **Temp** column for Access Control.cfg.
5. Right-click the newly created check mark in the **Temp** column and click **Open > in Workstation**.
6. In the Access Control.cfg window, click **Access Control Groups** to view its contents.



7. To add a new access control group:
 - a. Right-click **Access Control Groups** and click **Add new > Group**.
 - b. Right-click **Members** and click **Add new > Member**.

The members for the default groups are not pre-defined. By default, Administrator access is granted to 127.0.0.1 (local host), and **Sensor** access is granted to IP:*. All other access control group members must be defined.

 - c. Complete the parameters.
8. To add new members to an existing access control group:
 - a. Right-click **Members** under the appropriate access control group and click **Add new > Member**.
9. Save the file by right-clicking (**modified**) at the top of the window and then clicking **Save**.
10. To save the locally made changes to the Insight Server machine, in the **Server Files Manager**, right-click the check mark for **Access Control.cfg** in the **Temp** column, then click **Save to <server name>**.

User Administration of Group Member Access

Administrators can give workstation users the partial ability to manage access control for custom groups.

Self-administration of group member access gives rights to non-administrators to add and delete members in a custom group. The administrator creates a **User List** file and sets up group access in the **Access Control.cfg** file for the new group members.

Accessing the Servers Manager

Setting up the **User List** file and synching it with the **Communications.cfg** file is done in the **Servers Manager** workspace.

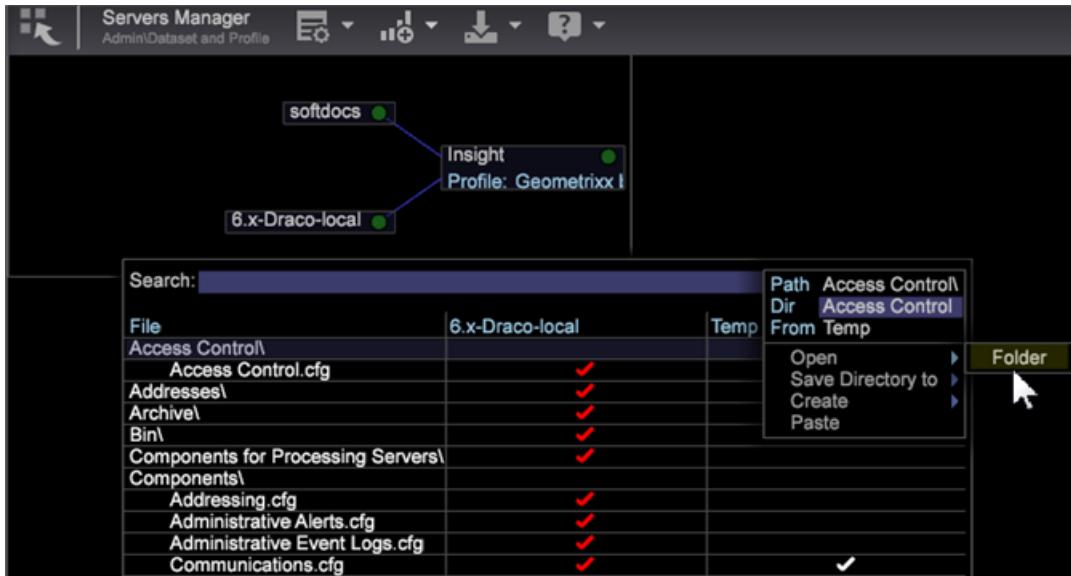
1. On the worktop, click the **Admin** tab > **Dataset and Profile** tab.
 2. Open the **Servers Manager** workspace.
 3. Right-click >*your server name*> in the diagram and select **Files**.
- The server files will open in a table with columns **File**, *<server name>*, and **Temp**.
4. **Make Local** by right-clicking in the **server** column of a server file (for this feature **Access Control** and **Components/Communications.cfg**).

A white checkmark will appear in the **Temp** column. You can edit in the Temp folder. Then right-click the checkmark and **Save To** the server. (It turns red when synched with server).

Create a User List.cfg file

The administrator needs to create a **User List.cfg** file in the **Access Control** folder.

1. Right-click **Access Control** row in the **Temp** column and select **Open > Folder**.



The Access Control folder in the **Temp** folder will open listing a single **Access Control.cfg** file.

2. Add another text file to this folder and name it **User List.cfg** (next to the **Access Control.cfg**).
3. Add the following parameters to the **User List.cfg** file.

The User List file should contain a vector of **AccessGroup** objects, and each **AccessGroup** object should have a name and a vector of strings called **Members**.

```
Access Control Groups = vector: 1 items
  0 = AccessGroup:
    Name = string: Group 1
    Members = vector: 1 items
      0 = string: CN:Joe User
```

You can then edit and add users this in the Workstation view of the **User List.cfg** file.



Here's the most basic parameters to add to the **User List.cfg** file. The Members can then be added in the Workstation view.

```
Access Control Groups = vector: 1 items
  0 = AccessGroup:
    Name = string:
    Members = vector: 0 items
```

Important: As with any **.cfg** file that you manually edit, make sure to use spaces instead of tabs and to pay close attention to the whitespace and syntax. A mistake in this file will cause *Adobe Insight Server* to ignore the User List file.

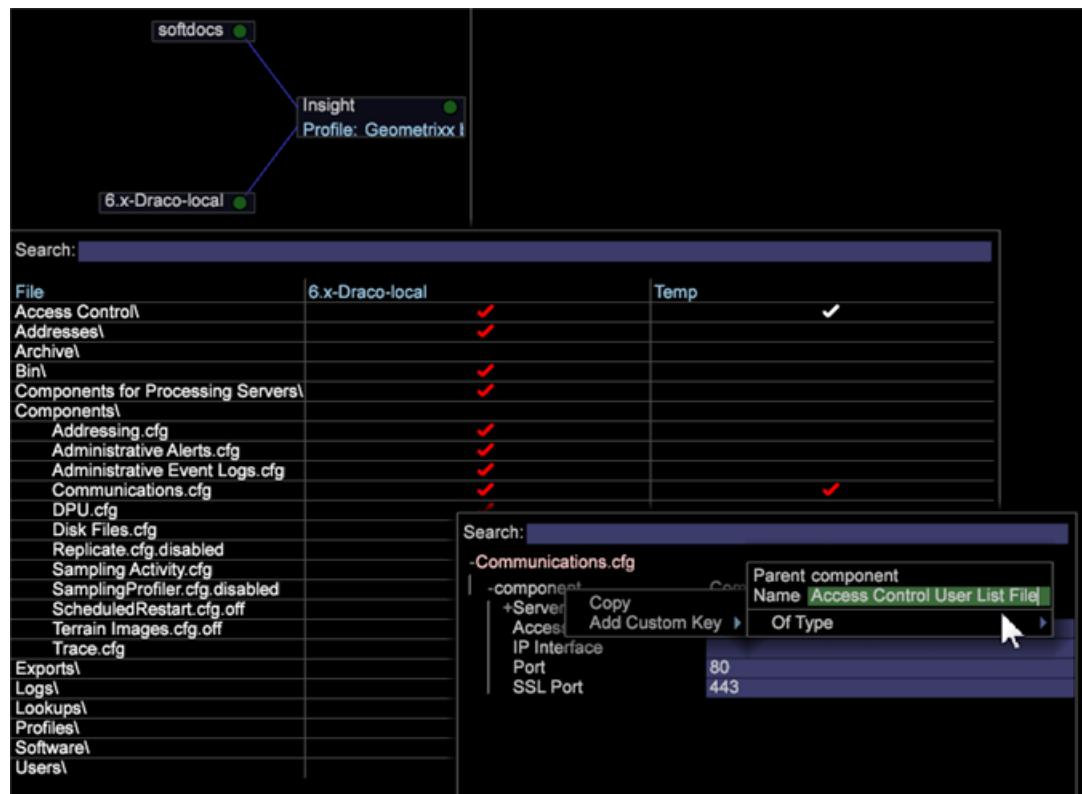
The **Name** field in each **Access Group** will be referenced within the **Access Control.cfg** file.

Note: Only valid members with directory service prefixes, such as **CN:** or **OU:** are accepted, and these cannot contain wildcard character (*).

Set up the **Communications.cfg** file

An administrator first enables this feature by opening the **Components > Communications.cfg** file and adding a new key with the name **Access Control User List File**. The string value of this key is the path where this new file will be located.

1. From the server files, click **Components** and right-click the checkmark in the **server** column. Click **Make Local**.
A white checkmark will appear in the **Temp** column.
2. Right-click the checkmark in the **Temp** column and select **Open > in Workstation**.
3. In the **Communication.cfg** file, right-click **component** and select **Add Custom Key**.



4. Type the **Name** as *Access Control User List File* and set **Of Type** as *String*.

Note: You cannot create the new list file as a Path. To remedy this, you need to save the file, open it in an editor (Notepad), and change "String" to "Path":

Before:

```
component = CommServer:  
Access Control File = Path: Access Control\\Access Control.cfg  
Access Control User List File = string: Access Control\\User List.cfg
```

After:

```
component = CommServer:  
Access Control File = Path: Access Control\\Access Control.cfg  
Access Control User List File = Path: Access Control\\User List.cfg
```

5. Save the **Communications.cfg** file and (if necessary) save it to the server. This will restart components in the server to make sure you haven't made any mistakes that could prevent the **Communications.cfg** file from being parsed.
6. If your system includes processing servers, modify the configuration file in the **Components for Processing Servers.cfg** file.
7. Right-click **Communications.cfg** and save to server.

The Data Workbench administrator can now confirm that the intended user(s) have access to the user list file and allow the users to manage the group. The user(s) will be able to open the User List file, edit it, and add and remove CN or OU members as needed.

Synch up the **Access Control.cfg** file

The administrator can then edit the **Access Control.cfg** and insert references to the group(s) defined by the *User List* file.

The references to the group(s) should be inserted just like any other member, but with the following syntax:

```
$(Group Name)
```

Where "Group Name" matches what's defined in the user list file, including white spaces.

Search: []

- Access Control.cfg (modified)
 - | -Access Control Groups
 - 0 Name AccessGroup Administrators
 - Members 0 OU:Insight Developers
 - +Read-Write Access
 - Read-Only Access
 - +1
 - +2
 - 3 Name AccessGroup Power Users
 - Members 0 \$(Power User Group)
 - +Read-Only Access
 - +Read-Write Access
- +4 AccessGroup (Cluster Servers)

At this point the Data Workbench administrator can confirm that select group users have access to the user list file. The select users can then open the **User List.cfg** file, edit it, and add and remove CN or OU members as needed.

Monitoring Disk Space

Navigation title: Monitoring Disk Space

You should regularly monitor available disk space so that all Insight Server machines continue to perform at the highest level possible.

Each Insight Server machine stores the following types of data:

- Dataset data
- Event data
- Integration data
- Operating system data
- Output data
- System data

For more information about data storage requirements than is provided in this section, see the *Minimum System Requirements* document.



Note: Adobe consultants can evaluate your use scenario to help you project the amount of data storage that your use of Adobe software applications generate and require. To request such an evaluation, please contact Adobe Support Services.

Monitoring Dataset Data Space

Navigation title: Monitoring Dataset Data Space

Information about dataset monitoring and adding new locations for dataset data storage.

Recommended Frequency: Every 5-10 minutes

By default, Insight Server writes its dataset to the `temp.db` file on the same drive as the Insight Server program files on the Data Processing Unit. The amount of dataset data per Insight Server machine is limited to the following, whichever occurs first:

- Five hundred (500) million records of data input to that dataset
- Five hundred (500) GB of dataset data stored
- One (1) MB of dataset data stored per any one root-level dimension (for example, 5,000 records per Visitor at an average 200 bytes per record)

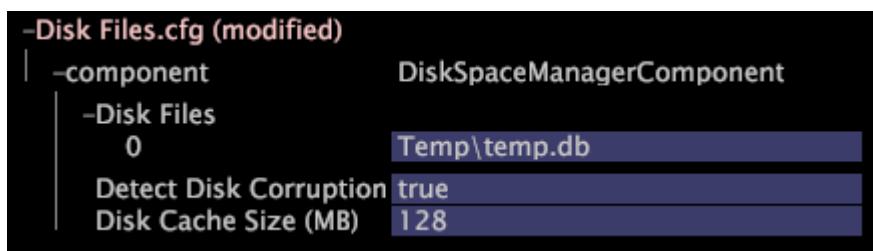
If you want Insight Server to maintain the dataset on a different drive, or if the amount of data you expect to collect requires the use of multiple drives, you must update the Disk Files configuration file (`Disk_Files.cfg`) to specify where you want Insight Server to write the `temp.db` file(s). The `Disk_Files.cfg` file lists the disk files (a vector of strings) and specifies the location of the dataset data used by Insight Server during reprocessing and operation. There is usually one file per physical drive.



Note: The contents of the `Disk_Files.cfg` file may have been modified while installing Insight Server. For more information, see [Configuring the Location of the Dataset \(temp.db\)](#) on page 808.

To add new locations for dataset data storage

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of the Insight Server you want to configure and click **Server Files**.
3. In the **Server Files Manager**, click **Components** to view its contents. The `Disk_Files.cfg` file is located within this directory.
4. Right-click the check mark in the *server name* column for `Disk_Files.cfg` and click **Make Local**. A check mark appears in the **Temp** column for `Disk_Files.cfg`.
5. Right-click the newly created check mark in the **Temp** column and click **Open > in Insight**.
6. In the `Disk_Files.cfg` window, click **component** to view its contents.



Note: The Detect Disk Corruption parameter is set to true by default. The Disk Cache Size (MB) parameter controls the amount of memory that Insight Server uses to increase disk access speed and is set to 128 by default. Please contact Adobe before changing either of these parameters.

7. To change the disk files on the Insight Server machine, right-click **Disk Files** and click **Add new > Disk File**.
To delete a disk file, right-click the disk file number and click **Remove**.
8. For the new disk file, enter the directory and name of the file to be used by Insight Server during reprocessing and operation.

-Disk Files.cfg	DiskSpaceManagerComponent
-component	
-Disk Files	
0	E:\Temp\temp.db
1	F:\Temp\temp.db
2	G:\Temp\temp.db
3	H:\Temp\temp.db
Detect Disk Corruption	true
Disk Cache Size (MB)	128



Note: The Detect Disk Corruption parameter is set to true by default. The Disk Cache Size (MB) parameter controls the amount of memory that Insight Server uses to increase disk access speed and is set to 128 by default. Please contact Adobe before changing either of these parameters.

- Save your changes to the server by doing the following:

- Right-click (**modified**) at the top of the window and click **Save**.
- In the **Server Files Manager**, right-click the check mark for the file in the **Temp** column and select **Save to > <server name>**.

Monitoring Event Data Space

Navigation title: Monitoring Event Data Space

Information about monitoring event data space and changing the log directory for **Sensor** data.

Recommended Frequency: Every 5-10 minutes

Insight Server stores one log file per **Sensor** per day on either the Data Processing Unit or the File Server Unit, depending on your configuration. The size of the log files and the amount of data storage space required for them depends on many variables, including, for example, the number of websites being logged and the number of requests your web servers receive per second.

A typical installation of Insight Server (or an Insight Server cluster) is capable of storing multiple terabytes of data, assuming that the implementation uses the hardware recommended by Adobe for the Insight Server machine(s).

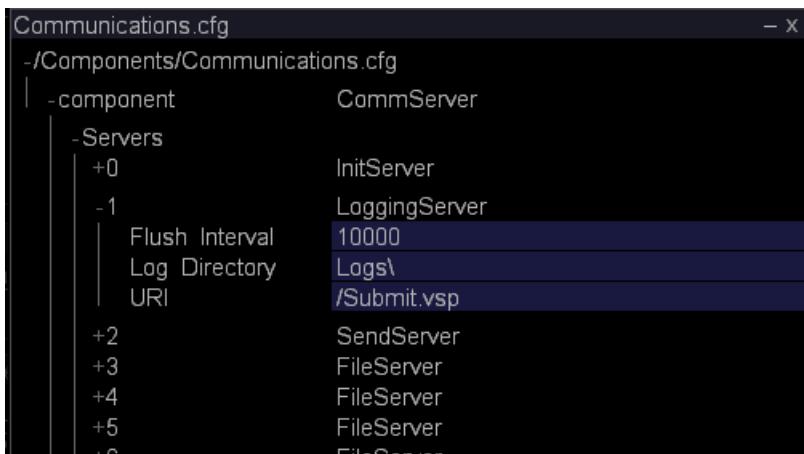
Typically, all log data remains present on the Insight Server machine. If it becomes necessary to make more data storage space available on the machine, you can move all but the most current day's log files to another machine or data storage medium (zip drive, tape, and so forth). Moving the data does not require you to stop Insight Server, and it does not affect the functionality available in any Insights that may be connected to Insight Server and working with continuous data. Provided that you do not process or reprocess an analysis dataset, you retain access to all previous data and new data continues to be available in Insight. If you process or reprocess an analysis dataset, you can not access the data until processing is complete.

By default, event data produced by **Sensor** and transmitted to Insight Server is stored in the **Logs** folder within the Insight Server installation directory. The Communications configuration file, **Communications.cfg**, specifies the location of event data log files that are read by Insight Server.

To change the log directory for Sensor data

- In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
- Right-click the icon of the Insight Server you want to configure and click **Server Files**.
- In the **Server Files Manager**, click **Components** to view its contents. The **Communications.cfg** file is located within this directory.

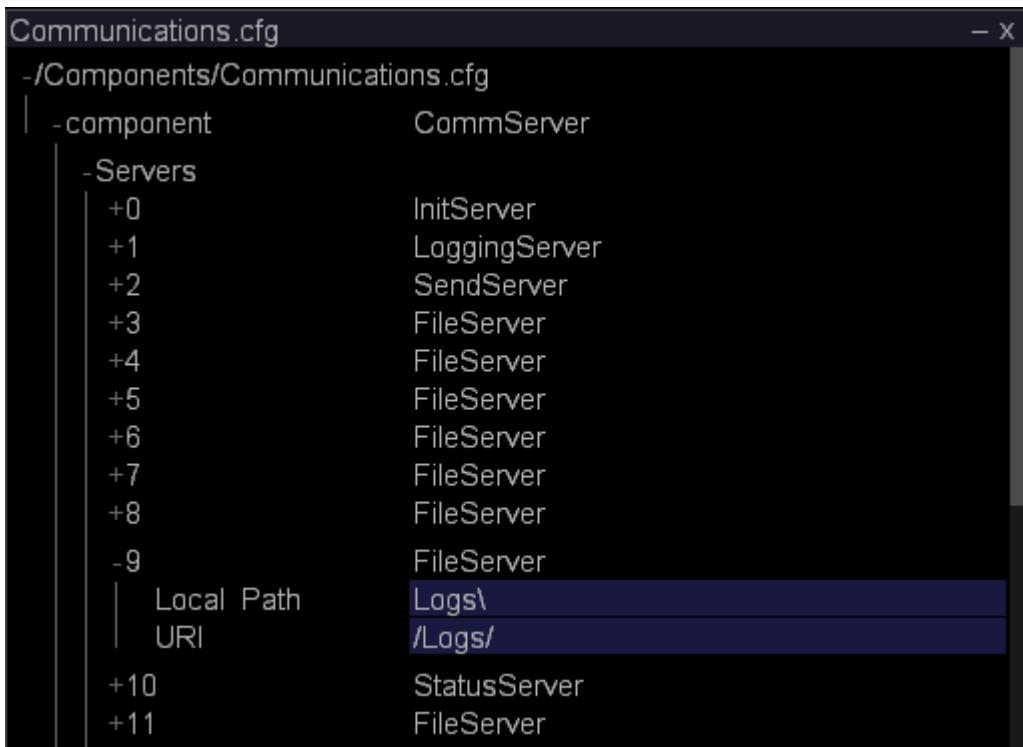
4. Right-click the check mark in the *server name* column for Communications.cfg and click **Make Local**. A check mark appears in the **Temp** column for Communications.cfg.
5. Right-click the newly created check mark in the **Temp** column and click **Open > in Insight**.
6. In the Communications.cfg window, click **component** to view its contents.
7. In the Communications.cfg window, click **Servers** to view its contents. Several types of servers may appear: File Servers, Logging Servers, Init Servers, Status Servers, Send Servers, or Replicate Servers.
8. Find the LoggingServer, which is where **Sensor** writes its log files to be processed by Insight Server, and click its number to view the menu.



The default log directory is the Logs folder within the Insight Server installation directory.

9. Edit the Log Directory parameter to reflect the desired location of the log files.

Note: Do not modify any other parameters for the LoggingServer.



Several FileServers may be listed under the Servers node, so you may need to view the contents of many of them (by clicking their numbers in the **Servers** list) to find the server with a Local Path of Logs\ to be modified.

10. Edit the Local Path to reflect the desired location of the .vsl files.



Note: Do not modify any other parameters for the FileServer.

Although the location of the log files has been changed in the Communications.cfg file, you can map these files to the Logs directory of the **Server Files Manager** by specifying /Logs/ as the URI for the FileServer.

11. Save your changes to the server by doing the following:

1. Right-click (**modified**) at the top of the window and click **Save**.
2. In the **Server Files Manager**, right-click the check mark for the file in the **Temp** column and select **Save to > <server name>**.

Backing Up Event Data

Navigation title:Backing Up Event Data

Event data must be backed up daily using your company's normal backup systems and disaster recovery procedures.

Recommended Frequency: Daily

Insight Server begins new log files at 12:00 AM GMT daily. Adobe recommends that you backup log files every day shortly after 12:00 AM GMT so that you capture all of the data from the previous day. For example, backing up all log files at 12:05 AM GMT on December 15th captures all of the data from December 14th. Insight Server does not need to be stopped during log file backups, and all functionality remains available.

Backing Up Integration, Operating System, Output, and System Data

Recommended Frequency: Daily

Integration, operating system, output, and system data must be backed up regularly and diligently using your company's normal backup and disaster recovery systems.

Monitoring Memory Usage

Navigation title:Monitoring Memory Usage

Information on assessing and monitoring the Address Space Load.

Monitoring the Address Space Load

Recommended Frequency: Daily

The Address Space load is a measure of the fraction of the maximum Address Space that a properly configured Insight Server uses. Even if the configuration parameters are changed to reduce memory usage, it usually does not decrease until the Insight Server service is restarted.

A safety margin is built into the Address Space load maximum to account for unexpected increases in Address Space utilization. You should never deliberately cut into this safety margin. It exists for emergency situations and not for the support of functionality added to your Adobe application.



Note: To make more Address Space available and avoid memory exhaustion errors, ensure that your operating system has the /3GB Switch enabled and that Low Fragmentation Heap is operating.

Errors logged to the Insight Server event data log can provide a clue that problems are developing with your Address Space load:

-
- “Requested X byte block is too large” errors indicate that something may be having an excessive impact on Address Space load, performance, and network bandwidth. Such large blocks can contribute greatly to Address Space usage, both by using a lot of memory and by requiring large contiguous blocks of Address Space.

To address this problem, you can reduce the cardinality of your largest dimensions, which increases your Address Space load. If you cannot reduce the cardinality of the dimensions, you should attempt to keep the Address Space load small enough so that you can handle an unexpected increase.

- “Memory budget exceeded” errors indicate that you might need to increase the Query Memory Limit. Memory used by queries is proportional to dimension cardinality and, in some cases, the lengths of the element names. If you increase the Query Memory Limit, you increase your total Address Space load and shrink the large dimensions.



Note: By default, use of large pages are allowed and memory-mapped lookup is disabled. In DWB 6.7 and later, this can be changed in dataset configuration. Any changes require a Server restart.

To assess Address Space load

To accurately assess the Address Space load for your system, Adobe recommends reprocessing the dataset, performing some normal queries without subsequently restarting Insight Server, and then viewing the measured Address Space load by following these steps.

If an Insight Server has not been reprocessed and queried significantly since it was last restarted, you should not draw conclusions from the Address Space load.

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of the Insight Server you want to configure and click **Detailed Status**.
3. In the Detailed Status interface, click **Memory Status** to view its contents. In the Address Space Load parameter, you can see the Address Space load expressed as a percentage and a parenthetical description indicating status.

The following table presents ranges and status for Address Space load. A recommended action is listed for each range.

Address Space Status

Address Space Load (%)	Status	Recommended Action
0-15	lean and mean	None.
15-33	light	None.
33-66	moderate	None.
66-100	heavy	To avoid memory exhaustion failures, do not add significant extra functionality or attempt to process more data.
100-125	reliability compromised	Adjust your dataset configuration to reduce Address Space load.
125 or greater	failure imminent	Adjust your dataset configuration to reduce Address Space load. You may not see the failure imminent status before failure occurs.

If you see an Address Space load above 100%, you should change the configuration as soon as possible to reduce Address Space usage.

Configuring Administrative Alerts

Navigation title:Configuring Administrative Alerts

Administrative alerts send email notifications to the specified email addresses when errors are detected by the Insight Server during the normal course of operation.

Recommended Frequency: Prior to production



Note: The use of administrative alerts requires that Insight Server has access to a forwarding SMTP server to send alerts by email.

The recipients of the email notifications should be key systems administration personnel and primary stakeholders.

The Administrative Alerts file, `Administrative Alerts.cfg`, is used to configure the administrative alerts for Insight Server.



Note: If you are running a cluster, you must create or modify alerts on the master Insight Server in the cluster.

To create or modify an administrative alert

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of the Insight Server you want to configure and click **Server Files**.
3. In the **Server Files Manager**, click **Components** to view its contents. The `Administrative Alerts.cfg` file is located within this directory.
4. Right-click the check mark in the *server name* column for `Administrative Alerts.cfg` and click **Make Local**. A check mark appears in the **Temp** column for `Administrative Alerts.cfg`.
5. Right-click the newly created check mark in the **Temp** column and click **Open > in Insight**.
6. In the `Administrative Alerts.cfg` window, click **component** to view its contents.
7. Fill out the parameters as desired. For a list of the parameters available in this file, see [Administrative Alerts Configuration Settings](#) on page 871.

The screenshot shows the **Administrative Alerts.cfg (modified)** file structure. The `-component AdminEmail` section is expanded, showing the `-Error Categories` section. The `+0` and `-1` entries under `-Error Categories` are collapsed. The `-Recipients` section is expanded, showing two entries: `0 support@omniture.com` and `1 docs@omniture.com`. The `Category` and `Throttle Delay (secs)` parameters are set to `Critical` and `0` respectively. The `+2` entry is collapsed. The `Error Categorization File` parameter is set to `Lookups\Error Categories.txt`. The `From` parameter is set to `support@mycompany.com`. The `Minimum Disk Space (MB)` parameter is set to `1000`. The `Sensor Alert Timeout (min)` parameter is set to `15`. The `Server Address`, `Server Password`, and `Server User` parameters are empty.

8. Save your changes to the server by doing the following:

-
1. Right-click (**modified**) at the top of the window and click **Save**.
 2. In the **Server Files Manager**, right-click the check mark for the file in the **Temp** column and select **Save to > <server name>**.
 9. (Optional) If you are working in a cluster and you want the same administrative alerts applied to each Data Processing Unit, you must copy and paste the updated **Administrative Alerts.cfg** file into the **Components for Processing Servers** folder within the master Insight Server installation directory.

Monitoring Administrative Events

Navigation title: Monitoring Administrative Events

You should regularly monitor your event log files to track Insight Server system event messages, which are logged to the <YYYYMMDD>-event.txt files located by default in the **Events** folder within the Insight Server installation directory.

Recommended Frequency: Every 5-10 minutes

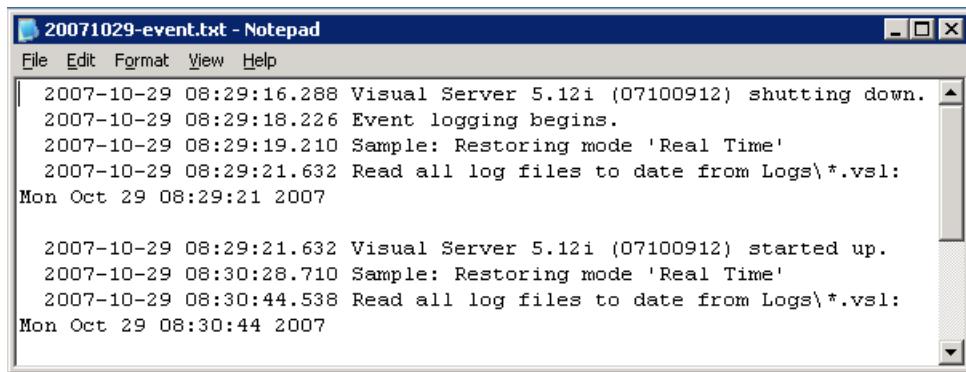
You can monitor these events using the **Server Files Manager** in Insight, your automated management tool, the *-event.txt files, or the Windows Event Viewer.



Note: The Administrative Event Logs are completely separate from your Windows Events Log, but contains some of the same events. The Administrative Event Logs contain information only about Insight Server events.

To view events.txt files through the Server Files Manager

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of an active Insight Server and click **Server Files**.
3. In the **Server Files Manager**, click **Events** to view its contents.
4. Right-click the check mark in the *server name* column next to the desired file and click **Make Local**. A check mark appears next to the file name in the **Temp** column.
5. Right-click the check mark in the **Temp** column and click **Open > in Notepad**. The event file appears in a new Microsoft Windows Notepad window.



The **Server.log** file in the **Trace** folder within the Insight Server installation directory contains more detailed logging information.

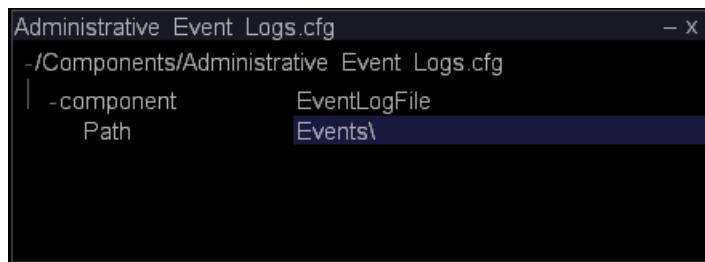
To view Events through the Windows Event Viewer

- Click **Start > Control Panel > Administrative Tools > Event Viewer**.

To change the Administrative Events Log directory

The Administrative Event Logs configuration file, `Administrative Events Log.cfg`, specifies the directory to which event logging is output.

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of the Insight Server you want to configure and click **Server Files**.
3. In the **Server Files Manager**, click **Components** to view its contents. The `Administrative Event Logs.cfg` file is located within this directory.
4. Right-click the check mark in the *server name* column for `Administrative Event Logs.cfg` and click **Make Local**. A check mark appears in the **Temp** column for `Administrative Event Logs.cfg`.
5. Right-click the newly created check mark in the **Temp** column and click **Open > in Insight**.
6. In the `Administrative Event Logs.cfg` window, click **component** to view its contents. The default path is the `Events` folder within the Insight Server installation directory.



7. In the **Path** parameter, type the name of the directory to which you want to output event logging data.
8. Save your changes to the server by doing the following:
 - a. Right-click (**modified**) at the top of the window and click **Save**.
 - b. In the **Server Files Manager**, right-click the check mark for the file in the **Temp** column and select **Save to > <server name>**.

Monitoring Audit Logs

Navigation title: Monitoring Audit Logs

Audit log files track all attempted connections to and disconnections from Insight Server, each of which is logged in the `<YYYYMMDD>-access.txt` files located by default in the `Audit` folder within the Insight Server installation directory.

Recommended Frequency: Daily or as needed for troubleshooting

Audit logs can be very helpful when troubleshooting issues connecting to Insight Server. You can monitor these logs using your automated management tool or by viewing the `access.txt` files directly.

To view access.txt files through the Server Files Manager

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of an active Insight Server and click **Server Files**.
3. In the **Server Files Manager**, click **Audit** to view its contents.
4. Right-click the check mark in the *server name* column next to the desired file and click **Make Local**. A check mark appears next to the file name in the **Temp** column.
5. Right-click the new check mark in the **Temp** column and click **Open > in Notepad**. The audit log appears in a new Microsoft Windows Notepad window.

```

20050511-access.txt - Notepad
File Edit Format View Help
10.1.0.30:1746 2005-05-11 12:24:32.724 Connection established.
10.1.0.30:1747 2005-05-11 12:24:32.724 Connection established.
10.1.0.30:1748 2005-05-11 12:24:32.740 Connection established.
10.1.0.30:1749 2005-05-11 12:24:32.740 Connection established.
10.1.0.30:1750 2005-05-11 12:24:32.740 Connection established.
10.1.0.30:1751 2005-05-11 12:24:32.756 Connection established.
10.1.0.30:1752 2005-05-11 12:24:32.756 Connection established.
10.1.0.30:1753 2005-05-11 12:24:32.756 Connection established.
10.1.0.30:1754 2005-05-11 12:24:32.756 Connection established.
10.1.0.30:1755 2005-05-11 12:24:32.756 Connection established.
10.1.0.30:1756 2005-05-11 12:24:32.771 Connection established.

```

Configuring Communications

Navigation title: Configuring Communications

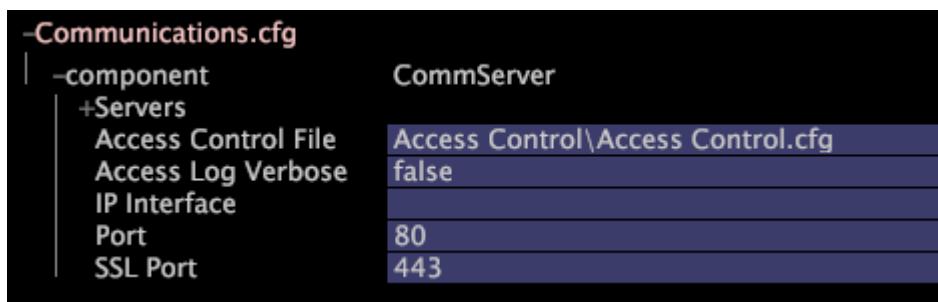
The Communications configuration file, Communications.cfg, contains Insight Server network settings and the path to the Access Control.cfg file.

These settings help you to connect to Insight Server.

Recommended Frequency: Only when necessary

To view and modify communications settings in Insight

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of the Insight Server you want to configure and click **Server Files**.
3. In the **Server Files Manager**, click **Components** to view its contents. The Communications.cfg file is located within this directory.
4. Right-click the check mark in the *server name* column for Communications.cfg and click **Make Local**. A check mark appears in the **Temp** column for Communications.cfg.
5. Right-click the newly created check mark in the **Temp** column and click **Open > in Insight**.
6. In the Communications.cfg window, click **component** to view its contents.
7. Change the settings, as necessary. For information about the parameters available in this file, see [Communications Configuration Settings](#) on page 872.



8. Save your changes to the server by doing the following:

1. Right-click (**modified**) at the top of the window and click **Save**.

-
2. In the **Server Files Manager**, right-click the check mark for the file in the **Temp** column and select **Save to > <server name>**.

Configuring DPU.cfg

Navigation title:Configuring DPU.cfg

The DPU configuration file, `DPU.cfg`, specifies various performance parameters for Insight Server.

How you set these parameters depends on your dataset size and many other factors. Please contact Adobe Consulting Services for help with performance tuning.

Recommended Frequency: Only when necessary

To change Insight Server DPU performance settings

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of the Insight Server you want to configure and click **Server Files**.
3. In the **Server Files Manager**, click **Components** to view its contents. The `DPU.cfg` file is located within this directory.
4. Right-click the check mark in the *server name* column for `DPU.cfg` and click **Make Local**. A check mark appears in the **Temp** column for `DPU.cfg`.
5. Right-click the newly created check mark in the **Temp** column and click **Open > in Insight**.
6. In the `DPU.cfg` window, click component to view its contents.
7. Change performance and path settings, as necessary. For a list of the parameters available in this file, see [DPU Performance Settings](#) on page 874.



Note: Please contact Adobe before changing any of the parameters in this file.

-DPU.cfg	
-component	ServerProfileManager
Execution Batch Count	2048
Execution Batches	128
Execution Time	0.100
Field Dump on Error	false
Profile Path	Profiles\
State Path	State\
Threads	5
User Path	Users\

8. Save your changes to the server by doing the following:
 1. Right-click (**modified**) at the top of the window and click **Save**.
 2. In the **Server Files Manager**, right-click the check mark for the file in the **Temp** column and select **Save to > <server name>**.

Restarting the Service

Navigation title:Restarting the Service

When restarting an Insight Server machine, you first must stop the Insight Server service.

Recommended Frequency: Only when necessary

To stop the Insight Server service

1. Click **Start > Control Panel > Administrative Tools > Services**.
2. Select the Insight Server service and click the **Stop** button.

When you restart the Insight Server machine, the Insight Server service starts automatically.

Transform

Navigation title: Transform

Steps to configure **Transform** on an Insight Server FSU.



Note: You must have an Insight Special Use License to configure the Insight FSU for **Transform** functionality. Please contact Adobe Customer Care for more information.

Configuring Transform

Navigation title: Configuring Transform

Transform functionality runs on an Insight Server FSU machine to enable the export of log source data for use by other applications.

Transform can read .vs1 files, log files, XML files, and ODBC data and export the data as .vs1 files, text files, or delimited text files that can be used by data warehouse loading routines, auditing agencies, or other targets. The data extraction and transformation can be performed on a continuous or other scheduled basis. Each Insight Server FSU that provides output of altered event data must run **Transform**.



Note: Typically, **Transform** is installed on an Insight Server FSU. However, your implementation may require installation on an Insight Server DPU. For more information, contact Adobe.

For information about the system requirements for installing, configuring, and operating **Transform**, see the *Minimum System Requirements* document.

Adobe distributes the **Transform** functionality as a profile within the .zip file for the Insight Server release package. The **Transform** profile is an internal profile that provides additional functionality to Insight Server. As with all other internal profiles provided by Adobe, the profile should not be changed. All customization must occur in your dataset or role-specific profiles or other profiles that you create.

The profile consists of the following files:

- Log Processing.cfg
- Insight Transform.cfg
- Insight Transform Mode.cfg
- a log processing dataset include file

All of these files are located in the Dataset folder for the profile.

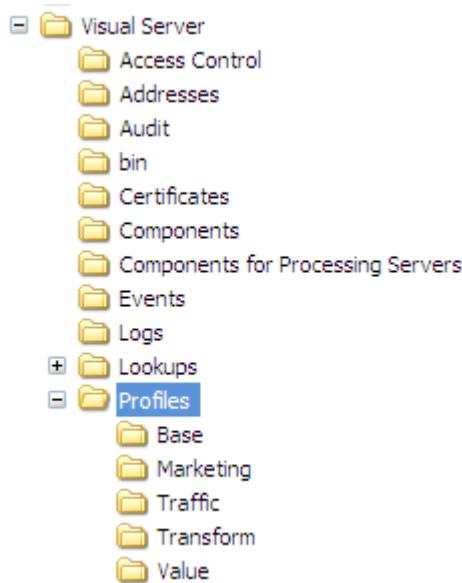
To install the Transform profile on Insight Server



Note: The following installation instructions assume that you have installed Insight and established a connection between Insight and the Insight Server on which you are installing **Transform**. If you have not done so, see the *Insight User Guide*.

1. Open the .zip file for the Insight Server release package, and open the Profiles folder within that .zip file.

-
2. Copy the **Transform** folder to the **Profiles** folder in your Insight Server installation directory. You want to end up with a . . . \Profiles\Transform folder on your Insight Server as shown in the following example.



Note: If you followed all of the steps for installing Insight Server (see *Server Administration and Installation* on page 787), you already may have a **Transform** folder in the **Profiles** directory.

3. Use the following steps to update the **profile.cfg** file for the profile with which you want to use **Transform**. The dataset reprocesses upon completion of these steps.
1. Open the **Profile Manager**.
 2. Right-click the check mark next to **profile.cfg** and click **Make Local**. A check mark for this file appears in the **User** column.
 3. Right-click the newly created check mark and click **Open > in Insight**. The **profile.cfg** window appears.
 4. In the **profile.cfg** window, right-click **Directories** and click **Add new > Directory**.
To add the new directory to the end of the list of directories, right-click the number or name of the last directory in the list and click **Add new > Directory**.
 5. Type the name of the new directory: **Transform**
 6. Right-click (**modified**) at the top of the window and click **Save**.
 7. In the **Profile Manager**, right-click the check mark for **profile.cfg** in the **User** column, then click **Save to > <profile name>**.



Note: Do not save the modified configuration file to any of the internal profiles provided by Adobe (including the profile), as your changes are overwritten when you install updates to these profiles.

Next Steps

Navigation title: Next Steps

After you have set up **Transform**, you must configure it to export log source data for use by other applications.

To configure **Transform**, you edit the `Insight Transform.cfg` and `Insight Transform Mode.cfg` files within the `Dataset` folder for the profile for which you want to export event data. For configuration instructions, see the **Transform** chapter of the *Dataset Configuration Guide*.

After you have configured **Transform** to export event data for your profile, you can view memory usage information in the Detailed Status interface. For more information, see the Administrative Interfaces chapter of the *Insight User Guide*.

Insight Server Replication Service

Navigation title:Insight Server Replication Service

Information about installing and configuring the Insight Server replication service.

Installing the Replication Service

Navigation title:Installing the Replication Service

The Insight Server **Replication Service** enables you to transmit the event data collected and stored on one Insight Server machine to another Insight Server machine.

Typically, the machine on which the data is collected and stored is configured to run as a **Repeater** machine. See [Repeater Functionality](#) on page 850. The **Replication Service**, which is configured by the `Replicate.cfg` file, enables an Insight Server machine to retrieve data from the **Repeater** machine and store it locally. The Insight Server machine is referred to as the target machine. Multiple Insight Server DPUs can connect to a single **Repeater** to request copies of the event data for inclusion in multiple datasets.

You can use the **Replication Service** in network infrastructures with multiple layers of firewalling to achieve single-port to single-port communications between components that are separated by firewalls.

To install the Replication Service

The `Replicate.cfg` file should be installed as part of your Insight Server installation. You can find the file within the `Components` folder of your Insight Server installation directory. If you do not have this file, contact Adobe.

Configuring the Replication Service

Navigation title:Configuring the Replication Service

You must configure the target Insight Server(s) to retrieve data from the **Repeater** on which the original event data is stored.

To configure the retrieval of data from a **Repeater** to a target Insight Server, you must edit the `Replicate.cfg` file provided in the `Components` folder on the target Insight Server(s) as described in the following procedure:

To configure the Replication Service on the target machine

1. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
2. Right-click the icon of the target Insight Server you want to configure and click **Server Files**.
3. In the **Server Files Manager**, click **Components** to view its contents. The `Replicate.cfg` file is located within this directory.
4. Right-click the check mark in the *server name* column for `Replicate.cfg` and click **Make Local**. A check mark appears in the **Temp** column for `Replicate.cfg`.
5. Right-click the newly created check mark in the **Temp** column and click **Open > in Insight**. The `Replicate.cfg` window opens.
6. In the `Replicate.cfg` window, click **Replicate.cfg**, then **component** to view its contents.
7. Edit the parameters using the following example and table as guides:

-Replicate.cfg (modified)	
-component	Replicate
-Directories	ReplicatedDirectory
-0	false
Flatten Paths	F:\Logs
Local Path	true
Recursive	/Logs/*.vsl
Remote URI	
-Server	serverInfo
Name	Visual Repeater
SSL Server Common Name	
Address	visualserver.mycompany.com
Port	80
SSL Client Certificate	Certificates\server_cert.pem
Use SSL	false
End Time	October 31, 2020 EST
Start Time	January 1, 2007 EST

For this parameter...	Specify...
Directories	The directories on the Repeater that are to be copied (replicated) to the target Insight Server. The Replication Service allows the replication of multiple directories from a single Repeater . To add a new directory, right-click Directories and click Add new > Directory .
Flatten Paths	True or false. The action defined by this parameter's setting depends on the setting of the Recursive parameter in this file: <ul style="list-style-type: none"> • If Recursive is false, Flatten Paths has no effect. Only the files at the top level of the directory specified by the Remote URI parameter are replicated. • If Recursive is true and Flatten Paths is false, the directory structure of the remote (Repeater) directory is duplicated exactly in the local path on the target Insight Server. • If both Recursive and Flatten Paths are true, no subdirectories are created in the local path. Instead, all files from the remote directory tree are placed in the top level of the local directory.  Note: If both Flatten Paths and Recursive are true and files in the various subdirectories on the remote machine share the same name(s), the Replication Service may stop or other undefined behavior may occur.
Local Path	The storage location for the files retrieved from Repeater . The path is relative to the Insight Server installation directory.
Recursive	True or false. If false, only the files at the top level of the directory specified by the Remote URI parameter are replicated. See Flatten Paths in this table.
Remote URI	The URI, including a file mask, to access the Repeater's file store. The Communications.cfg file on the Repeater should be configured so that the event data can be accessed using this URI. See Monitoring Event Data Space on page 836.
Server	Parameters for the Repeater from which the target Insight Server retrieves event data files.
Name	Optional. The name to identify the Repeater .
SSL Server Common Name	Required only if Use SSL is set to true. Common name of the Repeater on which the event data is stored. This name must match the common name listed in the communications certificate for the machine.
Address	Host name or numeric IP address of the Repeater on which the event data is stored. The server's common name is not a valid entry.

For this parameter...	Specify...
Port	Port used for data transmission. The default port is 80.
SSL Client Certificate	Required only if Use SSL is set to true. Name of the license certificate used to connect to the Repeater .
Use SSL	Determines whether SSL is used for the data transmission. The options are true or false, and the default value is false.  Note: Using SSL is not recommended because it can negatively affect performance. Note that SSL is not required unless the network connecting the Repeater to the target machines is insecure.
End Time, Start Time	(Optional) Limits the set of event data files copied to the target Insight Server to those that contain data in the range defined by Start Time and End Time. If Start Time is set, event data files in which all log entries are from earlier than the specified start time are not copied. If End Time is set, event data files in which all log entries from the time specified or later are not copied. If only part of the data in a file is in the specified range, then the entire file is copied to the target machine. Adobe recommends using one of the following formats for the time: <ul style="list-style-type: none"> • January 1 2013 HH:MM:SS EDT • Jan 1 2013 HH:MM:SS GMT  Note: You must specify a time zone. The time zone does not default to system time if not specified. If you wish to implement Daylight Saving Time or a similar clock-shifting policy, you must save the .dst file containing the appropriate rules in the Base\Dataset\Timezone directory on the Insight Server machine. For a list of supported time zone abbreviations and information about implementing Daylight Saving Time, see Time Zone Codes on page 874.  Note: To use these settings, the names of the event data files must begin with an ISO date (YYYYMMDD), and each file must contain data for the 24-hour period starting at 12 AM GMT on that date.

8. Save your changes to the server by doing the following:

1. Right-click (**modified**) at the top of the window and click **Save**.
2. In the **Server Files Manager**, right-click the check mark for the file in the **Temp** column and select **Save to > <server name>**.

This example illustrates how files are copied if both the Flatten Paths and Recursive parameters are set to true.

Suppose Remote URI is /RemoteRoot/ and Local Path is E:\LocalRoot\. On the remote (**Repeater**) machine, the files are organized as follows:

- /RemoteRoot/fileA.txt
- /RemoteRoot/Dir1/fileB.txt
- /RemoteRoot/Dir2/Subdir3/fileC.txt

When replication is complete, the local directory has the following files:

- E:\LocalRoot\fileA.txt
- E:\LocalRoot\fileB.txt
- E:\LocalRoot\fileC.txt

In the local directory, no subdirectories are created, and all files from the remote directory tree are placed in the top level of the local directory.



Note: If files in the various subdirectories on the remote machine share the same name(s), the **Replication Service** may stop or other undefined behavior may occur.

Next Steps

Navigation title:Next Steps

You have completed the installation of the Insight Server**Replication Service**.

If you have not already installed **Repeater**, see *Repeater Functionality* on page 850 for installation instructions.

Repeater Functionality

Navigation title:Repeater Functionality

To configure repeater functionality, you must purchase an Insight Special Use License.



Note: Please contact Adobe Customer Care for more information.

Configuring Repeater Functionality

Navigation title:Configuring Repeater Functionality

Repeater functionality enables an Insight Server FSU to receive **Sensor**-acquired event data from one or more sources and replicate the data to one or more Insight Server FSUs running the Insight Server**Replication Service**.

See *Insight Server Replication Service* on page 847. This feature enables data replication to redundant facilities for disaster recovery purposes. The target servers act as network clients, connecting to the source FSU to request copies of the event data for inclusion in multiple datasets.

You can use repeater functionality in network infrastructures with multiple layers of firewalling to achieve single-port to single-port communications between components that are separated by firewalls.

For information about the system requirements for installing, configuring, and operating **Repeater**, see the *Minimum System Requirements* document.

To install **Repeater**, you must perform the steps listed for the following two procedures:

- *Configuring an Insight Server FSU for Repeater* on page 850
- *Configuring Access Control for Target Machines* on page 851

If network firewalls allow access to your **Repeater** from Insight, you can create a connection as described in *Creating a Connection Between Insight and Repeater* on page 852.

Configuring an Insight Server FSU for Repeater

Navigation title:Configuring an Insight Server FSU for Repeater

Instructions to install and configure an Insight Server FSU for use with Repeater.

Complete the following tasks in order. Any exceptions or changes that you must make so that the Insight Server FSU can be used as a repeater server are noted after each step.

1. Install the Insight Server program files as described in *Installing Data Workbench Servers* on page 792.
Because the machine on which you install these files is used to run Repeater, it is helpful to give the installation directory a descriptive name such as D:\Adobe\Repeater.
2. Install the Insight Server digital certificate as described in *Downloading and Installing the Digital Certificates* on page 793.

After you have logged in to the Adobe License Server, remember to look for the certificate name that matches the server common name of the designated repeater machine.

3. Check the port settings in the Communications.cfg file as described in [Checking the Port Settings](#) on page 802.

If the assigned ports (Port and SSL Port) are used by another process running on the same machine, you must change the port assignments to an unused pair.

4. If necessary, change the log directory for the **Sensor** data to be collected and stored on this machine. For instructions, see [Monitoring Event Data Space](#) on page 836.
5. Modify the Access Control.cfg file to allow administrative access to Insight Server from Insight as described in [Updating the Access Control File](#) on page 803.
6. Modify the server.address file to define the server's network location as defined in [Defining the Server's Network Location](#) on page 804.
7. Set Windows memory utilization parameters.
8. Register Insight Server as a Windows service as described in [Registering Insight Server as a Windows Service](#) on page 811.

Configuring Access Control for Target Machines

Navigation title: Configuring Access Control for Target Machines

Target Insight Server machines running the Insight Server Replication Service must be able to read the log files on this repeater server.

Access to the target machines is granted using the Access Control.cfg file.

To configure Access Control for access by target Insight Server machines

1. Navigate to the Access Control folder in the directory where you installed repeater functionality.
Example: D:\Adobe\Repeater\Access Control
2. Open Access Control.cfg in a text editor such as Notepad.
3. Create an access group for the Insight Server machines that must access the log files on this repeater server. Give this access group a name something like “Replication Targets.”

The following file fragment shows how the access group should look.

```
 . . .
6 = AccessGroup:
    Members = vector: N items
        0 = string: IP:Machine0IPAddress
        1 = string: IP:Machine1IPAddress
    . . .
    N = string: IP:MachineNIPAddress
    Name = string: Replication Targets
    Read-Only Access = vector: 1 items
        0 = string: EventDataLocation
    Read-Write Access = vector: 0 items
. . .
```

1. In the Members section, specify the IP address for each machine.
2. Update the items count for the Members vector to reflect the number of machine IP addresses you inserted.
3. In the Read-Only Access section, specify the location of the event data that the replication targets access. Use forward slashes in the path specification (/). The default location is the Logs folder on the Repeater machine (/Logs/).
4. Update the items count for the Read-Only Access vector to reflect the number of locations you inserted.

-
4. Update the number of access groups in the Access Control Groups vector at the top of the file to reflect the addition of the new access group.

```
Access Control Groups = vector< n items
```

5. Save the file.

Creating a Connection Between Insight and Repeater

Navigation title:Creating a Connection Between Insight and Repeater

If network firewalls do not prevent access to the repeater server from Insight machines, you can create a connection between the repeater server and Insight so that you can manage the repeater server using Insight.

To create a connection between Insight and the repeater server

1. In Insight, on the **Admin** tab, click the **Configure Connections to Servers** thumbnail to open the Configure Connections to Servers workspace.
2. In the **Insight.cfg** window, right-click **Servers** and click **Add new > Server**.
3. For the new server, complete the following parameters:

For this parameter...	Specify...
Name	(Optional) The name that you want this Insight to use to represent the repeater server in its user interface. Example: Repeater.mycompany.com or 192.168.1.90
Address	The host name or numeric IP address of your repeater server. Example: Samantha.Smith.pem If you leave this parameter blank, Insight uses whatever certificate is present.
SSL Client Certificate	Optional unless you have more than one certificate. The name of the file that contains the digital certificate for this copy of Insight. (This is the file that you downloaded while installing Insight.) Example: Samantha.Smith.pem If you leave this parameter blank, Insight uses whatever certificate is present.
SSL Server Common Name	The common name assigned to the repeater server. This name must match the common name assigned to the repeater server within its license certificate. If you have access to the repeater's certificate file (Certificates\server_cert.pem), you can find the common name by opening the file with a text editor such as Notepad. The common name is identified in the CN field in the certificate.

4. Save the file by right-clicking (**modified**) at the top of the window and clicking **Save**. Insight will attempt to connect to the repeater server using the settings you have specified. If a connection is established, a green server icon appears in the **Servers Manager** interface. If a connection cannot be established, a red icon appears.

For more information about the **Servers Manager** interface, see the *Insight User Guide*.

Next Steps

Navigation title:Next Steps

Instructions to complete your installation of the repeater functionality.

- On the servers from which you want to collect data, install **Sensors** as described in the *Insight Sensor Guide* and have them transmit data to the repeater server.
- If you have not set up the **Replication Service** for Insight Server, see *Insight Server Replication Service* on page 847 for instructions to install the service on an Insight Server FSU.

Administering Repeater

Navigation title:Administering Repeater

The administrative tasks for repeater functionality are very similar to those for Insight Server.

The following administrative tasks apply; exceptions or changes that you must make so that the Insight Server DPU can be used with the repeater server are noted after each step.

- [Re-validating the Digital Certificate](#) on page 826
- [Confirming that the Service Is Running](#) on page 827

In the list of services in the Services control panel, look for “Repeater.”

- [Configuring Access Control](#) on page 827
- [Monitoring Disk Space](#) on page 834

The data types that apply to the repeater server are event, operating system, and system data. If you are using the repeater server for backup storage and it becomes necessary to make more data storage space available on the machine, you can move all but the most current day’s log files to another machine or data storage medium (zip drive, tape, and so forth). Moving the data does not require you to stop the repeater service.



Note: You should verify the integrity of the backup copies of your .vs1 files before deleting any of them from Repeater.

- [Configuring Administrative Alerts](#) on page 839
- [Monitoring Administrative Events](#) on page 841
- [Monitoring Audit Logs](#) on page 842
- [Configuring Communications](#) on page 843
- [Restarting the Service](#) on page 844

Repeater functionality is designed to run continuously. If you restart the machine, repeater restarts automatically. If you need to start and stop repeater manually, you can do so using the Services control panel in Windows.

General Systems Administration

Navigation title: General Systems Administration

Adobe recommends that you administer and monitor on a regular basis the general health of each server on which an Adobe software product is installed.

General recommendations and instructions for monitoring and administering all of your Adobe machines is provided.

For information about administering specific Adobe software products, see the individual chapters in this guide.



Note: This guide, its recommended frequencies, and its instructions assume that you are using monitoring tools.

Confirming Your Systems Are Healthy

You must ensure that the machines on which Adobe server products are installed meet the minimum system requirements as defined in the Minimum System Requirements document.

Recommended Frequency: Every 5-10 minutes

You also must monitor your systems according to the best practices for operating that particular hardware, including, but not limited to, monitoring the following:

- CPU usage
- Disk space
- Hardware system messages
- Internal system temperature
- Memory usage

-
- Power supply conditions
 - RAID or disk controller performance and errors

Adobe recommends that you configure your management tool to alert administrators when any system parameter on a server machine exceeds the threshold you have set.

For Insight Server machines, Adobe also recommends that you configure each Insight Server to indicate when it reaches the minimum disk space limit you have set. For more information about these alerts, see [Configuring Administrative Alerts](#) on page 839.

Monitoring Events for Errors

Navigation title:Monitoring Events for Errors

To detect system and application errors as soon as possible and address them before they cause major problems or outages, you should regularly monitor your event logs.

Recommended Frequency: Every 5-10 minutes

To monitor your Adobe server software products, your automated management tool can be set to monitor your event log for errors with the source “Adobe” and then alert appropriate personnel to issues that may require intervention.

In Windows, application error messages are output to the Application Event Log in Windows, which you can access using the Windows Event Viewer. In Unix, application error messages are output to the Unix syslog using the LOG_DAEMON facility.

To open the Windows Event Viewer

- Click Start > Control Panel > Administrative Tools > Event Viewer.

Monitoring Ports and Applications

Navigation title:Monitoring Ports and Applications

To more thoroughly monitor your implementation, you can monitor all of the ports on your server machines as well as the software products running on each of those ports.

Recommended Frequency: Every 5-10 minutes

Using an application or script, you can monitor the TCP port on which each application is running (typically port 80 or 443) to make sure that the application is bound to that port. To do so, you request an application status page from the machine you want to monitor.

To request the application status page

1. On the machine you want to monitor, modify the Access Controls to allow your monitoring application or script to access the machine. For instructions, see [Configuring Access Control](#) on page 827.
2. Connect to `https://IP Address/Status/`, where IP Address is the IP address of the machine for which you want to receive status.

Example: `https://127.0.0.1/Status/`

The machine should respond with a server status description. If it does not respond, check your event logs and contact Adobe Customer Care.

For more information about this type of advanced monitoring, please contact Adobe Consulting Services.

Upgrading and Uninstalling Adobe Software

Navigation title:Upgrading and Uninstalling Adobe Software

Information about upgrading and uninstalling the Insight Server, **Transform**, and **Repeater** software products.

For information about upgrading and uninstalling other Adobe software products, see the appropriate product guide.

Upgrading Insight Server

Navigation title:

Instructions to upgrade your Insight Server, **Transform**, or **Repeater** software.



Note: If you have Platform 4.x or later, you can use these instructions to upgrade your Insight Server, **Transform**, or **Repeater** software. To upgrade the profiles (including your dataset profiles) for your Adobe application (such as Site), contact Adobe for assistance with your upgrade.

If you are using Insight Site v3.6 or earlier, contact Adobe for assistance with your upgrade.

DWB Server upgrade: 6.2 to 6.3

Upgrading server components for Data Workbench 6.3.

Upgrade Server

If you have customized profiles that take precedence over the default files provided in the Base package, then you will need to update these customized files:

- **Update the Meta.cfg file** (`E:\..\Profiles\<your custom profile>\Context\meta.cfg`) to set updated password encryption for the File System Unit (FSU server), and to add entries for the Name Value Pair transformations to take advantage of [DWB Server upgrade: 6.2 to 6.3](#) on page 188.

1. Open the `meta.cfg` file on the FSU.
2. Change the data type for **Proxy Password** from "string" to "EncryptedString" in the *Workstation Configuration* section.

```
Proxy User Name = string:  
Proxy Password = EncryptedString:    (from Proxy Password = String)  
Use Address File = bool: true
```

3. Add new entries to enable the new Name Value Pair transformations: `BuildNameValuePair` and `ExtractNameValuePairs`.

Open a workspace and right-click **Admin > Profile Manager**.

Under **Context**, click the **meta.cfg** file in the **Base** column and then click **Make Local**. From the User table column, right-click and select **Open > in Workstation**.



- In the new window, click **metadata** and add acceptable children templates.

+90	meta
-91	meta
-acceptable children	Template
-0	Column
-value	
Column Name	
Field Name	
name	Name Value
path	BuildNameValuePair:Input Columns
-92	meta
-acceptable children	Template
-0	Column
-value	
Column Name	
Field Name	
name	Name Value
path	ExtractNameValuePair:Output Columns
+93	meta

- Open transformation and add new templates.

-transformations	
+0	Template
-1	Template
-value	BuildNameValuePair
Name	
Comments	Comment
Condition	AndCondition
Delimiter	
Input Columns	
Output	
name	BuildNameValuePair
+2	Template
+3	Template
+4	Template
+5	Template
+6	Template
-7	Template
-value	ExtractNameValuePair
Name	
Comments	Comment
Condition	AndCondition
Delimiter	
Input Field	
Output Columns	
name	ExtractNameValuePair
+8	Template

-
- **Update for Fast Merge improvements.** Add parameters or change values to the following configuration files to take advantage of speed improvements in Data Workbench during a transformation.

- **Communications.cfg** (E:\Server\Components\Communications.cfg)

```
18 = SourceListServer:  
    URI = string: /SourceListServer/  
    Listing Interval = int: 10 (new)
```

- **Disk Files.cfg** (at E:\Server\Components and E:\Server\Components for Processing Servers)

```
Disk Cache Size (MB) = double: 1024 (from double: 256)  
Disk Cache Read Limit (MB) = double: 768 (new)
```

- **Log Processing Mode.cfg** (E:\Server\Profiles\<your profile>\Dataset\Log Processing Mode.cfg)

```
(changed)  
Batch Bytes = int: 268435456  
Cloud Bytes = int: 268435456  
Real Time FIFO Bytes = int: 268435456  
  
(new)  
Cache Bytes = int: 32000000  
Fast Input Decision Ratio = double: 200  
Fast Input FIFO Bytes = int: 268435456  
FIFO Hash Mask = int: 16383  
Fast Merge Buffer Bytes = int: 536870912  
Slow Merge Buffer Bytes = int: 268435456  
Fast Merge Fan In = int: 64  
Key Cache Size Logarithm = int: 21  
Max Seeks = int: 512  
Output Old Buffer Bytes = int: 536870912  
Overflow FIFO Bytes = int: 67108864  
Paused = bool: false
```



Note: To take advantage of the Fast Merge improvements, make sure you have at least 8 GBs of RAM per DPU.

- **Adobe Target with DWB integration update.** A new export file, ExportIntegration.exe, replaces the existing TnTSend.exe file on the Insight Server (E:\Server\Scripts\TnTSend.exe). This new export file supports both *Adobe Target* integration and coordination with the new Master Marketing Profile (MMP) and *Adobe Audience Manager*.

You will need to update the following commands for Adobe Target exports.

```
Command = string: TnTSend.exe  
to  
Command = string: ExportIntegration.exe
```



Note: This will only affect exports created prior to version 6.3.

You can also try the following to employ the old export process:

- Create a new Test And Target Export in the workstation.
- Modify the old Test and Target export found in Server/Profiles/<your profile>/Export.

-
- **Update the Adobe SC profile.** Changes to the `Exclude Hit.cfg` file require a field to be declared in the associated `Decoding Instructions.cfg` file.



Note: If your Adobe SC profile includes a customized `Decoding Instructions.cfg` file, you will need to include a `DelimitedDecoder` parameter to your customized file.

```
0 = DelimitedDecoder:  
  Delimiter = string: \t  
  Fields = vector: x items  
  ...  
  5 = string:  
Changed to:  
  5 = string: x-hit_source
```

Adding the `DelimitedDecoder` field allows you to take advantage of feature updates and to avoid possible Log Processing problems resulting from these updates.

DWB Server upgrade: 6.1 to 6.2

Upgrading server components for Data Workbench 6.2 and 6.2.2.

Upgrade Issues for 6.2

- The Attribution profile is configured for users who have implemented the Adobe SC profile to employ the Analytics (SC/Insight) data feed. By default, the Marketing and Conversion events are employed as the default interactions evaluated in the rules-based models. See [Deploying the Attribution Profile](#) for additional information.
- For users of the Adobe SC profile upgrading to Data Workbench 6.2, if you are not using the default configurations, verify that the `x-bot_id` value in the `SC Fields.cfg` file is being decoded properly and that the `x-bot_id` field is listed properly in the `Decoding Instructions.cfg` and the `Exclude Hit.cfg` files. This will only be an issue if you have modified the configuration file from the default configuration.
- If you have deleted unused fields in the `Dataset > Log Processing > SC Fields.cfg` file for the Adobe SC profile, you will need to update to accommodate updated field values used for the Attribution profile (see [Deploying the Attribution Profile](#)).

Upgrade Issues for 6.2.2

- The **Browsers** and **Operating Systems** lookup files will not be updated within the legacy **Traffic** profile (for example, `Lookups\Traffic\Browsers.txt`). Instead, configuration of the **Traffic** profile will utilize the DeviceAtlas bundle (`Lookups\DeviceAtlas\DeviceAtlas.bundle`) to provide this configuration information.
- Data Workbench 6.2.1 will be the last release to provide a download of the 32-bit client application. All future client application downloads will be 64-bit and continue to require Windows 7 or newer. Memory limitations of the 32-bit application are addressed with the introduction of the 64-bit application beginning with the 6.1 release.



Note: The 32-bit version of the Data Workbench client application may experience potential issues related to memory limitations when running predictive models using the clustering and scoring features.

Data Workbench 6.0 to 6.1 Upgrade

Follow these steps to update to data workbench v6.1 from your data workbench v6.0x installation.

Step 1: [Server Upgrade](#) on page 859

Step 2: [Report Server Upgrade](#) on page 860

Step 3: [Client Upgrade](#) on page 861



Important: The server, report server, and client components are upgraded to run on 64-bit Windows operating systems.

Server Upgrade

Follow these steps to update the **Server v6.1** components:

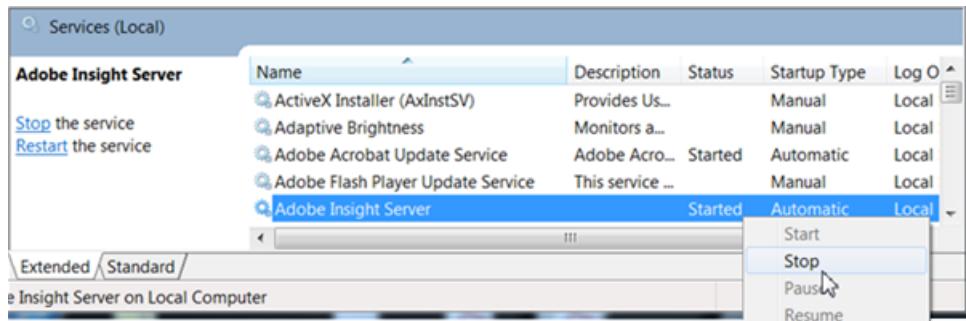
1. Using the **Software and Docs** profile, open the **Start Here** workspace and download all needed server packages to a local folder.

- Download **Server Packages \ v6.1** zip folders and extract all files.

The Server package includes **Lookup** and **Profile** folders with **Base** and **Transform** profiles to update the server.

- Download the **Profiles** folders.
- Download the **Lookup** folders.
- Download the **Report Server \ v6.1** package.
- Download additional **Sensor**, **Documentation**, and **Dashboard** files as needed for your system.

2. Stop the **Adobe Insight Server** service.



3. From the downloaded **Server** package:

- Replace the **Server\Bin** folder to update the **InsightServer64.exe** and supporting files.
- Replace the **Server\Profiles** folder. You can overwrite all files.
- Update the **Server\Lookups** folder. You will want to add the newly downloaded files to the custom files already located in the folder.
- Replace the **Server\Software** folder to update **Insight.exe** and **ReportServer.exe**
- Update the **Server\Scripts** folder to update **TnTSend.exe**.

4. If you employ **DeviceAtlas**, then you will need to [update the bundle](#) located in the **Server\Lookups** folder.
5. Configure the **Profile.cfg** file to ensure that the vector is updated to reflect the number of items for each profile.

For example, to enable the **Predictive Analytics** profile you will need to update this setting.

```
Directories = vector: 5 items
  0 = string: Base\\
  1 = string: Geography\\
  2 = string: Predictive Analytics\\
```

```
3 = string: Adobe SC\\
4 = string: Profile Name\\
```

6. Configure and save the PAServer.cfg file for the Predictive Analytics feature.

If you want to submit Predictive Analytics jobs to the servers, then you will need to configure the Server > Predictive Analytics > Dataset > PAServer.cfg file to manage server-side clustering submissions.

The custom profile should inherit the settings from the Predictive Analytics configuration profile, allowing you to configure and save the PAServer.cfg file based on your site's implementation.

7. Define the **Log Source ID**.

The **Recording of Rows per Log Source** was added in **v6.04** and defined in the custom profile's **Log Processing.cfg** file by adding a uniquely named **Log Source ID**.

```
Log Processing.cfg
  Log Source ID = string: <Name your ID Here>
```

If you do not have the Log Source ID defined, then you will get the following error:

```
Missing Log Source ID in log processing.cfg.
Log Source ID must be defined for all log sources.
```

8. Because the EventMessages.dll has been updated, it is required that you unregister and then register the **Adobe Insight Server** across the cluster.

- InsightServer64.exe /unregserver
- InsightServer64.exe /regserver

9. Start the **Adobe Insight Server** service across the cluster.

The server installation is now complete.

Report Server Upgrade



Important: Before upgrading to **Report Server v6.1**, you must first upgrade to **Server v6.1**.

1. Using the **Software and Docs** profile, download **v6.1** from the **Report Server** package to a local folder.
2. Copy **Report Server 6.1** from the downloaded package and replace the profile packages.



Note: The **Insight.zbin** file in the **install** folder is a backup file used for localization, and must be present in the **install** directory. This file or other **.zbin** files will be used depending on the command-line settings passed when starting up.

3. (optional) Data workbench currently supports English (-en-us) and Chinese (-zh-cn). You need to set a font to support single and double-byte characters:

```
Report Server.cfg - Add Fonts
  Fonts = vector: 2 items
  0 = string: SimSun
  1 = string: Arial
```

The Windows operating system must also have the listed fonts installed.

4. Configure Report Server v6.1 for localization.

- a. Stop the **Adobe Insight Report Server** service.
- b. Launch a command prompt as an "Administrator".
- c. Navigate to the Report Server install folder.

-
- d. Delete the Report Server service using the following command:

```
ReportServer.exe /unregserver
```

- e. Start the service based on language settings:

```
ReportServer.exe -RegServer -Locale -en-us (English)  
ReportServer.exe -RegServer -Locale -zh-cn (Simplified Chinese)
```

- 5. To verify that Report Server is running with the correct settings, open up **Windows Service Manager** and right-click **Adobe Insight Report Server - Properties**. The path to the executable will display the updated command-line settings.

The report server installation is now complete.

Client Upgrade



Important: Before upgrading to **Client v6.1**, the administrator must first upgrade to **Insight Server v6.1**.

- 1. Launch `Insight.exe` but do not connect to any profiles.
- 2. Edit the `Insight.cfg` file.

```
Update Software = bool: true
```

- 3. Connect to your profile.

Allow the client to synchronize with the server and your client will be upgraded with the latest v6.1client profiles, executables, and configuration files.



Note: The `Insight.zbin` file in the `install` folder is a backup file used for localization and must be present. This file or other `.zbin` files will be used depending on the command-line settings passed when starting up.

See [Installing the Input Method Editor](#) on page 133 to add an `insight.zbin` file required for localized settings.

Additional Client Settings

Before configuring `Insight.exe` and supporting files, you must exit the client application.

To install Simplified Chinese:

- 1. Create a shortcut that passes in the command-line setting to the `Insight.exe` file.

```
Insight.exe -zh-cn
```

- 2. Configure `Insight.cfg` to support single and double-byte font characters.

Data workbench currently supports both English and Simplified Chinese. You can select fonts to support both of these languages:

```
Fonts = vector: 2 items  
0 = string: SimSun  
1 = string: Arial
```

The Windows operating system must also have the requested fonts installed.

- 3. Launch the shortcut that you created to synchronize profiles and the updated `.zbin` file.

To employ the Input Method Editor (IME).

IME allows you to input international characters.

-
1. Update the `Insight.cfg` file with these settings:

```
Localized IME = bool: true
```

2. Launch the shortcut that you created to synchronize profiles and the updated `.zbin` file.

The client installation is now complete.

Installing the Input Method Editor

Data workbench now supports the Input Method Editor (IME) as a secondary text entry process for international languages.

IMEs allow you to enter international characters using a variety of methods suited for your local language. Data workbench provides an input dialog box that allows you to open and use your desired IME for text fields.



Note: For the data workbench 6.1 release, only the virtual Simplified Chinese keyboard will be supported. Inputting other languages through the IME could result in unexpected behavior.

Using an IME

To use the floating IME text input feature:

1. Click **Alt + Space** for any text input area.
2. Enter values using your system's IME.
3. Close the input dialog by selecting the **Enter** key or clicking the **OK** button.

The dialog will disappear and the characters will then appear in the selected field.

Updating the `Insight.cfg` file

To employ the IME, you must update the `Insight.cfg` file with this setting:

```
Localized IME = bool: true
```

If this setting does not exist in the configuration file, then pressing **Alt + Space** will not engage the IME feature.

Starting Insight in another language: To better support localized assets like a splash screen and to support multiple languages in the future, data workbench requires command-line arguments identifying the language to load. The default language is English.

Starting data workbench in Chinese requires you to invoke `Insight.exe` with the `"-zh-cn"` argument:

```
Insight.exe -zh-cn
```

(These command line arguments are not case sensitive.)

DeviceAtlas Distribution

The DeviceAtlas JSON file will now be distributed in a `.bundle` file (a renamed `.tar.gz`) along with `DeviceAtlas.dll` and `DeviceAtlas64.dll` files.

When the administrator upgrades the Insight Server to version 6.0, the `DeviceAtlas.bundle` file is included with the upgrade package in the Software and Docs profile (softdocs profile) located at:

```
Server Packages > v6.00 > Server_6.00.zip
```

The `DeviceAtlas.bundle` file is extracted to `Server\Lookups\DeviceAtlas`.

The `DeviceAtlas.bundle` file should be placed in a directory that is synchronized to the DPUs, and the `DeviceAtlas.cfg` file corresponding to the new `DeviceAtlasComponent` should be placed in the "Components for Processing Servers" directory on the synchronization master. When the `DeviceAtlas.bundle` file is changed, the very next `DeviceAtlas` lookup call will get results based on the updated API and/or JSON file.

Modify the Transformation.cfg file

The DeviceAtlas Transformations will no longer need to specify the path to the JSON file. Any previous DeviceAtlasTransformation that is defined in the transformation.cfg file should no longer include the File parameter that points to the obfuscated JSON file.

This example Transformation.cfg file shows the File argument that should be deleted to avoid confusion. (Leaving it there will not cause harm, but only potential confusion because it will be ignored.)

```
6 = DeviceAtlasTransformation:  
  Comments = Comment: 0 items  
  Condition = AndCondition: 0 items  
  
File = string: Lookups\\DeviceAtlas\\20110106_private.json.obfuscated  
^^ DELETE THE ABOVE LINE FROM ALL PREVIOUS TRANSFORMATIONS ^^  
  
Name = string: DeviceAtlas Lookup  
Outputs = vector: 4 items  
0 = Column:  
  Column Name = string: vendor  
  Field Name = string: x-vendor  
1 = Column:  
  Column Name = string: model  
  Field Name = string: x-model  
2 = Column:  
  Column Name = string: isBrowser  
  Field Name = string: x-isbrowser  
3 = Column:  
  Column Name = string: usableDisplayHeight  
  Field Name = string: x-usable-display-height  
User Agent = string: x-ua
```

Modify the DeviceAtlas.cfg file

This is an example of the component argument required in the DeviceAtlas.cfg file.

```
component = DeviceAtlasComponent:  
  DeviceAtlas Bundle File = string:Lookups\\DeviceAtlas\\DeviceAtlas.bundle  
  
  Unsyncronized Bundle Extraction Path = string: Temp\\DeviceAtlas\\
```

This DeviceAtlas.bundle file will be treated just like a configuration file from the perspective of the Profile Synchronization feature. In addition, the JSON data and DLL will be used at the Component level rather than at the individual Transformation level.

A new DeviceAtlasComponent, upon startup, finds the .bundle conglomeration, de-obfuscates the JSON file into memory, extracts the files into a temporary directory, and loads the appropriate DLL for the running platform. This component also monitors changes to the bundle file, and reloads the DLL and .cfg file automatically if it changes.

Running DeviceAtlas

Proper configuration makes a big difference in the time required for transformation. The transformation can be configured to run only once per visitor per session to allow DeviceAtlas to speed up the process.

If deployed using Log Processing.cfg:

Run the transformations twice.

1. Look up only the mobile_id field, then
2. Create conditions to ignore the mobile_id and then look up the rest of the fields.

If deployed using Transformation.cfg:

Deploy as in Step 1 in Log Processing above, or use cross-rows to support a conditional setting.

- Cross-Rows—Grab the previous session key. Then identify if the current session key is different from the one found with cross-rows. If so, then the DeviceAtlas transformation will only run on one record per session.

Data Workbench 5.5 to 6.1 Upgrade

Follow these steps to update to data workbench v6.1 from your Insight v5.5x installation.

Step 1: [Server Upgrade](#) on page 864

Step 2: [Report Server Upgrade](#) on page 865

Step 3: [Client Upgrade](#) on page 866



Important: The server, report server, and client components are upgraded to run on 64-bit Windows operating systems.

Server Upgrade

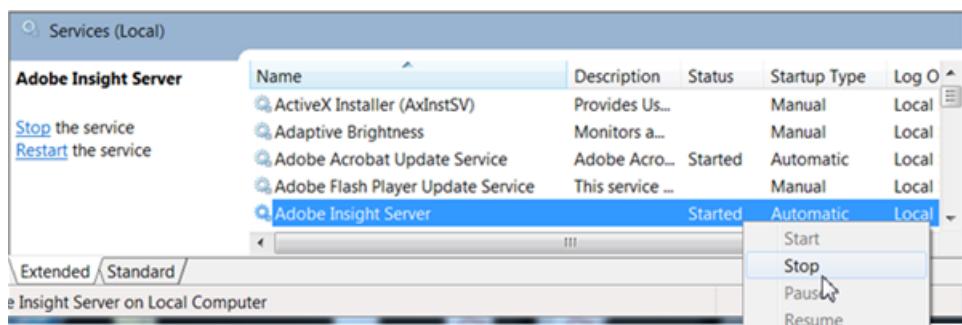
Follow these steps to update the **Server v6.1** components:

1. Using the **Software and Docs** profile, open the **Start Here** workspace and download all needed server packages to a local folder.
 - Download **Server Packages \ v6.1** zip folders and extract all files.

The **Server** package includes **Lookup** and **Profile** folders with the **Base** and **Transform** lookup files to add and replace to update the server.

- Download new **Profiles** folders.
- Download updated **Lookup** folders.
- Download the **Report Server \ v6.1** package.
- Download additional **Sensor**, **Documentation**, and **Dashboard** files as needed for your system.

2. Stop the **Adobe Insight Server** service.



3. From the downloaded **Server** package:
 - a. Replace the **Server\Bin** folder to update the **InsightServer64.exe** and supporting files.
 - b. Replace the **Server\Profiles** folder. You can overwrite all files.
 - c. Update the **Server\Lookups** folder. You will want to add the newly downloaded files to the custom files already located in the folder.
 - d. Replace the **Server\Software** folder to update **Insight.exe** and **ReportServer.exe**

-
- e. Update the Server\Scripts folder to update TnTSend.exe.
 4. If you employ **DeviceAtlas**, then you will need to *update the bundle* located in the Server\Lookups folder.
 5. Set Directories in the Profile.cfg file to ensure that the vector is updated to reflect the number of items for each profile.

For example, to enable the **Predictive Analytics** profile you will need to update this setting.

```
Directories = vector: 5 items
  0 = string: Base\\
  1 = string: Geography\\
  2 = string: Predictive Analytics\\
  3 = string: Adobe SC\\
  4 = string: Profile Name\\
```

6. Configure and save the PAServer.cfg file to upgrade the Predictive Analytics feature.

If you want to submit Predictive Analytics jobs to the servers, then you will need to configure the Server > Predictive Analytics > Dataset > PAServer.cfg file to manage server-side clustering submissions.

The custom profile should inherit the settings from the Predictive Analytics configuration profile, allowing you to configure and save the PAServer.cfg based on your site's implementation.

7. Define the **Log Source ID**.

The **Recording of Rows per Log Source** was added in **v6.04** and defined in the custom profile's Log Processing.cfg file by adding a uniquely named **Log Source ID**.

```
Log Processing.cfg
  Log Source ID = string: <Name your ID Here>
```

If you do not have the Log Source ID defined, then you will get the following error:

```
Missing Log Source ID in log processing.cfg.
Log Source ID must be defined for all log sources.
```

8. Because the EventMessages.dll has been updated, it is required that you unregister and then register the **Adobe Insight Server** across the cluster.
 - InsightServer64.exe /unregserver
 - InsightServer64.exe /regserver
9. Start the **Adobe Insight Server** service across the cluster.

The server installation is now complete.

Report Server Upgrade

 **Important:** Before upgrading to **Report Server v6.1**, you must first upgrade to **Server v6.1**.

1. Using the **Software and Docs** profile, download **v6.1** from the **Report Server** package to a local folder.
2. Copy **Report Server 6.1** from the downloaded package and replace the profile packages.



Note: The Insight.zbin file in the install folder is a backup file used for localization, and must be present in the install directory. This file or other .zbin files will be used depending on the command-line settings passed when starting up.

3. (optional) Modify the report server configuration file to support double-byte characters.

Data workbench currently supports English (-en-us) and Chinese (-zh-cn). You need to set a font to support single and double-byte characters:

```
Report Server.cfg - Add Fonts
  Fonts = vector: 2 items
    0 = string: SimSun
    1 = string: Arial
```

The Windows operating system must also have the listed fonts installed.

4. Configure Report Server v6.1.

- a. Stop the **Adobe Insight Report Server** service.
- b. Launch a command prompt as an "Administrator".
- c. Navigate to the Report Server install folder.
- d. Delete the Report Server service using the following command:

```
ReportServer.exe /unregserver
```

5. Start the service based on the language settings:

```
ReportServer.exe -RegServer -Locale -en-us (English)
ReportServer.exe -RegServer -Locale -zh-cn (Simplified Chinese)
```

6. To verify that Report Server is running with the correct settings, open up **Windows Service Manager** and right-click **Adobe Insight Report Server - Properties**. The path to the executable will display the updated command-line settings.

The report server installation is now complete.

Client Upgrade



Important: Before upgrading to **Client v6.1**, the administrator must first upgrade to **Server v6.1**.

1. Launch `Insight.exe` but DO NOT connect to any profiles.
2. Edit the `Insight.cfg` file to not update software automatically.

```
Update Software = bool: false
```

3. Connect to **Software and Docs** profile (`softdocs`).
4. Download `Software\Insight Client\v6.10`.
5. (optional) Modify `insight.cfg` to support double-byte characters.

Data workbench currently supports both English and Simplified Chinese. Select fonts to support both of these languages:

```
Fonts = vector: 2 items
0 = string: SimSun
1 = string: Arial
```

6. Exit out of the client.
7. Copy the files in the downloaded **v6.1** client package to the `Install` folder.



Note: The `Insight.zbin` file in the install folder is a backup file used for localization, and must be present in the install directory. This file or other `.zbin` files will be used depending on the command-line settings passed when starting up.

For example, to launch Simplified Chinese, create a shortcut that passes in the command-line setting.

```
Insight.exe -zh-cn
```

If you want to launch in English (default), then no command-line change is necessary.

8. Launch `Insight.exe` for English or the shortcut that you created for another language.
9. Connect to your profile and allow the client to synchronize with the server.
10. (optional) To employ the IME, make these changes to the `Insight.cfg` file:

```
Localized IME = bool: true
```

The Input Method Editor (IME) allows you to input international characters.

11. (optional) Edit the `Insight.cfg` file to automatically update software:

```
Update Software = bool: true
```

See instructions for implementing the IME.

12. Restart again after the profile synchronization to employ the most recent `.zbin` file.

The client installation is now complete.

DWB Server upgrade: 5.4 to 5.5

Upgrading server components for Data Workbench 6.1 from 5.4 installation.

Consequently, upgrading from Insight 5.4 to Insight 5.5 is relatively simple.

You can also upgrade directly from Insight 5.3 to Insight 5.5 using the steps below. Make sure you perform all upgrade tasks listed in the [DWB Server upgrade: 5.4 to 5.5](#) on page 867 section and the [DWB Server upgrade: 5.4 to 5.5](#) on page 867 section.

1. Stop the Insight Server services on all servers in the cluster except for the Insight Master Server.
 1. Copy the new `ReportServer.exe` and `Insight.exe` files to the `Software\Insight` folder.
 2. After the Insight client has updated, copy the `InsightServer.exe` and `InsightServer64.exe` files into the `\Bin` folder.
 3. Wait for the Insight Master Server to start, then verify the version running via the **Connections** visualization.
2. On the cluster's Master Server in the Insight client:
 1. Make a local copy of the existing **Base** profile and rename it (for example, `BaseBackup`).
 2. Copy the new **Base** profile to the `Profiles` folder.
 3. Repeat these two steps for the `Transform` folder.
3. Copy the `Scripts` folder from the server package onto the Master Server into the Server installation directory.
4. Copy the `Terrain_Images.cfg.off` file into the `Components` folder of the Master Server.

To upgrade client software from Insight 5.4 to 5.5

In the `Insight.cfg` file, make sure that the `Update Software` setting is set to TRUE.

Upgrading Transform

Navigation title:Upgrading Transform

Steps to upgrade the `Transform` folder.

1. Open the `.zip` file for the Insight Server release package, and open the `Profiles` folder within that `.zip` file.
2. Copy the `Transform` folder to the `Profiles` folder in your Insight Server installation directory. Doing this overwrites the existing `Transform` folder.
3. For each profile that inherits the **Transform** profile, confirm that the `profile.cfg` file has a “`Transform`” entry in the Directories vector.

Data reprocessing begins after synchronization of the profile.

Upgrading Repeater

Navigation title:Upgrading Repeater

Instructions to upgrade **Repeater** using Insight or to upgrade by copying files.

You can use one of the following methods to upgrade **Repeater** from Platform 4.x or later:

- If you created a connection between Insight and **Repeater** as described in *Creating a Connection Between Insight and Repeater* on page 852, you can use Insight to upgrade **Repeater**. See *Upgrading Repeater Using Insight* on page 868.

-or-

- If you could not or did not create a connection between Insight and **Repeater**, you must copy the upgrade files directly to the **Repeater** machine. See *Upgrading Repeater by Copying Files* on page 868.

Upgrading Repeater Using Insight

1. On the Insight computer, copy the `bin` folder from the Insight Server upgrade package to the Temp folder in the directory where Insight is installed.
2. Start Insight.
3. In Insight, on the **Admin > Dataset and Profile** tab, click the **Servers Manager** thumbnail to open the Servers Manager workspace.
4. Right-click the icon of the **Repeater** you want to upgrade and click **Server Files**.
5. In the **Server Files Manager**, do the following to upload the upgrade files to the server:
 - a. Locate the `bin` folder.
 - b. Right-click the check mark for the `bin` folder in the Temp directory and select **Save Directory to > <server name>**.



Note: You will see a message indicating that this step overwrites files of the same name that exist on the server. Click **Yes** to continue.



Note: If you need to upload additional files to complete your **Repeater** upgrade, Adobe will provide instructions to do so.

After you upload the upgrade files to the **Repeater** machine, **Repeater** automatically restarts itself and loads the new version.

Upgrading Repeater by Copying Files

1. Open the upgrade file provided by Adobe. Most likely, this file is a `.zip` file for upgrading Insight Server.
2. Go to the directory where you installed **Repeater** (for example, `D:\Adobe\Repeater`).
3. Copy the `bin` folder within the `.zip` file and paste it in your **Repeater** installation directory to overwrite the existing `bin` folder.



Note: If you need to upload additional files to complete your Insight Server upgrade, Adobe will provide instructions to do so.

After you copy the upgrade files to the **Repeater** machine, **Repeater** automatically restarts itself and loads the new version.

Uninstalling Your Software

Navigation title:

Instructions to uninstall Insight Server, **Transform**, or **Repeater**.

Uninstalling Insight Server Adobe

1. Unregister the Insight Server Windows service.
 - a. Open a command prompt and navigate to the bin sub-directory in the folder where you installed Insight Server.

Example: C:\Adobe\Server\bin

- b. At the command prompt, execute the following command to stop and unregister it as a service under Microsoft Windows:

```
InsightServer64.exe /unregserver
```

2. Delete the Insight Server installation directory.

Uninstalling Transform

1. Use the following steps to update the `profile.cfg` file for each profile with which you were using **Transform**.
 - a. Open the **Profile Manager**.
 - b. Right-click the check mark next to `profile.cfg` and click **Make Local**. A check mark for this file appears in the **User** column.
 - c. Right-click the newly created check mark and click **Open > in Insight**. The `profile.cfg` window appears.
 - d. In the `profile.cfg` window, delete the **Transform** profile entry from the Directories vector.
 - e. Right-click (**modified**) at the top of the window and click **Save**.
 - f. In the **Profile Manager**, right-click the check mark for `profile.cfg` in the **User** column, then click **Save to > <profile name>**.
2. Delete the **Transform** folder from the **Profiles** folder in your Insight Server installation directory.

Uninstalling Repeater

1. Unregister the **Repeater** Windows service.
 - a. Open a command prompt and navigate to the bin sub-directory in the folder where you installed **Repeater**.
- Example: D:\Adobe\Repeater\bin
- b. At the command prompt, execute the following command to stop and unregister it as a service under Microsoft Windows:

```
InsightServer64.exe /unregserver
```

2. Delete the **Repeater** installation directory.

Configuration Settings Reference

Navigation title: Configuration Settings Reference

Information about Insight Server directory structure, communications configuration settings, and more.

Insight Server Directory Structure

Navigation title: Insight Server Directory Structure

List of files installed with Insight Server and the files present after it has been registered, and run for the first time.

Files Included in the Installation Package

The following directories are included in the Insight Server installation package:

Directory	Description of Directory Contents
Access Control	Insight Server configuration file which specifies a list of Access Groups.
Addresses	Address(es) used for communication with Insight Server.
Audit	Daily access logs containing details regarding all attempted connections to Insight Server.
bin	Insight Server executable program files.
Certificates	SSL digital certificates.
Components	Insight Server component configuration files.
Components for Processing Servers	Insight Server component configuration files for processing Insight Servers within an Insight Server cluster.
Events	Daily event logs containing detailed event status messages, including error messages. Events captured and logged by Insight Server are also displayed in Windows Event Viewer.
Logs	Log files produced by Sensor(s) . “Logs” is the default logging directory, but an alternate directory may have been specified in the communications.cfg file.
Lookups	Lookup files, such as robot and search engine lists. Insight Server must load all lookup files into memory. The total size of all lookup files referenced in component configuration files, plus overhead (for example, 12 bytes per row for FlatFileLookup files), must not exceed the available physical or virtual memory that is available after all other software applications are loaded.
Profiles	Files related to each profile (configuration, workspace, and visualization files). Profiles are populated by data from a dataset. Datasets include event data (“Log Data”); such data may be captured by installed Sensors , transmitted by Web beacons or page tags, or input from data warehouses. Insight users with access to a given profile may use the set of processed data for that profile as well as the Workspaces and visualizations defined within that profile. Workspaces are work areas for system administration or analysis. A Workspace can contain multiple interfaces showing different details about system performance.
Software	Insight software updates. Report software updates are also stored here.

Directories and Files Created after Startup

The directories listed below are created after Insight Server is registered and run for the first time:

Insight Server Startup Files

Directory	Description
State	Processing information generated by Insight Server.

Directory	Description
Temp	Location of the temporary files used by Insight Server during reprocessing and operation. There is usually one file (named <code>temp.db</code> by default) per physical drive. Insight Server must be configured to write to this directory.
Trace	Log and event data about Insight Server. Useful for troubleshooting.
Users	Named (Insight) users with access to the profiles on the Server. A directory for each authorized named user is created within the directory <code>Users\</code> when the user first accesses Insight Server via Insight. The directory for each named user contains directories corresponding to all of the profiles that the user has accessed on that Insight Server as well as their local address files.

Administrative Alerts Configuration Settings

Navigation title:Administrative Alerts Configuration Settings

Instructions to configure administrative alerts for Insight Server, **Repeater**, or **Transform**.

Complete the parameters in the following file:

Product Name installation directory\Components\Administrative Alerts.cfg

Administrative Alert Configuration Settings

Parameter	Description
Category	The name of the category. A category of Default is required. See Error Categories in this table.
Error Categories	Lets you categorize errors in conjunction with the Error Categorization File. Each Error Category can have its own set of Recipients and its own Throttle Delay. For example, you might create a Critical category with a throttle delay of 0, so that every critical error is emailed immediately to the recipients specified in the Recipients list. Errors that do not match a substring in the Error Categorization File are assigned to the Default category. To add a new category, right-click on a number and click Add New > Error Category . You can also copy or remove them using the right-click action.
Error Categorization File	The name of the file you want to use to categorize each alert. You create this file using Notepad. This file should have three columns on each line, separated by tabs. The first column is a string to match in errors. A ^ sign matches the beginning and a \$ matches the end of the string; all other characters are matched literally. The second column is a category for errors that match, which is in Error Categories. The third is an alternate message, which is prepended to the actual error message in emails that are sent. If no file is specified, all errors categorize as Default. To see an example of this file, see the <code>Error_Categories.txt</code> file in the Lookups directory.
From	Address that appears in the “from” parameter of the email message. Example: <code>server_errors@mycompany.com</code>
Minimum Disk Space (MB)	The server generates an email alert when available disk storage in any directory used by the server drops below this value. The default value is 1000.
Sensor Alert Timeout (min)	The server generates an email alert when it has received no data from a configured and previously connected Sensor within this time window. The default value is 15.  Note: Sensor Alert Timeout works only if an existing connection to a Sensor is dropped. If the server's service is stopped and restarted and the Sensors do not connect, the server does not generate email alerts.
Server Address	The address of the SMTP server for outgoing email. Example: <code>mail.mycompany.com</code> An SMTP server is required for use of the described capabilities.

Parameter	Description
Server Password	The password for logging in to the SMTP server. This parameter is optional unless login is required to send mail. An SMTP server is required for use of the described capabilities.
Server User	The user ID/name for logging in to the SMTP server. This parameter is optional unless login is required to send mail. An SMTP server is required for use of the described capabilities.
Throttle Delay (secs)	The minimum number of seconds that must elapse between two errors in that category for an email to be sent. A value of 0 sends the email immediately.

Communications Configuration Settings

Navigation title: Communications Configuration Settings

Instructions to configure communications for Insight Server or **Repeater**.

Complete the parameters in the following file:

Product Name installation directory\Components\Communications.cfg



Note: Before modifying any parameters not listed in this table, please contact Adobe.

Communications Configuration Settings

Parameter	Description
Access Control File	Location of the <code>Access Control.cfg</code> file. The default location is the <code>Access Control</code> folder within the Insight Server or Repeater installation directory. Example: <code>Access Control File = Path: Access Control\\Access Control.cfg</code>
Access Log Directory	Folder to which you want to map the audit logs. Example: <code>Access Log Directory = string: Audit\\</code>  Note: You can map audit logs to another local drive (example: <code>string: P:\\Audit\\</code>), but do not map audit logs to a network drive.
Access LogVerbose	This parameter can be set to true or false. It is used to enable and disable audit log filtering. To ensure that every request is logged, set the parameter to True.
IP Interface	IP address to use when two network cards are available for accessing two different networks. Example: <code>I</code> <code>P Interface = string:</code> <code><IP Address></code>
Port	Non-secure (HTTP) port on which the Insight Server or Repeater listens. The default port is 80. Entering a value of 0 disables non-secure connections. Example: <code>Port = int: 80</code>

Parameter	Description
SSL Ciphers	<p>Some environments require stronger communication security than others. If you want to use a specific SSL cipher suite, you can specify it with this parameter.</p> <p>Example:</p> <pre>SSL Ciphers = string: AES256-SHA256</pre>
SSL Port	<p>Secure (via SSL) port on which the Insight Server or Repeater listens. The default port is 443. Entering a value of 0 disables secure connections.</p> <p>Example:</p> <pre>SSL Port = int: 443</pre>
<i>n</i> =LoggingServer:	Heading for Logging Server settings.
Customer Name	<p>Customer name to appear for Unspecified customers in administrative alerts, as in the following example: "No data received from sensor XYZ for customer 'Unspecified' in 15."</p> <p>Example:</p> <pre>1 = LoggingServer: Customer Name = string: CompanyAB</pre> <p>Using the example above, administrative alerts for Unspecified customers would now read as follows: "No data received from sensor XYZ for customer 'CompanyAB' in 15."</p>
<i>n</i> =FileServer: Local Path = string: Logs\\	<p>Folder in which you want to store the log files.</p> <p>Example:</p> <pre>9 = FileServer: Local Path = string: Logs\\</pre> <p>To be able to access this folder from the Server Files Manager, the location specified in this parameter must match the location that you specify in the Log Paths parameter in the Log Processing.cfg file. For more information about modifying the Logs directory in the Log Processing.cfg file, see Log Processing Configuration File chapter of the <i>Dataset Configuration Guide</i>.</p>
<i>n</i> =FileServer: Local Path = string: Audit\\	<p>Folder to which you want to map the audit logs.</p> <p>Example:</p> <pre>5 = FileServer: Local Path = string: Audit\\</pre> <p> Note: You can map audit logs to another local drive (example: string: P:\\Audit\\), but do not map audit logs to a network drive.</p> <p>To be able to access this folder from the Server Files Manager, the location specified in this parameter must match the location that you specify in the Access Log Directory parameter in this file.</p>
<i>n</i> =NormalizeServer:	<p>This parameter applies only to Insight Server.</p> <p>For more information about specifying the Centralized Normalization Server for your Insight Server cluster, see Log Processing Configuration File chapter of the <i>Dataset Configuration Guide</i>.</p>
<i>n</i> =ReportStatusServer: URI = string: /ReportStatus.vsp	<p>This parameter applies only to Insight Server.</p> <p>Enables you to view Report's status in the Detailed Status interface for Insight Server.</p>

DPU Performance Settings

Navigation title:DPU Performance Settings

Instructions to tune DPU performance.

Complete the following parameters in the *Insight Server installation directory\Components\DPU.cfg* file.

DPU.cfg Configuration Settings

Parameter	Description
Execution Batch Count	This is a tuning parameter. The default value is 65536. You can specify arbitrarily small execution batch counts. Please contact Adobe before making any changes to this value.
Execution Batches	This is a tuning parameter. The default value is 128. Please contact Adobe before making any changes to this value.
Execution Time	This is a tuning parameter. The default value is 0.100. Please contact Adobe before making any changes to this value.
Field Dump on Error	This parameter can be set to true or false. If set to true, Insight Server creates a file named Field Dump <number>.txt in its Trace directory whenever execution engine errors occur. The Field Dump <number>.txt is useful for troubleshooting.
Profile Path	Location in which to store files for all profiles. The default location is Profiles\ .
Query Memory Limit	Amount of memory (in bytes) that Insight Server reserves to store query results. The default value is 100000000 (100MB.) If more space for query results is required (for example, to allow more simultaneous users), the setting can be increased, but you must continue to check the Insight Server's memory load.
State Path	Location of system state files. The default location is State\ .
Threads	A performance tuning parameter for Insight Server machines with multiple processors. In general, for any n-core system, this value should be set to n. The default value is 1.
User Path	Location of authorized users' files. The default location is Users\ .

Time Zone Codes

Navigation title:Time Zone Codes

Format instructions about time-based parameters in Insight Server.

Most time-based parameters in Insight Server are specified in the following format:

Month DD, YYYY HH:MM:SS TimeZone

Example: August 13, 2013 22:30:00 EST

Time zones are expressed in a system-independent time zone format (Coordinated Universal Time) of the following format:

UTC +hhmm dstrules

The sign (+) can be either a plus (+) or a minus (-) sign, and *hhmm* is the offset from UTC in hours and minutes. The optional variable *dstrules* specifies a set of rules to implement Daylight Saving Time or a similar clock-shifting policy.

If you specify *dstrules*, a tab-delimited file named *<dstrules>.dst* must be present within the *Dataset\TimeZone* directory of either the Base profile (for configuration files that are not associated with a particular dataset) or the dataset profile (for configuration files that are dataset-specific). The file specifies a time-zone independent set of rules for Daylight Saving Time. You can have different sets of rules for different years. The **DST.dst** file provided by Adobe in the Base profile specifies the standard U.S. rules established by the Energy Policy Act of 2005 (in effect starting 2007) and the U.S. rules for prior years.

Sample time zone entries are listed below:

-
- U.S. Eastern Daylight Time: Time Zone = string: UTC -0500 DST
 - UTC time with no offset and no *dstrules* (corresponding to GMT): Time Zone = string: UTC -0000

When this format is used, the system time zone of Insight Server, Insight, and Report machines need not be the same as the specified time zone. In addition, all active dataset profiles on an Insight Server machine need not have the same time zone setting.

The following table contains the list of codes you can use to specify time zones in time-based parameters.

Time Zone Code Table



Note: If you are implementing Daylight Saving Time or a similar clock-shifting policy, you must save the .dst file containing the appropriate rules in the *profile name\Dataset\Timezone* directory on the Insight Server machine.

Time Zone Codes

Code	Time Zone	Offset from GMT
gmt	Greenwich Mean	0
est	Eastern Standard	5
edt	Eastern Daylight	5
cst	Central Standard	6
cdt	Central Daylight	6
mst	Mountain Standard	7
mdt	Mountain Daylight	7
pst	Pacific Standard	8
pdt	Pacific Daylight	8

Sensor Overview

Navigation title:Sensor Overview

Sensor is a lightweight, low-impact data acquisition software application that reliably, securely, and continuously collects event data from web, application, and data collection servers to be transmitted to any data workbench server.

Sensor's instrumentation ensures consistently accurate measurement of events that occur in your Internet channel.

How Does the Data Collection Process Work?

Navigation title:How Does the Data Collection Process Work?

Sensor automates the acquisition of data from your Internet channel by doing away with the bulk of human labor traditionally involved in data collection.

In many cases, using **Sensor** can vastly simplify your data management process.

Today's large Internet, extranet, and intranet sites often run on an array of web servers. The logs and data produced can be very large and cumbersome to manage. For example, if your site is running 30 web servers, typically one of your employees (or outsourced service provider's employees) would pull and consolidate each log file on

each of the 30 servers, then run reports on them. Installing **Sensor** on each of your web servers automates this entire process, reducing your expenses and making data available in real time.

To automate this process, **Sensor** collects raw information about the traffic on a website directly from each web server. The raw data that **Sensor** captures is called event data and is similar to the type of data that your web server records in its log files.

To capture this data, instrumentation within **Sensor** records information about each HTTP request that your web server processes. **Sensor** then buffers the information to protect against network failure and securely transmits the information via HTTP/S to the data workbench server that you specify.

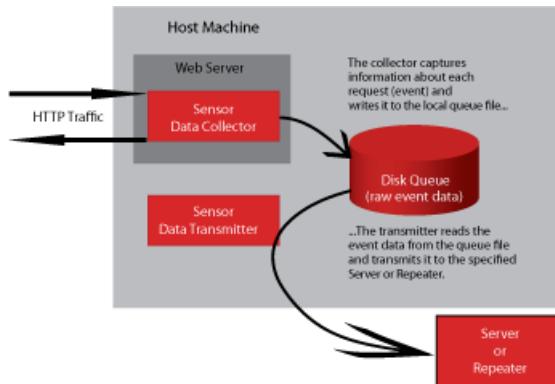
After the data workbench server receives the data, it processes and stores your log files in highly-compressed .vs1 format files, allowing you to easily maintain very large amounts of data on inexpensive hardware.

For information about the event data fields collected by **Sensor** in .vs1 files, see [Event Data Record Fields](#) on page 951.

What are Basic Components

Navigation title:What Are the Basic Components?

Sensor consists of three main components: Data Collector, Disk Queue, and Data Transmitter.



Data Collector

The data collector (collector) is an NSAPI, ISAPI, J2EE filter servlet, or Apache module that executes within the web server process.

It captures the raw event data about each HTTP request that the web server processes and deposits that information into the disk queue. If you run multiple instances of a web server on the same machine, each instance loads its own instance of the collector module; however, all instances of the collector write their event data to the same disk queue.

Disk Queue

The disk queue (queue) is a fault-tolerant, FIFO (first in, first out) memory-mapped queue file that buffers the raw event data that Sensor collects, providing temporary storage for collected event data on the web server where it is installed.

To prevent the queue from expanding without constraint (and thereby consuming all available disk space), the queue is maintained in a fixed-size file, meaning that it holds only as much event data as it has been given capacity to hold. The size of the queue file is configured in the QueueSize parameter of the Sensor configuration file, txlogd.conf, when the Sensor is installed. For information about the txlogd.conf parameters, see [Sensor Txlogd.conf File Parameters](#).

Once established, the physical length of the file does not grow or shrink. The collector simply deposits new event data into the queue and the transmitter pulls events from it. If the collector reaches the end of the file, it stops writing to the queue file.

Generally, the transmitter pulls events from the queue as quickly as the collector deposits them. However, if the connection between the transmitter and Insight Server is slow or unavailable, the queue can fill with untransmitted events. In this event, the collector stops collecting data until the transmitter draws down the queue. Information about requests that the web server processes during this time is permanently lost.

Determining the Queue Size

Before you install Sensor, you must determine how large the queue needs to be. To prevent permanent data loss, it is important to create a queue that is large enough to accommodate the number of events that could accumulate during longest likely outage of the connection to the Insight Server (that is, sufficient storage for several days of peak activity). The queue must be configured to hold enough event data so that system administrators have time to restore network accessibility to the target Insight Server, or repair or replace the Insight Server without losing any data. If the Sensor has failed and a valid and accessible queue file is not available to hold the event data, subsequent data is lost.



Note: It is important that the administrators of each machine on which Sensor runs understand the unique nature of the local queue file to ensure that they do not treat it as an ordinary log file that can be deleted, archived, or compressed.

Adobe recommends that the queue be configured to hold at least ten (10) peak days of event data that is produced by the server where the Sensor is installed. That is, take the amount of event data from any peak day in the last year and multiply it by ten.

This recommendation assumes the following:

- Your company's Information Technology team is monitoring each Sensor in the manner detailed in Administering Sensor, of this guide and is doing so at least once per day. If this is not the case, this period should be extended appropriately.
- Your company's Information Technology team is able to restore network accessibility or replace or repair any installed Insight Servers within 72 hours. If this is not the case, this period should be extended appropriately.
- The configuration of Sensor remains the same.
- No external events (for example, a large marketing campaign) will cause the amount of event data being generated by the web servers to increase significantly.

Your choice of queue size depends largely upon the desired level of system monitoring in light of your company's practices and policies regarding response times and weekend/holiday system administration. As larger queue sizes are better, Adobe recommends that your company make the queue as large as possible.



Note: Larger queue file sizes have no impact on performance.

For further recommendations about sizing the queue, contact Adobe Consulting Services.

Data Transmitter

The transmitter is an independent process (for example, a daemon on a UNIX-based computer or a service on a Windows computer) that executes on the same machine as the web server.

The transmitter reads the event data from the disk queue, compresses it, and sends it via HTTP/S to the Insight Server that you have specified, where it is processed and stored in .vsl files.

How Does Sensor Identify Visitors and Sessions?

Navigation title:How Does Sensor Identify Visitors and Sessions?
--

When collecting event data from a web server, **Sensor** automatically sets a persistent cookie for each visitor containing a small random identifier, without capturing any personally identifying information.

This Adobe cookie is used to identify the unique visitor and all of its related sessions.

By default, **Sensor** captures all of the W3C Extended Log File Format fields from each HTTP header, but you can configure **Sensor** to capture any header that is transmitted between the client and server. For information about the event data fields collected by **Sensor** in .vs1 files, see *Event Data Record Fields* on page 951.

Some visitors' browsers do not store cookies persistently, and a very small number of visitors' browsers do not accept cookies at all (even session cookies). Even though they account for only a fraction of a site's total traffic, they can result in significant miscounting if each page view by such a visitor is counted incorrectly as an entire session, as is done by some log file analysis software. Adobe addresses this problem by enabling you to analyze visitors with and without using cookies.

For more information about the Broken Sessions Filter, see the *Data Workbench Sensor Guide*.

Installing Sensor

Navigation title:Installing Sensor

You install **Sensor** on the same computer as the server whose activity you want to measure.

Each server from which you want to capture event data must run **Sensor**.

Sensor can be installed on a wide variety of supported web and application servers or on specialized data collection servers that are used to acquire information from pages, advertisements, and other Internet objects that have been tagged for measurement.



Note: **Sensor** does not degrade the performance of a properly configured web, application, or data collection server.

Adobe has designed **Sensor** to support a growing family of web and J2EE application servers including but not limited to AOLServer, Apache, iPlanet, JBoss, Microsoft IIS, Netscape Enterprise, Tomcat, and Weblogic running on common operating systems including, but not limited to, Microsoft Windows, AIX, Linux, and Solaris. **Sensor's** modular architecture enables Adobe to very quickly create new data acquisition logic for other applications as needed.

This chapter contains procedures for installing **Sensor** for web server/operating system combinations.

If you have not already installed the data workbench server that collects data from the **Sensor**, you must do so before performing the procedures in this chapter.

Microsoft IIS on Windows Server 2000 or later

Instructions about how to install and configure Sensor for Internet Information Services (IIS) 5.x or 6.x running under Microsoft Windows Server 2000 or later.

When using IIS 6.x, logging must be enabled for Sensor to function properly. If you have disabled logging to reduce disk I/O, you can enable logging without writing any data to the logs. To do so, enable logging, then clear all of the fields on the Advanced tab of the Properties for the W3C Extended Log File Format. If you need assistance, contact Adobe Consulting Services.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

To install and configure Sensor, you must perform the following high-level steps:

1. Install the Program Files

When running Sensor on Windows IIS, the program files and the disk queue file must reside in the same directory.

Therefore, before installing the program files, you must determine where you want to maintain the disk queue, because that is also where you must install the program files.

Use the following procedure to extract and install the program files for Sensor.

1. On your Windows machine, create a directory in which to install the Sensor program files. Keep in mind that your disk queue also resides in this directory, so be sure the device you choose has sufficient space to hold a queue of the size you need.

For example: C:\VisualSensor

2. Extract the contents of the installation file into the directory you just created. During this step, Sensor installs the following files:

File	Description
EventMessages.dll	Event Viewer messages.
qlog.dll	The collector module (an ISAPI filter).
TestExperiment.xls	An Excel spreadsheet file that architects can use to configure a controlled experiment. Sensor does not use this file.
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process.
TXLog.exe	The transmitter program.
txlogd.conf	The Sensor configuration file.



Note: The installation package contains a spreadsheet file called TestExperiment.xls. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

2. Edit the Configuration File

The txlogd.conf file contains the configuration parameters for Sensor.

You must edit the file to specify, among other things, the size of the disk queue, the address of the Insight Server, and the ID that will be attached to the data produced by this sensor. The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

1. Open the <SensorDirectory>/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.

For descriptions of txlogd.conf parameters, see Sensor Txlogd.conf File Parameters.

For examples of completed configuration files, see Sensor Sample Configuration Files.

2. Save and close the file.

3. Start the Transmitter and Create the Disk Queue

After you configure the txlogd.conf file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. From the Start menu in Windows, select Accessories > Command Prompt.
2. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

This command starts the transmitter, creates the disk queue, and registers Sensor as a Windows service.

3. To confirm that the transmitter is running correctly, click Start > Control Panel > Administrative Tools > Services.



Note: This command sequence might vary depending on which version of Windows you are using.

- a. In the service list, locate the entry for Sensor and confirm that its status is Started and its startup type is Automatic.
- b. Close the Services control panel.

4. To check whether transmitter experienced any errors during start-up, click Start > Control Panel > Administrative Tools > Event Viewer to open the Event Viewer.



Note: This command sequence might vary depending on which version of Windows you are using.

- a. In the left pane of the Event Viewer window, select the Applications log.
- b. In the right pane, look for events with “Adobe” in the Source column.
- c. If you find an error from “Adobe,” double-click the error to display the Event Properties window. This window provides detailed information about the error.

5. When you finish examining the Applications log, close the Event Viewer.

6. Verify that the transmitter has created the disk queue (Diskq2000.log) in the directory where you installed the Sensor program files and that it is the size that you specified in the QueueSize parameter in the txlogd.conf file.

If the queue has not been created correctly:

- a. Examine the txlogd.conf file and verify that the QueueSize parameter is set correctly.
- b. Check that the device on which you installed Sensor has sufficient space available to hold a file of the size specified in the QueueSize parameter.
- c. Using the Services control panel in Windows, stop the transmitter.
- d. Delete the queue file.
- e. Re-register Sensor as a Windows service: from the Start menu in Windows, select Accessories > Command Prompt. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

The transmitter is designed to run continuously. If you restart the machine, the transmitter restarts automatically. If you need to start and stop the transmitter manually, you can do so using the Services control panel in Windows.

Add the Collector to the Web Server

For IIS, the collector is an ISAPI filter that you add to your web server in IIS.

1. Open the IIS Manager using Start > Administrative Tools > Internet Information Services (IIS) Manager.
2. Expand the Local Computer and Web Sites nodes.
3. Right-click the website to which you want to add the collector and select Properties.
4. Select the ISAPI Filters tab and click Add.
5. In the Filter Name field, enter a display name for the filter. The suggested filter name is “Sensor.”
6. Click Browse, select the qlog.dll file (located in the directory where you installed Sensor), and click OK.
7. Click OK to add the filter.

After you add the filter, the collector is immediately operational and ready to collect data. An upward green arrow should appear in the Status column on the ISAPI Filters tab of the IIS Manager. You might not see the green arrow until traffic actually flows through the filter. In this case, you need to submit a request to the web server to confirm that the collector is operating correctly.

If the green arrow does not appear after traffic flows to the collector, complete the following steps:

1. Click Start > Administrative Tools > Event Viewer to check the Event Viewer for errors.



Note: This command sequence might vary depending on which version of Windows you are using.

2. In the left pane of the Event Viewer window, select the Application log.
3. In the right pane, look for events with “Adobe” in the Source column.
4. If you find an error, double-click the error to display the Event Properties window.

Capturing Additional Data

Web pages are often structured using ASP (Active Server Pages) programming language.

ASP is a Microsoft technology that runs within IIS. When a browser requests an ASP file, IIS passes the request to the ASP engine. The ASP engine reads the ASP file, line by line, and executes the scripts in the file. Finally, the ASP file is returned to the browser as plain HTML. ASP provides RESPOND or REQUEST objects which, in addition to other utilizations, allow the response or request of user queries or data submitted from HTML forms.

In certain cases, you may not want to append the values entered into forms to the URL that is displayed within the Address bar of a user’s browser or that is viewable within the HTML code itself. Simple server-side ASP Script lets you append form field names and their respective values to the log file without making them available within the user’s browser or embedding them into the HTML file. To capture the actual form values entered into particular forms within your website, a few lines of code must be added to append the form values to the log request.

Within the processing page of a form, include the following code to append the entered form values to the request data (in addition to writing the submitted form values to an external database or other location):

```
var sName= Request.Form("Name");
var sCity= Request.Form("City");
var sState= Request.Form("State");
var sZip= Request.Form("Zip");
```

```

Response.AppendToLog( "&v_1=" + sName );
Response.AppendToLog( "&v_2=" + sCity );
Response.AppendToLog( "&v_3=" + sState );
Response.AppendToLog( "&v_4=" + sZip );

```

This process would append the form values as defined to the request data for the Form Processing page. Within the log data, the appended values would be available as query strings of the Form Processing page as illustrated below. For example, v_1, v_2, v_3 and v_4 would now be query strings containing the data entered into the appropriate form fields. The syntax described in the previous example can be duplicated for any additional form fields and values that you want to capture:

http://www.myserver.com/path/to/formprocessingpage.asp?v_1=John+Smith&v_2=Los+Angeles&v_3=California&v_4=90210

If you want every form field and value to be captured and available for analysis, you can use the following syntax:

```

var formvalues = Response.Form;
Response.AppendToLog(formvalues);

```

This example would take all form fields present within the HTML along with their respective values and append them as query strings to the log entry for the Form Processing page. Please note that this would include any hidden fields present within the form.

The log data would be augmented as detailed in the following table:

Data Collected	Explanation	Example
v_1	Value associated with the NAME query string	v_1=John Smith
v_2	Value associated with the CITY query string	v_2=Los Angeles
v_3	Value associated with the STATE query string	v_3=California
v_4	Value associated with the ZIP query string	v_4=90210

Lotus Sametime on Windows Server 2000 or Later

Instructions about how to install and configure Sensor for Lotus Sametime for Windows 3.1 or later running under Microsoft Windows Server 2000 or later.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

To install and configure Sensor, you must perform the following high-level steps:

Install the Program Files

When running Sensor on Sametime, the program files and the disk queue file must reside in the same directory.

Therefore, before installing the program files, you must determine where you want to maintain the disk queue, because that is also where you must install the program files.

Use the following procedure to extract and install the program files for Sensor.

1. Stop the Lotus Domino Server and the Sametime Chat Logging service.
2. On your Windows machine, in the Lotus Domino directory, delete or backup the file named StChatLog.dll.
3. Extract the contents of the installation file into the Lotus Domino directory. During this step, Sensor installs the following files:

File	Description
EventMessages.dll	Event Viewer messages
stchatlog.dll	The collector module
TestExperiment.xls	An Excel spreadsheet file that architects can use to configure a controlled experiment Sensor does not use this file.
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process
TXLog.exe	The transmitter program
txlogd.conf	The Sensor configuration file



Note: The installation package contains a spreadsheet file called TestExperiment.xls. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Enable Logging on the Sametime Server

Steps that allow you to log on to the Sametime Server.

1. Use the Lotus Domino Administrator client to connect to the Lotus Domino server that is running Sametime.
2. In the Lotus Domino Administrator, click the Files tab, then double-click Sametime Configuration - stconfig.nsf to open the Sametime Configuration file.
3. In the Sametime Configuration file, open the Community Services form and double-click anywhere on the form to enter edit mode.
4. Set Chat Logging Flag to “strict” and Capture Service Type to “0x1000.”
5. Save and close the Community Services form, then close the Sametime Configuration file.
6. Restart the Sametime server.

Edit the Sensor Configuration File

The txlogd.conf file contains the configuration parameters for Sensor.

You must edit this file to specify, among other things, the size and location of the disk queue file, the address of the Insight Server, and the ID that will be attached to the event data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

- Open the <Sensor directory>/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.

-
- Save and close the file.

Start the Transmitter and Create the Disk Queue

After you configure the txlogd.conf file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. From the Start menu in Windows, select Accessories > Command Prompt.
2. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

This command starts the transmitter, creates the disk queue, and registers Sensor as a Windows service.

3. To confirm that the transmitter is running correctly, click Start > Control Panel > Administrative Tools > Services.



Note: This command sequence might vary depending on which version of Windows you are using.

- a. In the service list, locate the entry for Sensor and confirm that its status is Started and its startup type is Automatic.
 - b. Close the Services control panel.
4. To check whether transmitter experienced any errors during start-up, click Start > Control Panel > Administrative Tools > Event Viewer to open the Event Viewer.



Note: This command sequence might vary depending on which version of Windows you are using.

- a. In the left pane of the Event Viewer window, select the Applications log.
 - b. In the right pane, look for events with “Adobe” in the Source column.
 - c. If you find an error from “Adobe,” double-click the error to display the Event Properties window. This window provides detailed information about the error.
5. When you finish examining the Applications log, close the Event Viewer.
 6. Verify that the transmitter has created the disk queue (Diskq2000.log) in the directory where you installed the Sensor program files and that it is the size that you specified in the QueueSize parameter in the txlogd.conf file.

If the queue has not been created correctly:

1. Examine the txlogd.conf file and verify that the QueueSize parameter is set correctly.
2. Check that the device on which you installed Sensor has sufficient space available to hold a file of the size specified in the QueueSize parameter.
3. Using the Services control panel in Windows, stop the transmitter.
4. Delete the queue file.
5. Re-register Sensor as a Windows service: from the Start menu in Windows, select Accessories > Command Prompt. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

The transmitter is designed to run continuously. If you restart the machine, the transmitter restarts automatically. If you need to start and stop the transmitter manually, you can do so using the Services control panel in Windows.

7. Restart the Lotus Domino Server and the Sametime Chat Logging service.

Lotus Domino Server on Windows Server 2000 or Later

Instructions about how to install and configure Sensor for Lotus Domino Server 6 for Windows 3.1 or later running under Microsoft Windows Server 2000 or later.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

To install and configure Sensor, you must perform the following steps:

Install the Program Files

1. On your Lotus Domino machine, create a directory to install the Sensor program files. Keep in mind that your disk queue also resides in this directory, so be sure the device you choose has sufficient space to hold a queue of the size you need.

C:\VisualSensor

2. Extract the contents of the installation file into the Lotus Domino directory. During this step, Sensor installs the following files:

File	Description
EventMessages.dll	Event Viewer messages
stchatlog.dll	The collector module
TestExperiment.xls	An Excel spreadsheet file that architects can use to configure a controlled experiment Sensor does not use this file.
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process
TXLog.exe	The transmitter program
txlogd.conf	The Sensor configuration file

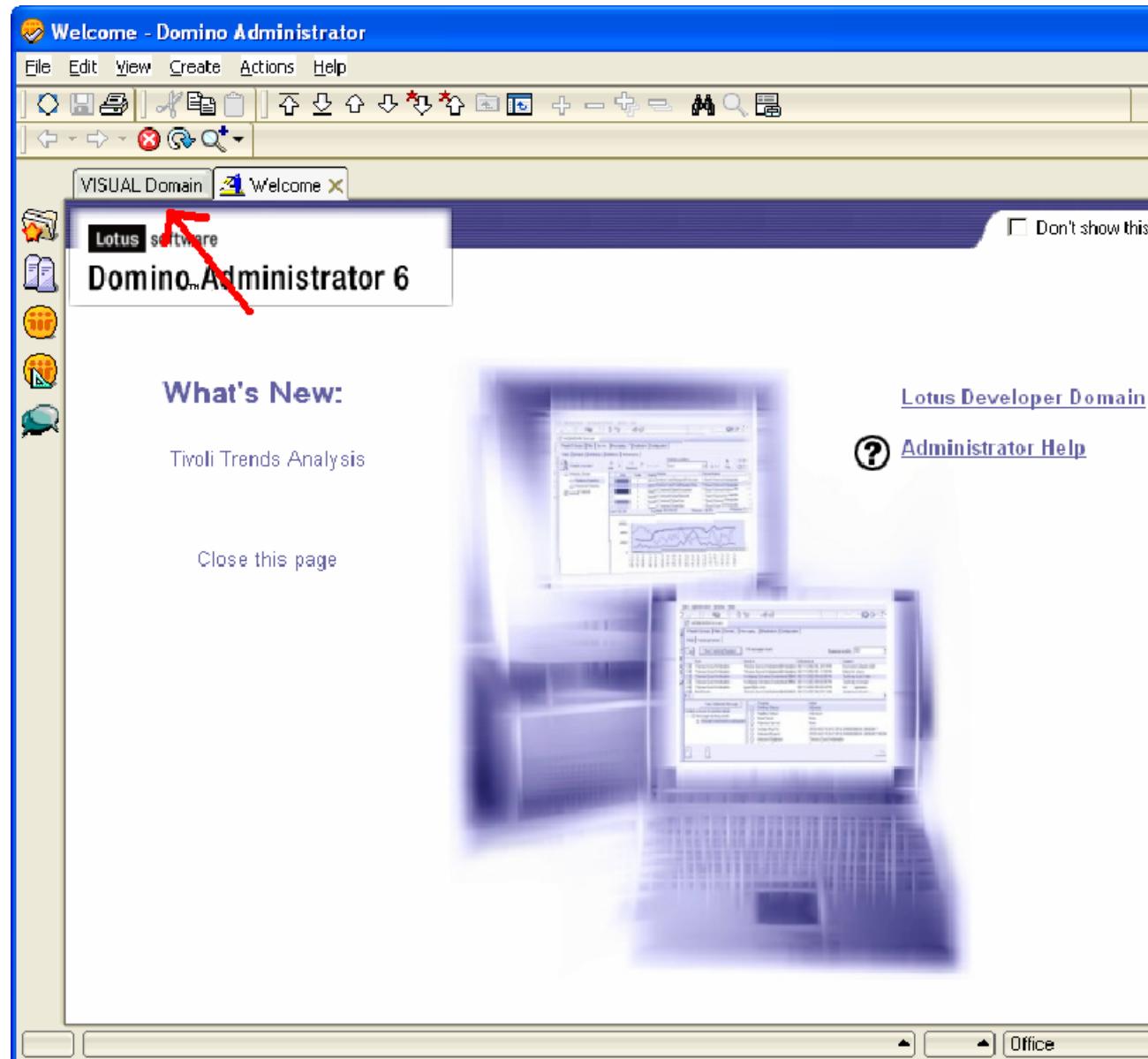


Note: The installation package contains a spreadsheet file called TestExperiment.xls. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

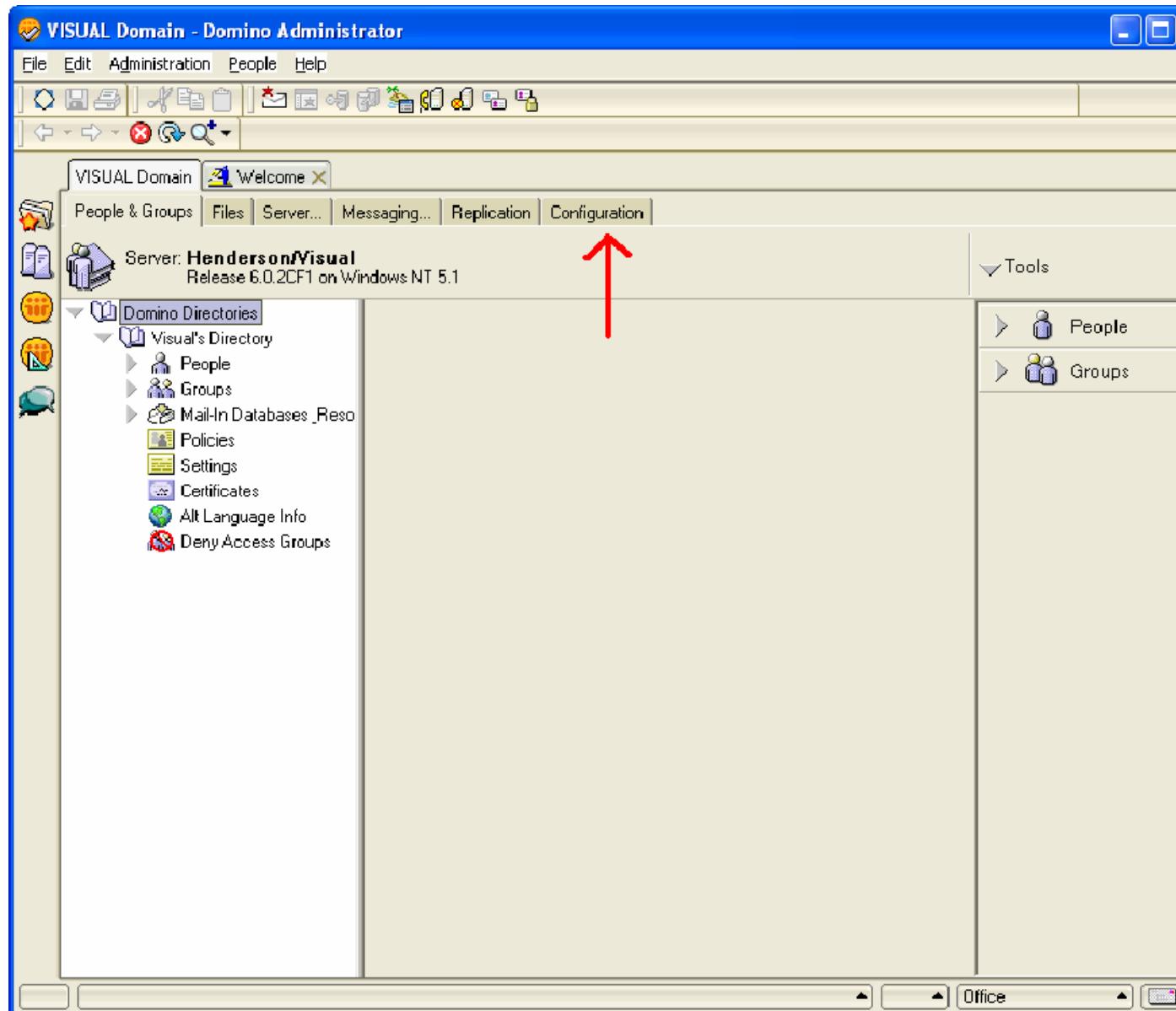
Configure the Lotus Domino Server

Steps to configure the Lotus Domino Server.

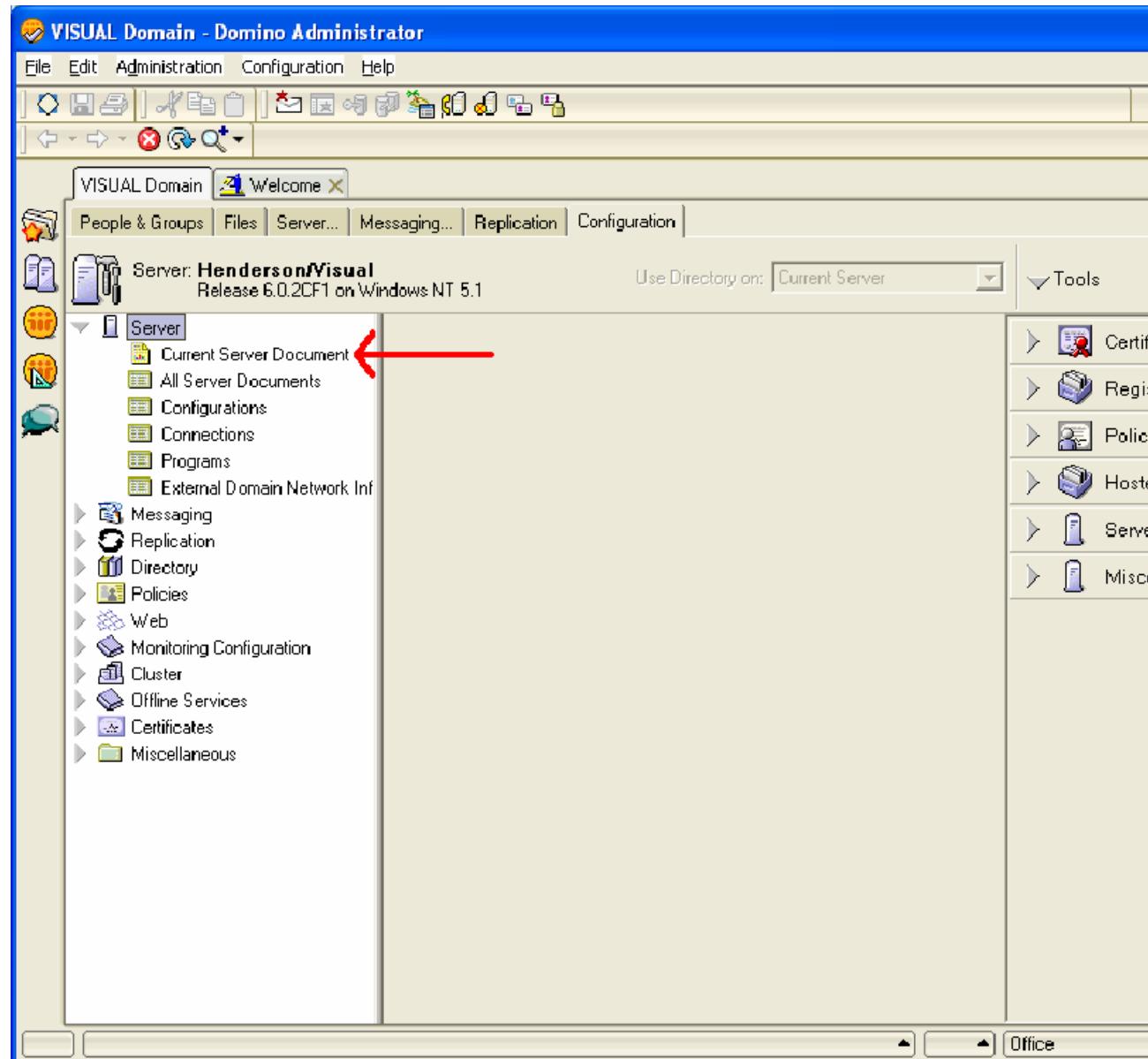
1. Log in to the Lotus Domino Administrator and click **Domain**.



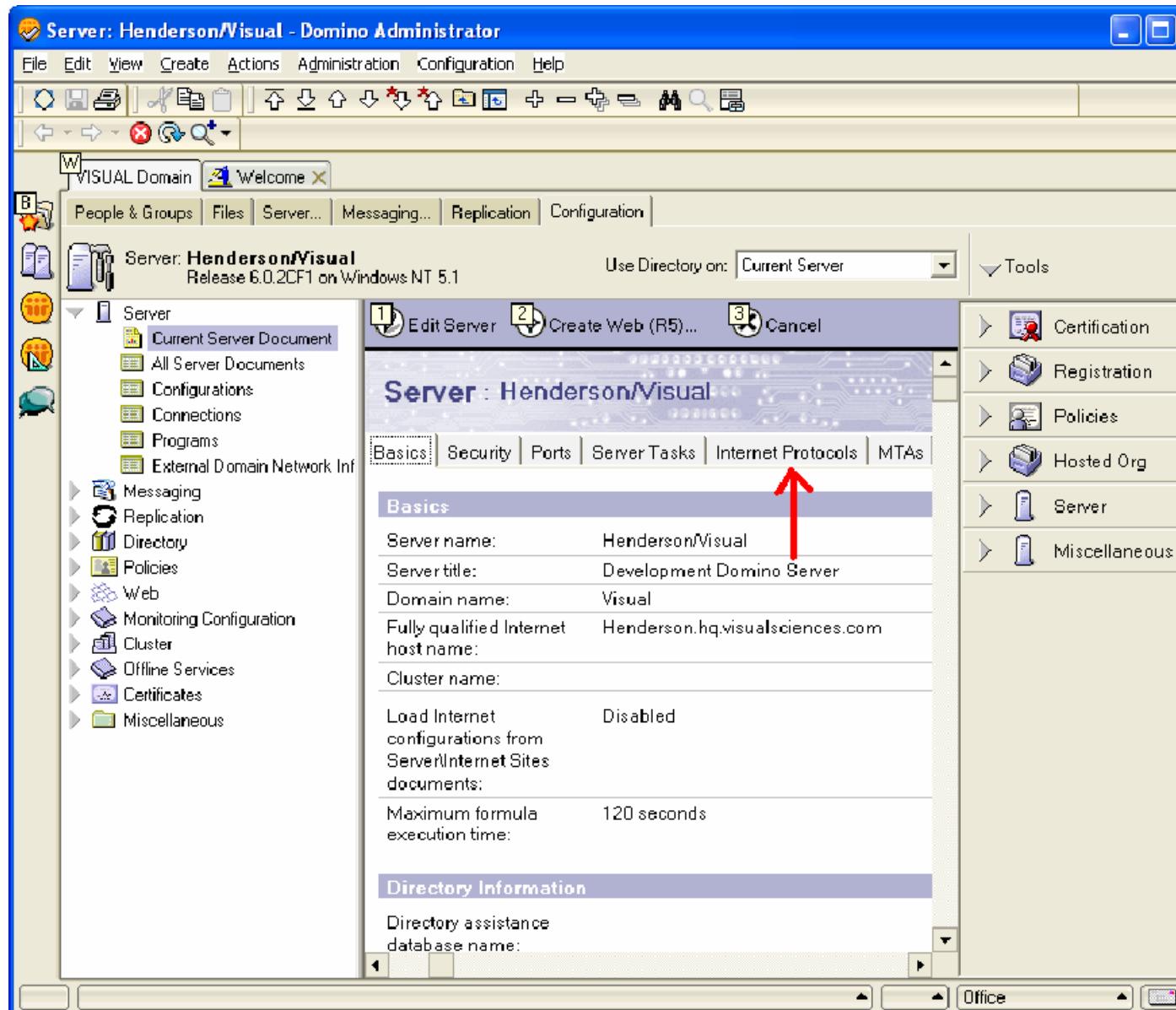
2. In the Lotus Domino Administrator, click **Configuration**.



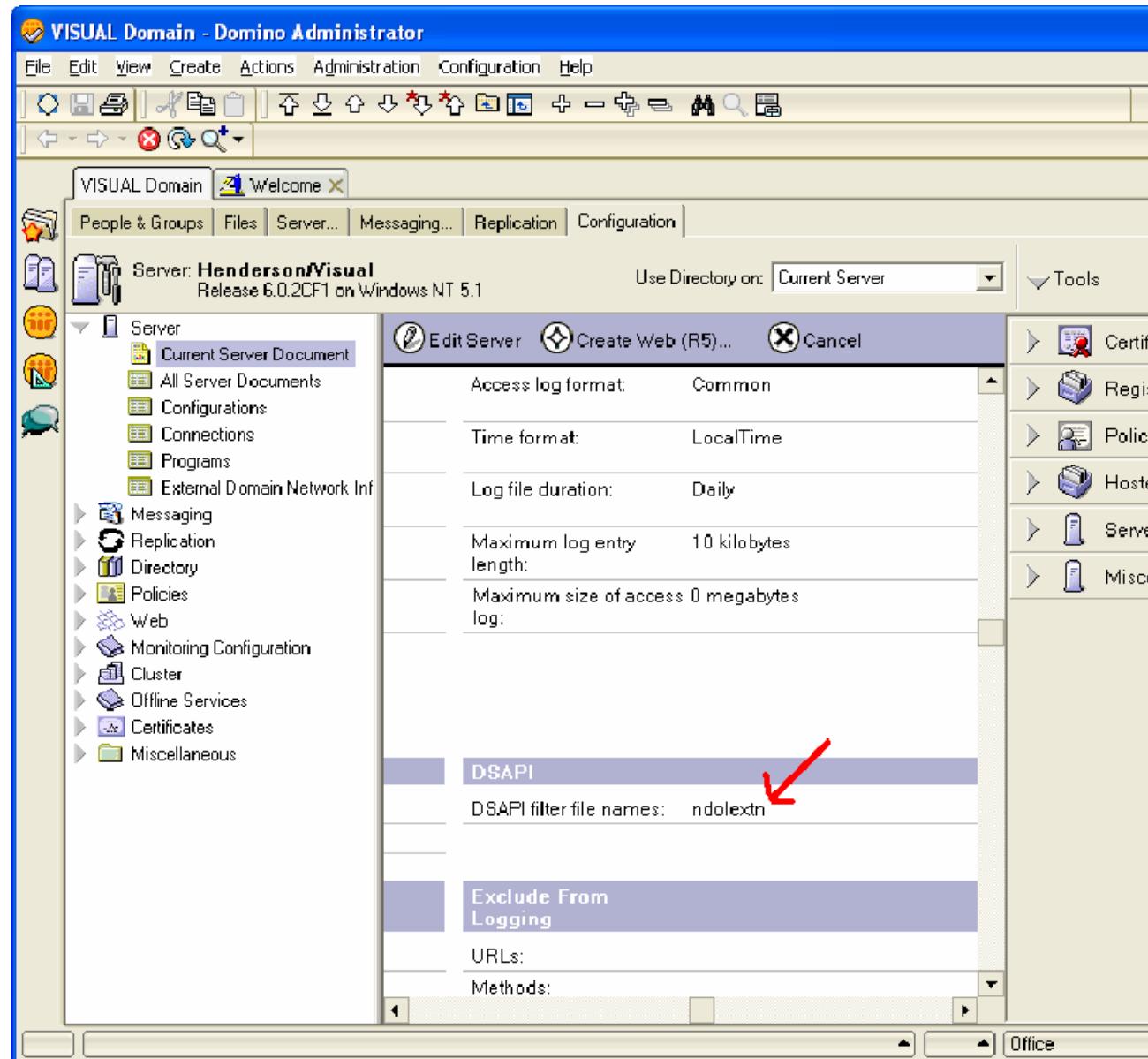
3. Expand the Server node and click **Current Server Document**.



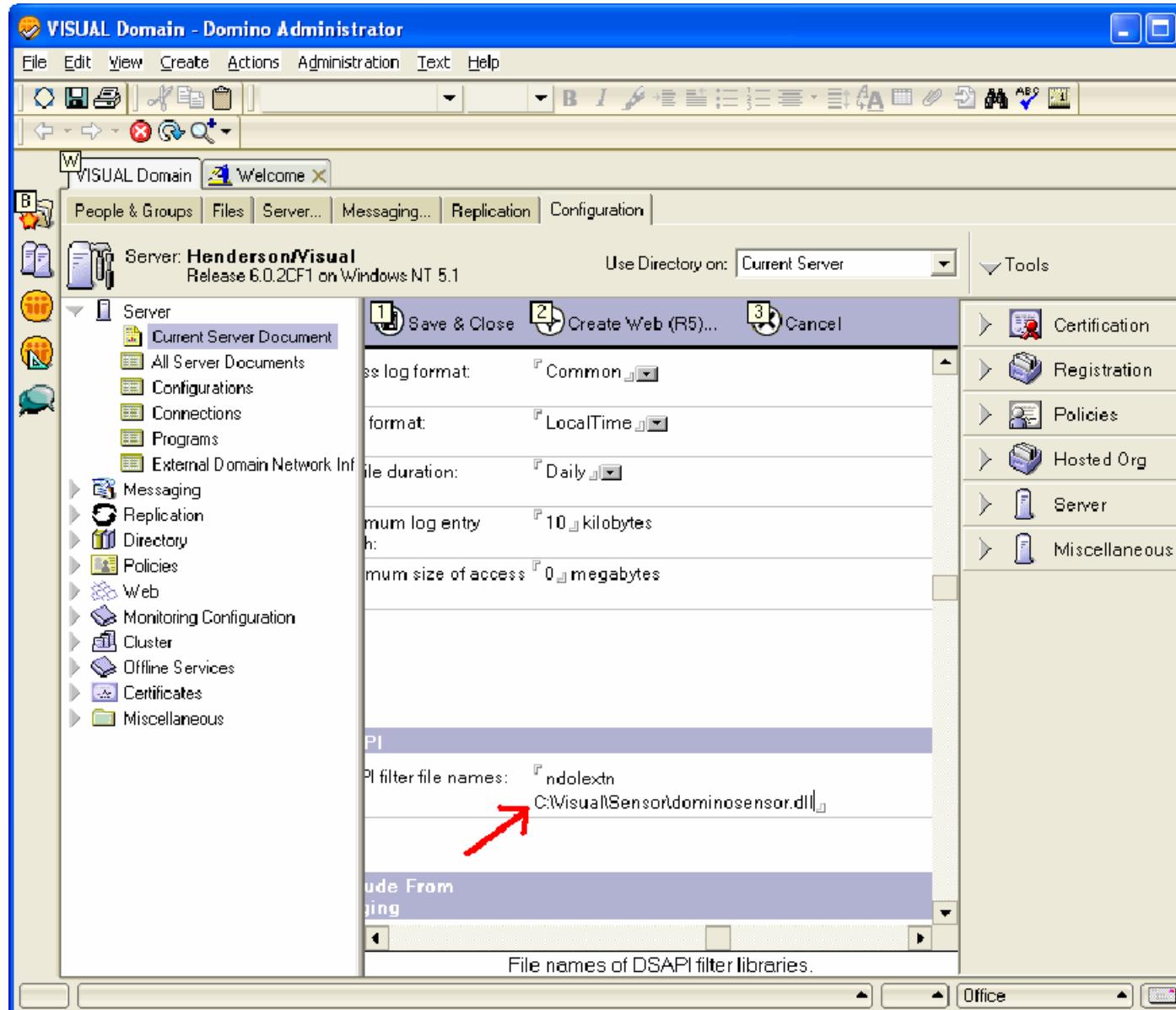
4. Click **Current Server Document**, then click **Internet Protocols**.



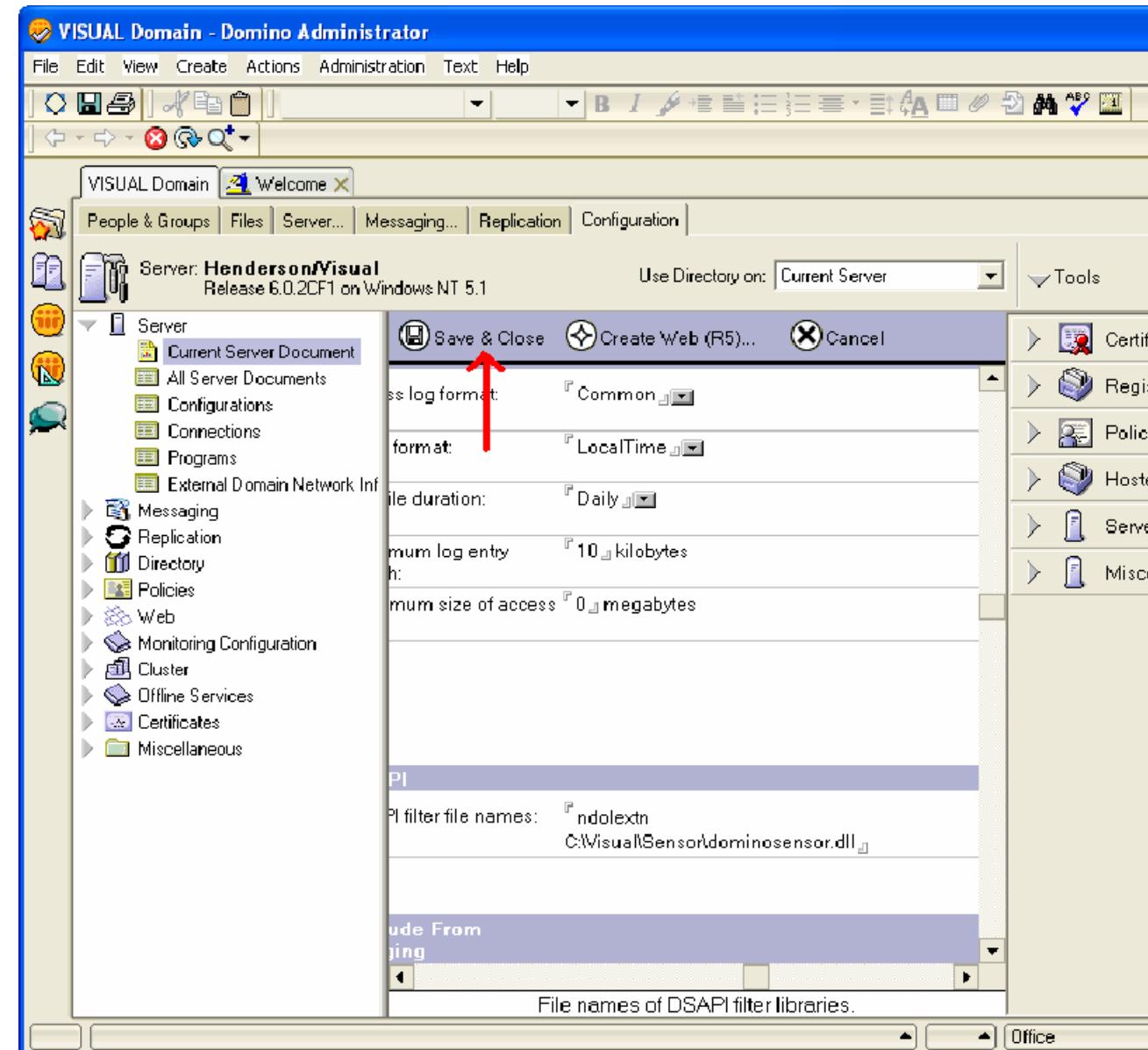
5. On the HTTP tab, under the DSAPI section, double-click after the word ndolextn.



6. Press **Enter** and type in the path to the `dominosensor.dll` file.



7. Click Save & Close.



Edit the Sensor Configuration File

The txlogd.conf file contains the configuration parameters for Sensor.

You must edit this file to specify, among other things, the size and location of the disk queue file, the address of the Insight Server, and the ID that will be attached to the event data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

-
- Open the <Sensor directory>/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
 - Save and close the file.

Start the Transmitter and Create the Disk Queue

After you configure the txlogd.conf file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. From the Start menu in Windows, select **Accessories > Command Prompt**.
2. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

This command starts the transmitter, creates the disk queue, and registers Sensor as a Windows service.

3. To confirm that the transmitter is running correctly, click **Start > Control Panel > Administrative Tools > Services**.



Note: This command sequence might vary depending on which version of Windows you are using.

- a. In the service list, locate the entry for Sensor and confirm that its status is Started and its startup type is Automatic.
 - b. Close the Services control panel.
4. To check whether transmitter experienced any errors during start-up, click Start > Control Panel > Administrative Tools > Event Viewer to open the Event Viewer.



Note: This command sequence might vary depending on which version of Windows you are using.

- a. In the left pane of the Event Viewer window, select the Applications log.
 - b. In the right pane, look for events with “Adobe” in the Source column.
 - c. If you find an error from “Adobe,” double-click the error to display the Event Properties window. This window provides detailed information about the error.
5. When you finish examining the Applications log, close the Event Viewer.
 6. Verify that the transmitter has created the disk queue (`Diskq2000.log`) in the directory where you installed the Sensor program files and that it is the size that you specified in the QueueSize parameter in the `txlogd.conf` file.

If the queue has not been created correctly:

1. Examine the `txlogd.conf` file and verify that the QueueSize parameter is set correctly.
2. Check that the device on which you installed Sensor has sufficient space available to hold a file of the size specified in the QueueSize parameter.
3. Using the Services control panel in Windows, stop the transmitter.
4. Delete the queue file.
5. Re-register Sensor as a Windows service: from the Start menu in Windows, select Accessories > Command Prompt. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

The transmitter is designed to run continuously. If you restart the machine, the transmitter restarts automatically. If you need to start and stop the transmitter manually, you can do so using the Services control panel in Windows.

IBM HTTP Server on AIX 5.1 or Later

Instructions about how to install and configure IBM HTTP server on IBM AIX 5.1 or later running under Microsoft Windows Server 2000 or later.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

To install and configure Sensor, you must perform the following high-level steps:

Install the Program Files

On your IBM AIX server, create a directory in which to install the Sensor program files. Keep in mind that your disk queue also resides in this directory, so be sure the device you choose has sufficient space to hold a queue of the size you need.

1. Log on as the root user or as a user with root authority.
2. Decompress and unpack the installation file using the following command:

```
tar -zxf installationFilename
```
3. Copy the unpacked program files to the directories identified in the following table:

File	Description	Target Directory
mod_visual_sciences.so	The collector load module.	<i>IBMHttpServer/modules</i>
txlogd	The transmitter program. --OR-- <i>/usr/local/sbin</i>	<i>/usr/local/bin</i>
txlogd.conf	The Sensor configuration file.	<i>/etc</i>
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process	<i>/usr/local/visual_sciences</i>



Note: The installation package contains a spreadsheet file called TestExperiment.xls. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Permissions on the Program Files



Note: Incorrect permissions on the program files cause the majority of problems encountered when installing Sensor. Please make sure that you set the permissions exactly as stated in this section.

By default, the program files in the tar file have the following permissions. Depending on how your system is configured, these settings might be altered (unmasked) when you extract the files. To reset the permissions to the recommended default settings, use the chmod commands below. Check that the directories into which you have installed the files permit at least this level of access.

File	Default Permissions	chmod command
libvisual_sciences.so	rwx --x --x	chmod 711
J2EECollector.jar	rw- rw- r--	chmod 664
txlogd	rwx --x --x	chmod 711
txlogd.conf	rw- rw- r--	chmod 664
trust_ca_cert.pem	rw- rw- r--	chmod 664

Edit the Sensor Configuration File

The txlogd.conf file contains the configuration parameters for Sensor.

You must edit this file to specify, among other things, the size and location of the disk queue file, the address of the Insight Server, and the ID that will be attached to the event data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

- Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
- Save and close the file.

To edit the Sensor configuration file

1. Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
2. Save and close the file.

Start the Transmitter and Create the Disk Queue

After you configure the txlogd.conf file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. If the directory in which the disk queue resides does not already exist, create it. Make sure that the directory provides both the collector module and the transmitter program with read/write access to the file.

For more information about the permissions required by the disk queue files, see Sensor UNIX File Permissions.

2. On the computer where Sensor is installed, execute the following command to start the transmitter:

```
/usr/local/bin/txlogd -ic -f /etc/txlogd.conf
```

-
- The “i” option in this command starts the transmitter in “interactive mode.” This mode displays transmitter messages on the screen and also enables you to interact with the transmitter using keyboard commands.
 - The “c” option directs the transmitter to create the disk queue.
 - The “f” option specifies the location of the configuration file.

For additional information about the options you can use when starting the transmitter, see Sensor Transmitter Command-Line Options.

3. Verify that the transmitter has created the disk queue in the location specified in the QueueFile parameter and of the size specified in the QueueSize parameter.
4. If the queue has not been created correctly, type Ctrl+C to terminate the transmitter, then do the following:
 - a. Examine the txlogd.conf file and verify that the QueueFile and QueueSize parameters are set correctly.
 - b. Check that the device to which the disk queue is assigned is operational and has sufficient space available to hold a file of the size specified in the QueueSize parameter.
 - c. Make any necessary corrections and repeat this procedure.

Add the Collector to the Web Application

For WebSphere servers, the collector operates as a filter in the servlet container.

To add the collector to your web app, open the web.xml file for the web server whose events Sensor captures.

If Sensor is capturing data for multiple web servers on the server computer, you must perform the following procedure for each web server.

1. Using a text editor, open the httpd.conf file for the web server whose events Sensor captures.
2. Add the following <filter> and <filter-mapping> elements to the descriptor file. If you did not install txlogd.conf in the /etc directory, you need to enter the correct path to this file in the <param-value> element.

```
<filter>
  <filter-name>VSCollectorFilter</filter-name>
  <description></description>
  <filter-class>
    com.visualsciences.collector.VSCollectorFilter
  </filter-class>
  <init-param>
    <param-name>configPath</param-name>
    <param-value>C:/VisualSensor/txlogd.conf</param-value>
  <description></description>
  </init-param>
</filter>

<filter-mapping>
  <filter-name>VSCollectorFilter</filter-name>
  <url-pattern>/*</url-pattern>
</filter-mapping>
```



Note: These lines are case sensitive. Type them exactly as they appear above.

3. Restart the web application. The collector is loaded with the application and will begin collecting event data and writing it to the disk queue.

Declare the Location of the Collector and Shared Object Files

Edit the Websphere startup script to declare the location of the J2EECollector.jar and libvisual_science.so files.

1. Open the setupCmdLine.sh file in the Websphere /bin directory.

-
2. After the line that defines the \$WAS_CLASSPATH variable, add the following line:

```
WAS_CLASSPATH="$WAS_CLASSPATH":$WAS_HOME/lib/J2EECollector.jar
```

3. After the case block that defines the \$WAS_LIBPATH variable, add the following line:

```
WAS_LIBPATH="$WAS_LIBPATH":/usr/local/visual_sciences
```

4. Save the setupCmdLine.sh file.

Test the Sensor

Verify that the collector is collecting event data and the transmitter is transmitting it to the target Insight Server.



Note: To verify that the transmitter can successfully send event data to the Insight Server, make sure that the target Insight Server is installed and running before you begin the following test.

1. If the transmitter is not already running, restart it using the following command:

```
/usr/local/bin/txlogd -i -f /etc/txlogd.conf
```

2. Open a browser (on any machine) and request a page from the web server on which Sensor is running (be sure to select a page that Sensor is monitoring).
3. After you issue the request, check the transmitter's console for messages indicating that it is sending event data to the target Insight Server.

4. If Sensor is not transmitting data successfully, verify that:

- The target Insight Server is running.
- The ServerAddress and ServerPort parameters are set correctly in txlogd.conf. If you specified ServerAddress using a server name, try using its numeric IP address instead.
- The value of the CertName parameter matches the common name that appears on the digital certificate of the target Insight Server exactly.

Add the Transmitter to System Startup Script

Information about automatically loading the transmitter to your system startup script.

To ensure that the transmitter loads automatically when the web server machine is restarted, add the following command (which launches the transmitter) to your system startup script:

```
/usr/local/bin/txlogd -f /etc/txlogd.conf
```

This command starts the transmitter as a daemon. Operating and error messages that the transmitter generates are written to syslog.

Capturing Additional Data

Sensors for all platforms can collect any of the data available in the HTTP request and response headers.

The Sensors for the J2EE Platform provide a mechanism for collecting data that is not available on other platforms. The collector for the J2EE platform (J2EE collector) sits on the application layer, which enables it to collect sensitive data that is available only to the application and should not be exposed through page tagging or in the headers.



Note: While page tags and header modification can hide the data, it is still available to those who examine the source code of a page or look at the headers using browser plug-in tools.

For example, the J2EE collector can be used to capture cost per click (CPC) data for links displayed on a page, sensitive partner information on a page, and many other data points. The J2EE environment makes it easy for you to modify your WEBAPP to capture this custom data using our collector class.

When a Sensor for the J2EE Platform receives a request, it invokes a collector class that imports the appendToLog function. The appendToLog function appends to the initial request the query string parameters specified in the appendToLog function. This results in the URI of the initial request containing additional query string name-value pairs that correspond to the names and values of the data that is being captured. For example, CPC=20 would be appended to the initial request when the value of a particular ad placement or click-through link is 20 cents. Insight Server processes these values into the dataset for analysis. One additional benefit to this collection methodology is that it allows the collection of additional data without creating extra log entries, as might be created using page tagging methodologies.

For more information about processing, see the Dataset Configuration Guide.

To capture additional data from a page:

1. Add the following code to the top of the .jsp page from which you want to capture data:

```
<%@ page import="com.visualsciences.collector.VSCollector" %>
```

2. Use the appendToLog() method of the collector object to append the desired name-value pairs to the requested .jsp page's query string. The following example appends "A=1" and "B=2" to the requested .jsp page's query string for the /index.jsp page:

```
<html>
<body>
    <h1>Hello World</h1>
    <%
        VSCollector collector = new VSCollector(request, response);
        collector.appendToLog("A", "1");
        collector.appendToLog("B", "2");
    %>
</body>
</html>
```

The resulting request URI is /index.jsp?A=1&B=2.

3. Repeat this procedure for each .jsp page from which you want to capture additional data.

Apache Server 1.3.x on Linux, Sun Solaris, FreeBSD, or Mac OS X

Detailed instructions for installing and configuring Sensor for an Apache Server 1.3.x on RedHat Linux 7.x or later, SUSE Linux 9.x or later, Sun Solaris SPARC 2.6 or later, Sun Solaris x86 9 or later, FreeBSD 4 or later, or Mac OS X PowerPC.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

To install and configure Sensor, you must perform the following high-level steps:

Install the Program Files

Instructions to extract and install the program files for Sensor to the server machine.

1. Log on as the root user or as a user with root authority.
2. Decompress and unpack the installation file using the following command:

- On Linux:

```
tar -zxf installationfilename.tar.gz
```

- On Solaris:

```
unzip -d installationFilename.tar.gz
tar -xf installationFilename.tar
```

3. Copy the unpacked program files to the directories identified in the following table:

File	Description	Target Directory
mod_visual_sciences.so	The collector load module	apachePath/libexec
txlogd	The transmitter program	/usr/local/bin --OR-- /usr/local/sbin
txlogd.conf	The Sensor configuration file	/etc
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process	/usr/local/visual_sciences



Note: The installation package contains a spreadsheet file called TestExperiment.xls. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you may choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Permissions on the Program Files

Incorrect permissions on the program files cause the majority of problems encountered when installing Sensor. Please make sure that you set the permissions exactly as stated in this section.

By default, the program files in the tar file have the following permissions. Depending on how your system is configured, these settings might be altered (unmasked) when you extract the files. To reset the permissions to the recommended default settings, use the chmod commands below. Check that the directories into which you have installed the files permit at least this level of access.

File	Default Permissions	chmod command
mod_visual_sciences.so	rwx r-x r-x	chmod 755
txlogd	rwx --x --x	chmod 711
txlogd.conf	rw- rw- r--	chmod 664
trust_ca_cert.pem	rw- rw- r--	chmod 664

Edit the Sensor configuration file

The txlogd.conf file contains the configuration parameters for Sensor.

You must edit the file to specify, among other things, the size of the disk queue, the address of the Insight Server, and the ID that will be attached to the data produced by this sensor.

The configuration file contains required parameters and optional parameters.

-
- Required parameters are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
 - Optional parameters are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

1. Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
2. Save and close the file.

Start the Transmitter and Create the Disk Queue

Instructions to create the disk queue after you configure the txlogd.conf file.

1. If the directory in which the disk queue resides does not already exist, create it. Make sure that the directory provides both the collector module and the transmitter program with read/write access to the file.
2. On the computer where Sensor is installed, execute the following command to start the transmitter:

```
/usr/local/bin/txlogd -ic -f /etc/txlogd.conf
```

- The “i” option in this command starts the transmitter in interactive mode. This mode displays transmitter messages on the screen and also enables you to interact with the transmitter using keyboard commands.
 - The “c” option directs the transmitter to create the disk queue.
 - The “f” option specifies the location of the configuration file.
3. Verify that the transmitter has created the disk queue in the location specified in the QueueFile parameter and of the size specified in the QueueSize parameter.
 4. If the queue has not been created correctly, type Ctrl+C to terminate the transmitter, then do the following:
 - a. Examine the txlogd.conf file and verify that the QueueFile and QueueSize parameters are set correctly.
 - b. Check that the device to which the disk queue is assigned is operational and has sufficient space available to hold a file of the size specified in the QueueSize parameter.
 - c. Make any necessary corrections and repeat this procedure.

Add the Collector to the Web Server

For Apache servers, the collector is a dynamic shared object that you load into your web server process.

To add the collector to your web server, you must edit the httpd.conf file as described below and restart your web server.

If Sensor is capturing data for multiple web servers on the server computer, you must perform the following procedure for each web server.

1. Using a text editor, open the httpd.conf file for the web server whose events Sensor captures.
2. Add the following lines to the end of the file:

```
LoadModule visual_sciences_module libexec/mod_visual_sciences.so
VisualSciencesConfig /etc/txlogd.conf
AddModule mod_visual_sciences.c
```



Note: These lines are case sensitive. Type them exactly as they appear above.

3. Restart the web server. The collector is loaded with the web server and will begin collecting event data and writing it to the disk queue.

Test the Sensor

Start the transmitter and verify that it can successfully connect to the Insight Server and transmit event data to it.



Note: To verify that the transmitter can successfully send event data to the Insight Server, make sure that the target Insight Server is installed and running before you begin the following test.

1. If the transmitter is not already running, restart it using the following command:

```
/usr/local/bin/txlogd -i -f /etc/txlogd.conf
```
2. Open a browser (on any machine) and request a page from the web server on which Sensor is running (be sure to select a page that Sensor is monitoring).
3. After you issue the request, check the transmitter's console for messages indicating that it is sending event data to the target Insight Server.
4. If Sensor is not transmitting data successfully, verify that:
 - The target Insight Server is running.
 - The `ServerAddress` and `ServerPort` parameters are set correctly in `txlogd.conf`.
 - If you specified `ServerAddress` using a server name, try using its numeric IP address instead. The value of the `CertName` parameter matches the common name that appears on the digital certificate of the target Insight Server exactly.

Add the Transmitter to Your System Startup Script

Information to ensure that the transmitter loads automatically when the web server machine is restarted.

Add the following command (which launches the transmitter) to your system startup script.

```
/usr/local/bin/txlogd -f /etc/txlogd.conf
```

This command starts the transmitter as a daemon. Operating and error messages that the transmitter generates are written to syslog.



Note: Some Solaris users might encounter an “unable to acquire mutex” error. For Sensor to function properly on these systems, the following line needs to either be added to or edited in the file `/etc/system`:
`semsys:seminfo_semmnu=1024`

The default Solaris setting is 60. Based on tests conducted with Sensor, which uses three semaphores for each instance, Adobe recommends that you use 1024 as your setting. This number is high enough for Sensor to function along with any other applications on the server that may require semaphores, but does not affect performance. To support this recommendation, please note that Adrian Cockcroft stated the following in his book Sun Performance and Tuning (Prentice Hall, October 1994): “Databases tend to use lots of shared memory and semaphore settings. These do not affect performance; as long as they are big enough, the programs will run.”

Apache Server 2.0.40, 2.0.42 or later, and Apache Server 2.2 or 2.4 on Linux, Solaris, or FreeBSD

Instructions about how to install and configure Apache Server 2.0.40, 2.0.42 or later, Apache Server 2.2, or Apache Server 2.4 on Linux, Sun Solaris, or FreeBSD.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

To install and configure Sensor, you must perform the following high-level steps:

The following Apache Servers are supported:

- Apache Server 2.0.40 running under RedHat Linux 7.x or later, or Sun Solaris SPARC 2.6 or later.
- Apache Server 2.0.40, 2.0.42 or Later, Apache Server 2.2, or Apache Server 2.4 on Linux, Sun Solaris, or FreeBSD
- Apache Server 2.0.42 or later running under RedHat Linux 7.x or later, Sun Solaris SPARC 2.6 or later, SUSE Linux 9.x or later, or FreeBSD 5.3.
- Apache Server 2.0.42 or later running under 64-bit versions of RedHat Linux ES 4 and ES 5.
- Apache Server 2.2 running under RedHat Linux 7.x or later or Sun Solaris SPARC 2.6 or later.
- Apache Server 2.4 running under RedHat Linux 7.x or later, or Sun Solaris x86_64, or FreeBSD



Note: Although the instructions to install Sensors on web servers running Apache Server versions 2.0.40, 2.0.42 or later (32-bit and 64-bit), or 2.2 are the same (except where noted in the following procedures), the installation files for each version differ. Before installing the Sensor, ensure that you have received the correct installation files for the Apache Server and operating system versions that you are running.

Install the Program Files

Procedure to extract and install the program files for Sensor.

1. Log on as the root user or as a user with root authority.
2. Decompress and unpack the installation file using the following command:

- On Linux:

```
tar -zxf installationFilename  
unzip -d installationFilename.tar.gz  
tar -xf installationFilename.tar
```

- On Solaris:

3. Copy the unpacked program files to the directories identified in the following table:

File	Description	Target Directory
mod_visual_sciences.so	The collector load module.	<i>IBMHttpServer/modules</i>
txlogd	The transmitter program.	<i>/usr/local/bin</i> --OR-- <i>/usr/local/sbin</i>
txlogd.conf	The Sensor configuration file.	<i>/etc</i>
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process	<i>/usr/local/visual_sciences</i>



Note: The installation package contains a spreadsheet file called TestExperiment.xls. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Permissions on the Program Files



Note: Incorrect permissions on the program files cause the majority of problems encountered when installing Sensor. Please make sure that you set the permissions exactly as stated in this section.

By default, the program files in the tar file have the following permissions. Depending on how your system is configured, these settings might be altered (unmasked) when you extract the files. To reset the permissions to the recommended default settings, use the chmod commands below. Check that the directories into which you have installed the files permit at least this level of access.

File	Default Permissions	chmod command
mod_visual_sciences.so	rwx r-x r-x	chmod 775
txlogd	rwx --x --x	chmod 711
txlogd.conf	rw- r-- r--	chmod 664
trust_ca_cert.pem	rw- r-- r--	chmod 664

Enable Logging on the Sametime Server

Steps that allow you to log on to the Sametime Server.

1. Use the Lotus Domino Administrator client to connect to the Lotus Domino server that is running Sametime.
2. In the Lotus Domino Administrator, click the Files tab, then double-click Sametime Configuration - stconfig.nsf to open the Sametime Configuration file.
3. In the Sametime Configuration file, open the Community Services form and double-click anywhere on the form to enter edit mode.
4. Set Chat Logging Flag to “strict” and Capture Service Type to “0x1000.”
5. Save and close the Community Services form, then close the Sametime Configuration file.
6. Restart the Sametime server.

Edit the Sensor Configuration File

The txlogd.conf file contains the configuration parameters for Sensor.

You must edit this file to specify, among other things, the size and location of the disk queue file, the address of the Insight Server, and the ID that will be attached to the event data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

- Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
- Save and close the file.

To edit the Sensor configuration file

1. Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
2. Save and close the file.

Start the Transmitter and Create the Disk Queue

After you configure the txlogd.conf file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. If the directory in which the disk queue resides does not already exist, create it. Make sure that the directory provides both the collector module and the transmitter program with read/write access to the file.

For more information about the permissions required by the disk queue files, see Sensor UNIX File Permissions.

2. On the computer where Sensor is installed, execute the following command to start the transmitter:

```
/usr/local/bin/txlogd -ic -f /etc/txlogd.conf
```

- The “i” option in this command starts the transmitter in “interactive mode.” This mode displays transmitter messages on the screen and also enables you to interact with the transmitter using keyboard commands.
- The “c” option directs the transmitter to create the disk queue.
- The “f” option specifies the location of the configuration file.

For additional information about the options you can use when starting the transmitter, see Sensor Transmitter Command-Line Options.

3. Verify that the transmitter has created the disk queue in the location specified in the QueueFile parameter and of the size specified in the QueueSize parameter.
4. If the queue has not been created correctly, type Ctrl+C to terminate the transmitter, then do the following:
 - a. Examine the txlogd.conf file and verify that the QueueFile and QueueSize parameters are set correctly.
 - b. Check that the device to which the disk queue is assigned is operational and has sufficient space available to hold a file of the size specified in the QueueSize parameter.
 - c. Make any necessary corrections and repeat this procedure.

Add the Collector to the Web Server

For IBM HTTP servers, the collector is a dynamic shared object that you load into your web server process.

To add the collector to your web server, you must edit the httpd.conf file as described below and restart your web server.

If Sensor is capturing data for multiple web servers on the server computer, you must perform the following procedure for each web server.

1. Using a text editor, open the httpd.conf file for the web server whose events Sensor captures.
2. Add the following two lines to the end of the file:

```
LoadModule visual_sciences_module modules/mod_visual_sciences.so  
VisualSciencesConfig /etc/txlogd.conf
```



Note: These lines are case sensitive. Type them exactly as they appear above.

3. Restart the web server process (you do not have to reboot the entire server computer, simply restart the web server process). The collector is loaded with the web server and begins collecting event data and writing it to the disk queue.

Test the Sensor

Verify that the collector is collecting event data and the transmitter is transmitting it to the target Insight Server.



Note: To verify that the transmitter can successfully send event data to the Insight Server, make sure that the target Insight Server is installed and running before you begin the following test.

1. If the transmitter is not already running, restart it using the following command:

```
/usr/local/bin/txlogd -i -f /etc/txlogd.conf
```

2. Open a browser (on any machine) and request a page from the web server on which Sensor is running (be sure to select a page that Sensor is monitoring).
3. After you issue the request, check the transmitter's console for messages indicating that it is sending event data to the target Insight Server.
4. If Sensor is not transmitting data successfully, verify that:
 - The target Insight Server is running.
 - The ServerAddress and ServerPort parameters are set correctly in txlogd.conf. If you specified ServerAddress using a server name, try using its numeric IP address instead.
 - The value of the CertName parameter matches the common name that appears on the digital certificate of the target Insight Server exactly.

Add the Transmitter to System Startup Script

Information about automatically loading the transmitter to your system startup script.

To ensure that the transmitter loads automatically when the web server machine is restarted, add the following command (which launches the transmitter) to your system startup script:

```
/usr/local/bin/txlogd -f /etc/txlogd.conf
```

This command starts the transmitter as a daemon. Operating and error messages that the transmitter generates are written to syslog.

Apache Server 1.3, 2, 2.2, or 2.4 on Windows Server 2000 or Later

Instructions for installing and configuring Sensor for Apache Server 1.3, Apache Server 2.0.42 or later, or Apache Server 2.2 running under Microsoft Windows Server 2000 or later.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

- Apache Server 1.3
- Apache Server 2.0.42 or later
- Apache Server 2.2 running under Microsoft Windows Server 2000 or later

Install the Program Files

Before installing the program files, you must determine where you want to maintain the disk queue because that is also where you must install the program files. Procedure to extract and install the program files for Sensor.

To install and configure Sensor, you must perform the following steps:

1. On your Windows machine, create a directory in which to install the Sensor program files. Keep in mind that your disk queue also resides in this directory, so be sure the device you choose has sufficient space to hold a queue of the size you need.

```
C:\VisualSensor
```

-
2. Extract the contents of the installation file into the directory you just created. During this step, Sensor installs the following files:

File	Description
EventMessages.dll	Event Viewer messages.
mod_visual_sciences.dll	The collector module.
TestExperiment.xls	An Excel spreadsheet file that architects can use to configure a controlled experiment. Sensor does not use this file.
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process.
TXLog.exe	The transmitter program
txlogd.conf	The Sensor configuration file



Note: The installation package contains a spreadsheet file called TestExperiment.xls. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Edit the Sensor Configuration File

The txlogd.conf file contains the configuration parameters for Sensor.

You must edit this file to specify, among other things, the size and location of the disk queue file, the address of the Insight Server, and the ID that will be attached to the event data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

- Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
- Save and close the file.

To edit the Sensor configuration file

1. Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
2. Save and close the file.

Start the Transmitter and Create the Disk Queue

After you configure the txlogd.conf file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

-
1. If the directory in which the disk queue resides does not already exist, create it. Make sure that the directory provides both the collector module and the transmitter program with read/write access to the file.

For more information about the permissions required by the disk queue files, see Sensor UNIX File Permissions.

2. On the computer where Sensor is installed, execute the following command to start the transmitter:

```
/usr/local/bin/txlogd -ic -f /etc/txlogd.conf
```

- The “i” option in this command starts the transmitter in “interactive mode.” This mode displays transmitter messages on the screen and also enables you to interact with the transmitter using keyboard commands.
- The “c” option directs the transmitter to create the disk queue.
- The “f” option specifies the location of the configuration file.

For additional information about the options you can use when starting the transmitter, see Sensor Transmitter Command-Line Options.

3. Verify that the transmitter has created the disk queue in the location specified in the QueueFile parameter and of the size specified in the QueueSize parameter.
4. If the queue has not been created correctly, type Ctrl+C to terminate the transmitter, then do the following:
 - a. Examine the txlogd.conf file and verify that the QueueFile and QueueSize parameters are set correctly.
 - b. Check that the device to which the disk queue is assigned is operational and has sufficient space available to hold a file of the size specified in the QueueSize parameter.
 - c. Make any necessary corrections and repeat this procedure.

Add the Collector to the Web Server

For Apache servers, the collector is a dynamic shared object that you load into your web server process.

To add the collector to your web server, you must edit the httpd.conf file as described below and restart your web server.



Note: If Sensor is capturing data for multiple web servers on the server computer, you must perform the following procedure for each web server.

1. Using a text editor, open the httpd.conf file for the web server whose events Sensor captures.
2. Add the following two lines to the end of the file:

```
LoadModule visual_sciences_module modules/mod_visual_sciences.so  
VisualSciencesConfig /etc/txlogd.conf
```



Note: These lines are case sensitive. Type them exactly as they appear above.

3. Restart the web server process (you do not have to reboot the entire server computer, simply restart the web server process). The collector is loaded with the web server and begins collecting event data and writing it to the disk queue.

JBoss, Tomcat, and WebLogic Servers on RedHat Linux or Sun Solaris

Detailed instructions for installing and configuring Sensor for J2EE implementations running on RedHat Linux 7.x or later, Sun Solaris SPARC 2.6 or later, or Sun Solaris x86 9 or later.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

Supported J2EE implementations include:

- JBoss Server 3.2.x or later
- Apache Jakarta Tomcat Server 4.1 or later
- WebLogic Server 6.x or later

To install and configure Sensor, you must perform the following steps:

Install the Program Files

Procedure to extract and install the program files for Sensor.

1. Log on as the root user or as a user with root authority.
2. Decompress and unpack the installation file using the following command:

- On Linux:

```
tar -zxf installationFilename.tar.gz
```

- On Solaris:

```
unzip -d installationFilename.tar.gz  
tar -xf installationFilename.tar
```

3. Copy the unpacked program files to the directories identified in the following table:

File	Description	Target Directory
mod_visual_sciences.so	The collector load module.	<i>IBMHttpServer/modules</i>
txlogd	The transmitter program.	<i>/usr/local/bin</i> --OR-- <i>/usr/local/sbin</i>
txlogd.conf	The Sensor configuration file.	<i>/etc</i>
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process	<i>/usr/local/visual_sciences</i>



Note: The installation package contains a spreadsheet file called TestExperiment.xls. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Permissions on the Program Files



Note: Incorrect permissions on the program files cause the majority of problems encountered when installing Sensor. Please make sure that you set the permissions exactly as stated in this section.

By default, the program files in the tar file have the following permissions. Depending on how your system is configured, these settings might be altered (unmasked) when you extract the files. To reset the permissions

to the recommended default settings, use the chmod commands below. Check that the directories into which you have installed the files permit at least this level of access.

File	Default Permissions	chmod command
mod_visual_sciences.so	rwx r-x r-x	chmod 775
txlogd	rwx --x --x	chmod 711
txlogd.conf	rw- r-- r--	chmod 664
trust_ca_cert.pem	rw- r-- r--	chmod 664

Edit the Sensor Configuration File

The txlogd.conf file contains the configuration parameters for Sensor.

You must edit this file to specify, among other things, the size and location of the disk queue file, the address of the Insight Server, and the ID that will be attached to the event data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

- Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
- Save and close the file.

To edit the Sensor configuration file

1. Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
2. Save and close the file.

Start the Transmitter and Create the Disk Queue

After you configure the txlogd.conf file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. If the directory in which the disk queue resides does not already exist, create it. Make sure that the directory provides both the collector module and the transmitter program with read/write access to the file.

For more information about the permissions required by the disk queue files, see Sensor UNIX File Permissions.

2. On the computer where Sensor is installed, execute the following command to start the transmitter:

```
/usr/local/bin/txlogd -ic -f /etc/txlogd.conf
```

- The “i” option in this command starts the transmitter in “interactive mode.” This mode displays transmitter messages on the screen and also enables you to interact with the transmitter using keyboard commands.
- The “c” option directs the transmitter to create the disk queue.
- The “f” option specifies the location of the configuration file.

For additional information about the options you can use when starting the transmitter, see Sensor Transmitter Command-Line Options.

-
3. Verify that the transmitter has created the disk queue in the location specified in the QueueFile parameter and of the size specified in the QueueSize parameter.
 4. If the queue has not been created correctly, type Ctrl+C to terminate the transmitter, then do the following:
 - a. Examine the txlogd.conf file and verify that the QueueFile and QueueSize parameters are set correctly.
 - b. Check that the device to which the disk queue is assigned is operational and has sufficient space available to hold a file of the size specified in the QueueSize parameter.
 - c. Make any necessary corrections and repeat this procedure.

Add the Collector to the Web Server

For Apache servers, the collector is a dynamic shared object that you load into your web server process.

To add the collector to your web server, you must edit the httpd.conf file as described below and restart your web server.



Note: If Sensor is capturing data for multiple web servers on the server computer, you must perform the following procedure for each web server.

1. Using a text editor, open the httpd.conf file for the web server whose events Sensor captures.
2. Add the following <filter> and <filter-mapping> elements to the descriptor file. If you did not install txlogd.conf in the /etc directory, you need to enter the correct path to this file in the <param-value> element:

```
<filter>
  <filter-name>VSCollectorFilter</filter-name>
  <description></description>
  <filter-class>
    com.visualsciences.collector.VSCollectorFilter
  </filter-class>
  <init-param>
    <param-name>configPath</param-name>
    <param-value>/etc/txlogd.conf</param-value>
  <description></description>
  </init-param>
</filter>

<filter-mapping>
  <filter-name>VSCollectorFilter</filter-name>
  <url-pattern>/*</url-pattern>
</filter-mapping>
```



Note: These lines are case sensitive. Type them exactly as they appear above.

3. Restart the web server process (you do not have to reboot the entire server computer, simply restart the web server process). The collector is loaded with the web server and begins collecting event data and writing it to the disk queue.

Test the Sensor

Verify that the collector is collecting event data and the transmitter is transmitting it to the target Insight Server.



Note: To verify that the transmitter can successfully send event data to the Insight Server, make sure that the target Insight Server is installed and running before you begin the following test.

-
1. If the transmitter is not already running, restart it using the following command:

```
/usr/local/bin/txlogd -i -f /etc/txlogd.conf
```

2. Open a browser (on any machine) and request a page from the web server on which Sensor is running (be sure to select a page that Sensor is monitoring).
3. After you issue the request, check the transmitter's console for messages indicating that it is sending event data to the target Insight Server.
4. If Sensor is not transmitting data successfully, verify that:
 - The target Insight Server is running.
 - The ServerAddress and ServerPort parameters are set correctly in txlogd.conf. If you specified ServerAddress using a server name, try using its numeric IP address instead.
 - The value of the CertName parameter matches the common name that appears on the digital certificate of the target Insight Server exactly.

Add the Transmitter to System Startup Script

Information about automatically loading the transmitter to your system startup script.

To ensure that the transmitter loads automatically when the web server machine is restarted, add the following command (which launches the transmitter) to your system startup script:

```
/usr/local/bin/txlogd -f /etc/txlogd.conf
```

This command starts the transmitter as a daemon. Operating and error messages that the transmitter generates are written to `syslog`.

The default Solaris setting is 60. Based on tests conducted with Sensor, which uses three semaphores for each instance, Adobe recommends that you use 1024 as your setting. This number is high enough for Sensor to function along with any other applications on the server that may require semaphores, but does not affect performance. To support this recommendation, please note that Adrian Cockcroft stated the following in his book Sun Performance and Tuning (Prentice Hall, October 1994): "Databases tend to use lots of shared memory and semaphore settings. These do not affect performance; as long as they are big enough, the programs will run."

Capturing Additional Data

Sensors for all platforms can collect any of the data available in the HTTP request and response headers.

The Sensors for the J2EE Platform provide a mechanism for collecting data that is not available on other platforms. The collector for the J2EE platform (J2EE collector) sits on the application layer, which enables it to collect sensitive data that is available only to the application and should not be exposed through page tagging or in the headers.



Note: While page tags and header modification can hide the data, it is still available to those who examine the source code of a page or look at the headers using browser plug-in tools.

For example, the J2EE collector can be used to capture cost per click (CPC) data for links displayed on a page, sensitive partner information on a page, and many other data points. The J2EE environment makes it easy for you to modify your WEBAPP to capture this custom data using our collector class.

When a Sensor for the J2EE Platform receives a request, it invokes a collector class that imports the `appendToLog` function. The `appendToLog` function appends to the initial request the query string parameters specified in the `appendToLog` function. This results in the URI of the initial request containing additional query string name-value pairs that correspond to the names and values of the data that is being captured. For example, CPC=20 would be appended to the initial request when the value of a particular ad placement or click-through link is 20 cents. Insight Server processes these values into the dataset for analysis. One additional benefit to this collection

methodology is that it allows the collection of additional data without creating extra log entries, as might be created using page tagging methodologies.

For more information about processing, see the Dataset Configuration Guide.

To capture additional data from a page

1. Add the following code to the top of the .jsp page from which you want to capture data:

```
<%@ page import="com.visualsciences.collector.VSCollector" %>
```

2. Use the appendToLog() method of the collector object to append the desired name-value pairs to the requested .jsp page's query string. The following example appends "A=1" and "B=2" to the requested .jsp page's query string for the /index.jsp page:

```
<html>
<body>
    <h1>Hello World</h1>
    <%
        VSCollector collector = new VSCollector(request, response);
        collector.appendToLog("A", "1");
        collector.appendToLog("B", "2");
    %>
</body>
</html>
```

The resulting request URI is /index.jsp?A=1&B=2.

3. Repeat this procedure for each .jsp page from which you want to capture additional data.

JBoss Server on Windows Server 2000 or Later

Detailed instructions for installing and configuring Sensor for JBoss Server 4.0.5 or later running under Microsoft Windows Server 2000 or later.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

Supported J2EE implementations include:

- JBoss Server 4.0.5 or later running on Microsoft Windows Server 2000 or later.

To install and configure Sensor, you must perform the following steps:

Install the Program Files

Procedure to extract and install the program files for Sensor.

1. On your JBoss Server, create a directory in which to install the Sensor program files. Keep in mind that your disk queue resides in this directory, so be sure the device you choose has sufficient space to hold a queue of the size you need.

```
C:\VisualSensor
```

2. Extract the contents of the installation file into the directory you just created. During this step, Sensor installs the following files:

File	Description	Target Directory
mod_visual_sciences.so	The collector load module.	<i>IBMHttpServer/modules</i>
txlogd	The transmitter program.	<i>/usr/local/bin</i>

File	Description	Target Directory
		--OR-- <i>/usr/local/sbin</i>
txlogd.conf	The Sensor configuration file.	<i>/etc</i>
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process	<i>/usr/local/visual_sciences</i>



Note: The installation package contains a spreadsheet file called `TestExperiment.xls`. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Edit the Sensor Configuration File

The `txlogd.conf` file contains the configuration parameters for Sensor.

You must edit this file to specify, among other things, the size and location of the disk queue file, the address of the Insight Server, and the ID that will be attached to the event data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

- Open the `/etc/txlogd.conf` file in a text editor and set the required parameters as well as any desired optional parameters.
- Save and close the file.

To edit the Sensor configuration file

1. Open the `/etc/txlogd.conf` file in a text editor and set the required parameters as well as any desired optional parameters.
2. Save and close the file.

Start the Transmitter and Create the Disk Queue

After you configure the `txlogd.conf` file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. From the Start menu in Windows, select Accessories > Command Prompt.
2. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

This command starts the transmitter, creates the disk queue, and registers Sensor as a Windows service.

-
3. To confirm that the transmitter is running correctly, click Start > Control Panel > Administrative Tools > Services.



Note: This command sequence might vary depending on which version of Windows you are using.

- a. In the service list, locate the entry for Sensor and confirm that its status is Started and its startup type is Automatic.
- b. Close the Services control panel.
4. To check whether transmitter experienced any errors during start-up, click Start > Control Panel > Administrative Tools > Event Viewer to open the Event Viewer.
 - a. In the left pane of the Event Viewer window, select the Applications log.
 - b. In the right pane, look for events with “Adobe” in the Source column.
 - c. If you find an error from “Adobe,” double-click the error to display the Event Properties window. This window provides detailed information about the error.
5. When you finish examining the Applications log, close the Event Viewer.
6. Verify that the transmitter has created the disk queue (Diskq2000.log) in the directory where you installed the Sensor program files and that it is the size that you specified in the QueueSize parameter in the txlogd.conf file.

If the queue has not been created correctly:

1. Examine the txlogd.conf file and verify that the QueueSize parameter is set correctly.
2. Check that the device on which you installed Sensor has sufficient space available to hold a file of the size specified in the QueueSize parameter.
3. Using the Services control panel in Windows, stop the transmitter.
4. Delete the queue file.
5. Re-register Sensor as a Windows service: from the Start menu in Windows, select Accessories > Command Prompt. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

The transmitter is designed to run continuously. If you restart the machine, the transmitter restarts automatically. If you need to start and stop the transmitter manually, you can do so using the Services control panel in Windows.

Add the Collector to the Web Server

For JBoss servers, the collector operates as a filter in the servlet container.

To add the collector to your web server, you must edit the web.xml file as described below and restart your web application.

1. Using a text editor, open the web.xml file for the web server whose events Sensor captures.
2. Add the following <filter> and <filter-mapping> elements to the descriptor file. If you did not install txlogd.conf in the /etc directory, you need to enter the correct path to this file in the <param-value> element:

```
<filter>
  <filter-name>VSCollectorFilter</filter-name>
  <description></description>
  <filter-class>
    com.visualsciences.collector.VSCollectorFilter
  </filter-class>
<init-param>
```

```
<param-name>configPath</param-name>
  <param-value>C:/VisualSensor/txlogd.conf</param-value>
  <description></description>
  </init-param>
</filter>

<filter-mapping>
  <filter-name>VSCollectorFilter</filter-name>
  <url-pattern>/*</url-pattern>
</filter-mapping>
```



Note: These lines are case sensitive. Type them exactly as they appear above.

3. Restart the web server process (you do not have to reboot the entire server computer, simply restart the web server process). The collector is loaded with the web server and begins collecting event data and writing it to the disk queue.

Modify the Startup Script

Before modifying the startup script, ensure that the JAVA_HOME variable is defined in the Windows environment.

In the run.bat file (for example, C:\jboss-4.0.5.GA\bin\run.bat), add the following lines near the end of the file just before the “echo” lines that precede the startup command of the JBoss server:

```
set
JBoss_CLASSPATH=%JBoss_CLASSPATH%;C:\jboss-4.0.5.GA\server\default\lib\jax.servlet.jar;C:\VisualSciences\J2EECollector.jar
set JAVA_OPTS=%JAVA_OPTS% -Djava.library.path=C:\VisualSciences
```

Capturing Additional Data

You can capture additional measurement data from J2EE-based web applications using the appendToLog() functionality.

1. Add the following code to the top of the .jsp page from which you want to capture data:

```
<%@ page import="com.visualsciences.collector.VSCollector" %>
```

2. Use the appendToLog() method of the collector object to append the desired name-value pairs to the requested .jsp page's query string. The following example appends “A=1” and “B=2” to the requested .jsp page's query string for the /index.jsp page:

```
<html>
<body>
  <h1>Hello World</h1>
  <%
    VSCollector collector = new VSCollector(request, response);
    collector.appendToLog("A", "1");
    collector.appendToLog("B", "2");
  %>
</body>
</html>
```

The resulting request URI is /index.jsp?A=1&B=2.

3. Repeat this procedure for each .jsp page from which you want to capture additional data.

Microsoft IIS on Windows Server 2008 or Later

Install and configure Sensor for Microsoft IIS 7.x or 8.x running under Microsoft Windows Server 2008 or later.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

To install and configure Sensor, you must perform the following high-level steps:

1. *Install the program files* on page 916
2. *Edit the Sensor configuration file* on page 916
3. *Start the transmitter and create the disk queue* on page 917
4. *Add the collector to the web server* on page 918
5. *Capture additional data* on page 918

Install the program files

When running Sensor on Windows IIS, the program files and the disk queue file must reside in the same directory.

Before installing the program files, first determine where you want to maintain the disk queue because that is where you must install the program files.

Use the following procedure to extract and install the program files for Sensor.

1. On your Windows machine, create a directory in which to install the Sensor program files. Keep in mind that your disk queue also resides in this directory, so be sure the device you choose has sufficient space to hold a queue of the size you need.

For example: C:\VisualSensor

2. Extract the contents of the installation file into the directory you just created. During this step, Sensor installs the following files:

File	Description
EventMessages.dll	Event Viewer messages.
qlog.dll	The collector module (an ISAPI filter).
TestExperiment.xls	An Excel spreadsheet file that architects can use to configure a controlled experiment. Sensor does not use this file.
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process.
TXLog.exe	The transmitter program.
txlogd.conf	The Sensor configuration file.



Note: The installation package contains a spreadsheet file called TestExperiment.xls. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Edit the Sensor configuration file

The txlogd.conf file contains the configuration parameters for Sensor.

You must edit the file to specify, among other things, the size of the disk queue, the address of the Insight Server, and the ID that will be attached to the data produced by this sensor. The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

1. Open the <SensorDirectory>/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.

For descriptions of txlogd.conf parameters, see *Sensor Txlogd.conf File Parameters* on page 945.

2. Save and close the file.

Start the transmitter and create the disk queue

After you configure the txlogd.conf file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. From the Start menu in Windows, select Accessories > Command Prompt.
2. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

This command starts the transmitter, creates the disk queue, and registers Sensor as a Windows service.

3. To confirm that the transmitter is running correctly, click Start > Control Panel > Administrative Tools > Services.



Note: This command sequence might vary depending on which version of Windows you are using.

- a. In the service list, locate the entry for Sensor and confirm that its status is Started and its startup type is Automatic.
- b. Close the Services control panel.

4. To check whether transmitter experienced any errors during start-up, click Start > Control Panel > Administrative Tools > Event Viewer to open the Event Viewer.



Note: This command sequence might vary depending on which version of Windows you are using.

- a. In the left pane of the Event Viewer window, select the Applications log.
- b. In the right pane, look for events with “Adobe” in the Source column.
- c. If you find an error from “Adobe,” double-click the error to display the Event Properties window. This window provides detailed information about the error.

5. When you finish examining the Applications log, close the Event Viewer.

6. Verify that the transmitter has created the disk queue (Diskq2008.log) in the directory where you installed the Sensor program files and that it is the size that you specified in the QueueSize parameter in the txlogd.conf file.

If the queue has not been created correctly:

- a. Examine the txlogd.conf file and verify that the QueueSize parameter is set correctly.

-
- b. Check that the device on which you installed Sensor has sufficient space available to hold a file of the size specified in the QueueSize parameter.
 - c. Using the Services control panel in Windows, stop the transmitter.
 - d. Delete the queue file.
 - e. Re-register Sensor as a Windows service: from the Start menu in Windows, select Accessories > Command Prompt. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

The transmitter is designed to run continuously. If you restart the machine, the transmitter restarts automatically. If you need to start and stop the transmitter manually, you can do so using the Services control panel in Windows.

Add the collector to the web server

For IIS, the collector is an ISAPI filter that you add to your web server in IIS.

1. Open the IIS Manager using **Start > Administrative Tools > Internet Information Services (IIS) Manager**.
2. Expand the **Local Computer** and **Sites** nodes.
3. Select the website, and in the right-pane, double-click **ISAPI Filters**.
4. Under the **Actions** pane, click **Add**.
5. In the **Filter Name** field, enter a display name for the filter. The suggested filter name is “Sensor.”
6. Click **Browse**, select the qlog.dll file (located in the directory where you installed Sensor), and click **OK**.
7. Click **OK** to add the filter.

After you add the filter, the collector is immediately operational and ready to collect data.

If the green arrow does not appear after traffic flows to the collector, complete the following steps:

1. Click Start > Administrative Tools > Event Viewer to check the Event Viewer for errors.



Note: This command sequence might vary depending on which version of Windows you are using.

2. In the left pane of the Event Viewer window, select the **Application** log.
3. In the right pane, look for events with “Adobe” in the **Source** column.
4. If you find an error, double-click the error to display the **Event Properties** window.

Capture additional data

Web pages are often structured using ASP (Active Server Pages) programming language.

ASP is a Microsoft technology that runs within IIS. When a browser requests an ASP file, IIS passes the request to the ASP engine. The ASP engine reads the ASP file, line by line, and executes the scripts in the file. Finally, the ASP file is returned to the browser as plain HTML. ASP provides RESPOND or REQUEST objects which, in addition to other utilizations, allow the response or request of user queries or data submitted from HTML forms.

In certain cases, you may not want to append the values entered into forms to the URL that is displayed within the Address bar of a user’s browser or that is viewable within the HTML code itself. Simple server-side ASP Script lets you append form field names and their respective values to the log file without making them available within the user’s browser or embedding them into the HTML file. To capture the actual form values entered into particular forms within your website, a few lines of code must be added to append the form values to the log request.

Within the processing page of a form, include the following code to append the entered form values to the request data (in addition to writing the submitted form values to an external database or other location):

```
var sName= Request.Form("Name");
var sCity= Request.Form("City");
var sState= Request.Form("State");
var sZip= Request.Form("Zip");

Response.AppendToLog("&v_1=" + sName);
Response.AppendToLog("&v_2=" + sCity);
Response.AppendToLog("&v_3=" + sState);
Response.AppendToLog("&v_4=" + sZip);
```

This process would append the form values as defined to the request data for the Form Processing page. Within the log data, the appended values would be available as query strings of the Form Processing page as illustrated below. For example, v_1, v_2, v_3 and v_4 would now be query strings containing the data entered into the appropriate form fields. The syntax described in the previous example can be duplicated for any additional form fields and values that you want to capture:

http://www.myserver.com/path/to/formprocessingpage.asp?v_1=John+Smith&v_2=Los+Angeles&v_3=California&v_4=90210

If you want every form field and value to be captured and available for analysis, you can use the following syntax:

```
var formvalues = Response.Form;
Response.AppendToLog(formvalues);
```

This example would take all form fields present within the HTML along with their respective values and append them as query strings to the log entry for the Form Processing page. Please note that this would include any hidden fields present within the form.

The log data would be augmented as detailed in the following table:

Data Collected	Explanation	Example
v_1	Value associated with the NAME query string	v_1=John Smith
v_2	Value associated with the CITY query string	v_2=Los Angeles
v_3	Value associated with the STATE query string	v_3=California
v_4	Value associated with the ZIP query string	v_4=90210

Tomcat Server on Windows Server 2000 or Later

Detailed instructions for installing and configuring Sensor for Apache Jakarta Tomcat 4.1 or later running under Windows Server 2000 or later.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

Supported J2EE implementations include:

- JBoss Server 4.0.5 or later running on Microsoft Windows Server 2000 or later.

To install and configure Sensor, you must perform the following steps:

Install the Program Files

Procedure to extract and install the program files for Sensor.

-
1. On your Tomcat Server, create a directory in which to install the Sensor program files. Keep in mind that your disk queue resides in this directory, so be sure the device you choose has sufficient space to hold a queue of the size you need.

C:\VisualSensor

2. Extract the contents of the installation file into the directory you just created. During this step, Sensor installs the following files:

File	Description	Target Directory
visual_sciences.dll	The collector load module.	In any directory.
J2EECollector.jar	The collector load module libraries	WEB-INF/lib
txlogd.exe	The transmitter program.	In any directory
txlogd.conf	The Sensor configuration file.	In any directory
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process	In any directory



Note: The installation package contains a spreadsheet file called `TestExperiment.xls`. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Edit the Sensor Configuration File

The `txlogd.conf` file contains the configuration parameters for Sensor.

You must edit this file to specify, among other things, the size and location of the disk queue file, the address of the Insight Server, and the ID that will be attached to the event data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

- Open the `/etc/txlogd.conf` file in a text editor and set the required parameters as well as any desired optional parameters.
- Save and close the file.

To edit the Sensor configuration file

1. Open the `/etc/txlogd.conf` file in a text editor and set the required parameters as well as any desired optional parameters.
2. Save and close the file.

Start the Transmitter and Create the Disk Queue

After you configure the txlogd.conf file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. From the Start menu in Windows, select Accessories > Command Prompt.
2. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

This command starts the transmitter, creates the disk queue, and registers Sensor as a Windows service.

3. To confirm that the transmitter is running correctly, click Start > Control Panel > Administrative Tools > Services.



Note: This command sequence might vary depending on which version of Windows you are using.

- a. In the service list, locate the entry for Sensor and confirm that its status is Started and its startup type is Automatic.
- b. Close the Services control panel.
4. To check whether transmitter experienced any errors during start-up, click Start > Control Panel > Administrative Tools > Event Viewer to open the Event Viewer.
 - a. In the left pane of the Event Viewer window, select the Applications log.
 - b. In the right pane, look for events with “Adobe” in the Source column.
 - c. If you find an error from “Adobe,” double-click the error to display the Event Properties window. This window provides detailed information about the error.
5. When you finish examining the Applications log, close the Event Viewer.
6. Verify that the transmitter has created the disk queue (Diskq2000.log) in the directory where you installed the Sensor program files and that it is the size that you specified in the QueueSize parameter in the txlogd.conf file.

If the queue has not been created correctly:

1. Examine the txlogd.conf file and verify that the QueueSize parameter is set correctly.
2. Check that the device on which you installed Sensor has sufficient space available to hold a file of the size specified in the QueueSize parameter.
3. Using the Services control panel in Windows, stop the transmitter.
4. Delete the queue file.
5. Re-register Sensor as a Windows service: from the Start menu in Windows, select Accessories > Command Prompt. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

The transmitter is designed to run continuously. If you restart the machine, the transmitter restarts automatically. If you need to start and stop the transmitter manually, you can do so using the Services control panel in Windows.

Add the Collector to the Web Server

For JBoss servers, the collector operates as a filter in the servlet container.

To add the collector to your web server, you must edit the `web.xml` file as described below and restart your web application.

-
1. Using a text editor, open the web.xml file for the web server whose events Sensor captures.
 2. Add the following <filter> and <filter-mapping> elements to the descriptor file. If you did not install txlogd.conf in the /etc directory, you need to enter the correct path to this file in the <param-value> element:

```
<filter>
  <filter-name>VSCollectorFilter</filter-name>
  <description></description>
  <filter-class>
    com.visualsciences.collector.VSCollectorFilter
  </filter-class>
  <init-param>
    <param-name>configPath</param-name>
    <param-value>C:/VisualSensor/txlogd.conf</param-value>
  <description></description>
  </init-param>
</filter>

<filter-mapping>
  <filter-name>VSCollectorFilter</filter-name>
  <url-pattern>/*</url-pattern>
</filter-mapping>
```



Note: These lines are case sensitive. Type them exactly as they appear above.

3. Restart the web server process (you do not have to reboot the entire server computer, simply restart the web server process). The collector is loaded with the web server and begins collecting event data and writing it to the disk queue.

Modify the Java Library Path

Instructions for adding the visual_science.dll to the Tomcat java library path.

1. On your Windows server, navigate to Tomcat installation directory. (Tomcat > bin)
2. Under bin folder, run Tomcat9w.exe (common daemon service manager).

In the Java tab, under Java options, add a new line:

```
-Djava.library.path=C:\Sensor directory
```

Where the C:\Sensor directory is the directory containing the visual_science.dll file.

Capturing Additional Data

You can capture additional measurement data from J2EE-based web applications using the appendToLog() functionality.

1. Add the following code to the top of the .jsp page from which you want to capture data:

```
<%@ page import="com.visualsciences.collector.VSCollector" %>
```

2. Use the appendToLog() method of the collector object to append the desired name-value pairs to the requested .jsp page's query string. The following example appends "A=1" and "B=2" to the requested .jsp page's query string for the /index.jsp page:

```
<html>
<body>
  <h1>Hello World</h1>
  <%
    VSCollector collector = new VSCollector(request, response);
    collector.appendToLog("A", "1");
    collector.appendToLog("B", "2");
```

```
%>  
</body>  
</html>
```

The resulting request URI is /index.jsp?A=1&B=2.

3. Repeat this procedure for each .jsp page from which you want to capture additional data.

WebLogic Server on Windows Server 2000 or Later

Detailed instructions for installing and configuring Sensor for WebLogic Server 6.x or later running under Microsoft Windows Server 2000 or later.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

To install and configure Sensor, you must perform the following steps:

Install the Program Files

Procedure to extract and install the program files for Sensor.

1. On your WebLogic Server, create a directory in which to install the Sensor program files. Keep in mind that your disk queue resides in this directory, so be sure the device you choose has sufficient space to hold a queue of the size you need.

```
C:\VisualSensor
```

2. Extract the contents of the installation file into the directory you just created. During this step, Sensor installs the following files:

File	Description	Target Directory
visual_sciences.dll	The collector load module.	In any directory.
J2EECollector.jar	The collector load module libraries	WEB-INF/lib
txlogd.exe	The transmitter program.	In any directory
txlogd.conf	The Sensor configuration file.	In any directory
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process	In any directory



Note: The installation package contains a spreadsheet file called `TestExperiment.xls`. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Edit the Sensor Configuration File

The `txlogd.conf` file contains the configuration parameters for Sensor.

You must edit this file to specify, among other things, the size and location of the disk queue file, the address of the Insight Server, and the ID that will be attached to the event data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

- Open the `/etc/txlogd.conf` file in a text editor and set the required parameters as well as any desired optional parameters.
- Save and close the file.

To edit the Sensor configuration file

1. Open the `/etc/txlogd.conf` file in a text editor and set the required parameters as well as any desired optional parameters.
2. Save and close the file.

Start the Transmitter and Create the Disk Queue

After you configure the `txlogd.conf` file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. From the Start menu in Windows, select Accessories > Command Prompt.
2. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

This command starts the transmitter, creates the disk queue, and registers Sensor as a Windows service.

3. To confirm that the transmitter is running correctly, click Start > Control Panel > Administrative Tools > Services.



Note: This command sequence might vary depending on which version of Windows you are using.

- a. In the service list, locate the entry for Sensor and confirm that its status is Started and its startup type is Automatic.
 - b. Close the Services control panel.
4. To check whether transmitter experienced any errors during start-up, click Start > Control Panel > Administrative Tools > Event Viewer to open the Event Viewer.
 - a. In the left pane of the Event Viewer window, select the Applications log.
 - b. In the right pane, look for events with “Adobe” in the Source column.
 - c. If you find an error from “Adobe,” double-click the error to display the Event Properties window. This window provides detailed information about the error.
 5. When you finish examining the Applications log, close the Event Viewer.
 6. Verify that the transmitter has created the disk queue (`Diskq2000.log`) in the directory where you installed the Sensor program files and that it is the size that you specified in the `QueueSize` parameter in the `txlogd.conf` file.

If the queue has not been created correctly:

-
1. Examine the txlogd.conf file and verify that the QueueSize parameter is set correctly.
 2. Check that the device on which you installed Sensor has sufficient space available to hold a file of the size specified in the QueueSize parameter.
 3. Using the Services control panel in Windows, stop the transmitter.
 4. Delete the queue file.
 5. Re-register Sensor as a Windows service: from the Start menu in Windows, select Accessories > Command Prompt. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

The transmitter is designed to run continuously. If you restart the machine, the transmitter restarts automatically. If you need to start and stop the transmitter manually, you can do so using the Services control panel in Windows.

Add the Collector to the Web Server

For JBoss servers, the collector operates as a filter in the servlet container.

To add the collector to your web server, you must edit the web.xml file as described below and restart your web application.

1. Using a text editor, open the web.xml file for the web server whose events Sensor captures.
2. Add the following <filter> and <filter-mapping> elements to the descriptor file. If you did not install txlogd.conf in the /etc directory, you need to enter the correct path to this file in the <param-value> element:

```
<filter>
  <filter-name>VSCollectorFilter</filter-name>
  <description></description>
  <filter-class>
    com.visualsciences.collector.VSCollectorFilter
  </filter-class>
  <init-param>
    <param-name>configPath</param-name>
    <param-value>C:/VisualSensor/txlogd.conf</param-value>
  <description></description>
  </init-param>
</filter>

<filter-mapping>
  <filter-name>VSCollectorFilter</filter-name>
  <url-pattern>/*</url-pattern>
</filter-mapping>
```



Note: These lines are case sensitive. Type them exactly as they appear above.

3. Restart the web server process (you do not have to reboot the entire server computer, simply restart the web server process). The collector is loaded with the web server and begins collecting event data and writing it to the disk queue.

Modify the Startup Script

Instructions for startup script modification.

In the script that is used to start WebLogic (for example, C:\bea\user_projects\mydomain\startServer.cmd), edit the “set JAVA_OPTIONS=” line to set the java.library.path definition to the directory containing the visual_sciences.dll file.

```
set JAVA_OPTIONS=-Djava.library.path=C:\Sensor  
directory
```

Capturing Additional Data

You can capture additional measurement data from J2EE-based web applications using the appendToLog() functionality.

1. Add the following code to the top of the .jsp page from which you want to capture data:

```
<%@ page import="com.visualsciences.collector.VSCollector" %>
```

2. Use the appendToLog() method of the collector object to append the desired name-value pairs to the requested .jsp page's query string. The following example appends “A=1” and “B=2” to the requested .jsp page's query string for the /index.jsp page:

```
<html>  
<body>  
    <h1>Hello World</h1>  
    <%  
        VSCollector collector = new VSCollector(request, response);  
        collector.appendToLog("A", "1");  
        collector.appendToLog("B", "2");  
    %>  
</body>  
</html>
```

The resulting request URI is /index.jsp?A=1&B=2.

3. Repeat this procedure for each .jsp page from which you want to capture additional data.

Netscape Enterprise on Linux or Solaris

Instructions for installing and configuring Sensor on the family of web servers that evolved from the original Netscape Enterprise Web Server running on Linux or Solaris machines. Includes Netscape Enterprise, iPlanet, Sun ONE, and Sun Java System Servers on Linux or Solaris.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

Sensor supports the following servers running under RedHat Linux 7.x or later or Sun Solaris SPARC 2.6 or later:

- Netscape Enterprise Server 3.6 or later
- iPlanet Web Server 4.0 or later

Sensor supports these servers running under RedHat Linux 7.x or later or Sun Solaris 8.x or later:

- Sun ONE Web Server 6.0 or later
- Sun Java System Web Server 6.1 or later

Sensor supports these servers running under Sun Solaris x86 9 or later:

- Sun Java System Web Server 6.1 or later



Note: The installation file for this family of web servers is listed as the “Netscape Solaris Sensor” or “Netscape LINUX Sensor” on the Adobe download site.

To install and configure Sensor, you must perform the following steps:

Install the Program Files

Procedure to extract and install the program files for Sensor.

1. Log on as the root user or as a user with root authority.
2. Decompress and unpack the installation file using the following command:

```
gunzip installationfilename.tar.gz  
tar -xf installationfilename.tar
```

3. Copy the unpacked program files to the directories identified in the following table:

File	Description	Target Directory
aol_visual_sciences.so	The collector load module.	/usr/local/aolserver/visual_sciences
txlogd	The transmitter program. --OR-- /usr/local/sbin	/usr/local/bin
txlogd.conf	The Sensor configuration file.	/etc
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process	/usr/local/visual_sciences



Note: The installation package contains a spreadsheet file called TestExperiment.xls. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Permissions on the Program Files



Note: Incorrect permissions on the program files cause the majority of problems encountered when installing Sensor. Please make sure that you set the permissions exactly as stated in this section.

By default, the program files in the tar file have the following permissions. Depending on how your system is configured, these settings might be altered (unmasked) when you extract the files. To reset the permissions to the recommended default settings, use the chmod commands below. Check that the directories into which you have installed the files permit at least this level of access.

File	Default Permissions	chmod command
mod_visual_sciences.so	rwx r-x r-x	chmod 775
txlogd	rwx --x --x	chmod 711
txlogd.conf	rw- r-- r--	chmod 664
trust_ca_cert.pem	rw- r-- r--	chmod 664

Edit the Sensor Configuration File

The `txlogd.conf` file contains the configuration parameters for Sensor.

You must edit this file to specify, among other things, the size and location of the disk queue file, the address of the Insight Server, and the ID that will be attached to the event data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

- Open the `/etc/txlogd.conf` file in a text editor and set the required parameters as well as any desired optional parameters.
- Save and close the file.

To edit the Sensor configuration file

1. Open the `/etc/txlogd.conf` file in a text editor and set the required parameters as well as any desired optional parameters.
2. Save and close the file.

Start the Transmitter and Create the Disk Queue

After you configure the `txlogd.conf` file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. If the directory in which the disk queue resides does not already exist, create it. Make sure that the directory provides both the collector module and the transmitter program with read/write access to the file.

For more information about the permissions required by the disk queue files, see Sensor UNIX File Permissions.

2. On the computer where Sensor is installed, execute the following command to start the transmitter:

```
/usr/local/bin/txlogd -ic -f /etc/txlogd.conf
```

- The “`i`” option in this command starts the transmitter in “interactive mode.” This mode displays transmitter messages on the screen and also enables you to interact with the transmitter using keyboard commands.
- The “`c`” option directs the transmitter to create the disk queue.
- The “`f`” option specifies the location of the configuration file.

For additional information about the options you can use when starting the transmitter, see Sensor Transmitter Command-Line Options.

3. Verify that the transmitter has created the disk queue in the location specified in the `QueueFile` parameter and of the size specified in the `QueueSize` parameter.
4. If the queue has not been created correctly, type `Ctrl+C` to terminate the transmitter, then do the following:
 - a. Examine the `txlogd.conf` file and verify that the `QueueFile` and `QueueSize` parameters are set correctly.
 - b. Check that the device to which the disk queue is assigned is operational and has sufficient space available to hold a file of the size specified in the `QueueSize` parameter.
 - c. Make any necessary corrections and repeat this procedure.

Add the Collector to the AOL Server

For AOLServer, the collector is a dynamic shared object that you load into your web server process.

To add the collector to your AOL server, you must edit the configuration file for your server as described below and restart your AOL server. Usually the server's configuration file is named nsd.tcl and is located in the directory where AOL Server is installed.

1. Open the configuration file in a text editor and locate the following section:

```
ns_section "ns/server/${servername}/modules"
```

2. Add the following line. (Add as a single statement. Ignore word wrap shown below.)

```
ns_param aol_visual_sciences  
/usr/local/aolserver/visual_sciences/aol_visual_sciences.so
```

3. Create a new section as follows.

```
ns_section "ns/server/${servername}/module/aol_visual_sciences"
```

4. To this new section, add the line:

```
ns_param VisualSciencesConfig /etc/txlogd.conf
```



Note: These lines are case sensitive. Type them exactly as they appear above.

5. Restart the AOL server. The collector is loaded and will begin collecting event data and writing it to the disk queue.

Test the Sensor

Verify that the collector is collecting event data and the transmitter is transmitting it to the target Insight Server.



Note: To verify that the transmitter can successfully send event data to the Insight Server, make sure that the target Insight Server is installed and running before you begin the following test.

1. If the transmitter is not already running, restart it using the following command:

```
/usr/local/bin/txlogd -i -f /etc/txlogd.conf
```

2. Open a browser (on any machine) and request a page from the web server on which Sensor is running (be sure to select a page that Sensor is monitoring).

3. After you issue the request, check the transmitter's console for messages indicating that it is sending event data to the target Insight Server.

4. If Sensor is not transmitting data successfully, verify that:

- The target Insight Server is running.
- The ServerAddress and ServerPort parameters are set correctly in txlogd.conf. If you specified ServerAddress using a server name, try using its numeric IP address instead.
- The value of the CertName parameter matches the common name that appears on the digital certificate of the target Insight Server exactly.

Add the Transmitter to System Startup Script

Information about automatically loading the transmitter to your system startup script.

To ensure that the transmitter loads automatically when the web server machine is restarted, add the following command (which launches the transmitter) to your system startup script:

```
/usr/local/bin/txlogd -f /etc/txlogd.conf
```

This command starts the transmitter as a daemon. Operating and error messages that the transmitter generates are written to `syslog`.

The default Solaris setting is 60. Based on tests conducted with Sensor, which uses three semaphores for each instance, Adobe recommends that you use 1024 as your setting. This number is high enough for Sensor to function along with any other applications on the server that may require semaphores, but does not affect performance. To support this recommendation, please note that Adrian Cockcroft stated the following in his book *Sun Performance and Tuning* (Prentice Hall, October 1994): "Databases tend to use lots of shared memory and semaphore settings. These do not affect performance; as long as they are big enough, the programs will run."

Sun Java Server on Windows Server 2000 or Later

Instructions for installing and configuring Sensor for Sun Java System Application Server Standard Edition 7 running under Windows Server 2000 or later.

The program files for Sensor are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the Sensor installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.

Sensor supports the following servers running under RedHat Linux 7.x or later or Sun Solaris SPARC 2.6 or later:

To install and configure Sensor, you must perform the following steps:

Install the Program Files

Procedure to extract and install the program files for Sensor to the server.

1. On your Netscape Enterprise, iPlanet, Sun ONE, or Sun Java System Server, create a directory in which to install the Sensor program files. Keep in mind that your disk queue resides in this directory, so be sure the device you choose has sufficient space to hold a queue of the size you need.

```
C:\VisualSensor
```

2. Extract the contents of the installation file into the directory you just created. During this step, Sensor installs the following files:

File	Description	Target Directory
saf_visual_sciences.dll	The collector load module.	/usr/local/aolserver/visual_sciences
txlogd	The transmitter program.	/usr/local/bin --OR-- /usr/local/sbin
txlogd.conf	The Sensor configuration file.	/etc
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process	/usr/local/visual_sciences



Note: The installation package contains a spreadsheet file called `TestExperiment.xls`. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file,

so it is not necessary to install the file on the machine where Sensor is running (although you can choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Edit the Sensor Configuration File

The txlogd.conf file contains the configuration parameters for Sensor.

You must edit this file to specify, among other things, the size and location of the disk queue file, the address of the Insight Server, and the ID that will be attached to the event data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- **Required parameters** are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- **Optional parameters** are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the Sensor configuration file

- Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
- Save and close the file.

To edit the Sensor configuration file

1. Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
2. Save and close the file.

Start the Transmitter and Create the Disk Queue

After you configure the txlogd.conf file, you can start the transmitter program, register it as a Windows service, and create the disk queue.

1. From the Start menu in Windows, select Accessories > Command Prompt.
2. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

This command starts the transmitter, creates the disk queue, and registers Sensor as a Windows service.

3. To confirm that the transmitter is running correctly, click Start > Control Panel > Administrative Tools > Services. In the service list, locate the entry for Sensor and confirm that its status is Started and its startup type is Automatic. Then close the Services control panel.
4. To check whether transmitter experienced any errors during start-up, click Start > Control Panel > Administrative Tools > Event Viewer to open the Event Viewer.
 - a. In the left pane of the Event Viewer window, select the Applications log.
 - b. In the right pane, look for events with “Adobe” in the Source column.
 - c. If you find an error from “Adobe,” double-click the error to display the Event Properties window. This window provides detailed information about the error.
5. When you finish examining the Applications log, close the Event Viewer.

-
6. Verify that the transmitter has created the disk queue (Diskq2000.log) in the directory where you installed the Sensor program files and that it is the size that you specified in the QueueSize parameter in the txlogd.conf file.

If the queue has not been created correctly:

1. Examine the txlogd.conf file and verify that the QueueSize parameter is set correctly.
2. Check that the device on which you installed Sensor has sufficient space available to hold a file of the size specified in the QueueSize parameter.
3. Using the Services control panel in Windows, stop the transmitter.
4. Delete the queue file.
5. Re-register Sensor as a Windows service: from the Start menu in Windows, select Accessories > Command Prompt. In the command prompt window, navigate to the directory in which you installed Sensor and execute the following command:

```
txlog /regserver
```

The transmitter is designed to run continuously. If you restart the machine, the transmitter restarts automatically. If you need to start and stop the transmitter manually, you can do so using the Services control panel in Windows.

Modify the Startup Script

In the init.conf file (for example, C:\Sun\AppServer7\domains\domain1\server1\config\init.conf), add the following lines at the end of the file:

```
Init fn="load-modules" shlib="C:/VisualSciences/saf_visual_sciences.dll"
funcs="vys-cookie,vys-log,vys-init,vys-content-type"
Init fn="vys-init" config-file="C:/VisualSciences/txlogd.conf"
```

In the obj.conf file (for example, C:\Sun\AppServer7\domains\domain1\server1\config\server1-obj.conf), add the following lines directly below the existing “<Object name="default">” line:

```
NameTrans fn="vys-cookie"
ObjectType fn="vys-content-type"
AddLog fn="vys-log"
```

WebSphere on AIX

Detailed instructions for installing and configuring **Sensor** for WebSphere 5.x running on AIX 5.1 or later.

The program files for **Sensor** are packaged in an installation file that you obtain from the Adobe download site. If you do not already have the **Sensor** installation file for your particular web server, download it (or obtain it from your Adobe representative) before you begin the following procedures.



Note: The **Sensor** for WebSphere servers does not support controlled experimentation. For information about controlled experimentation, see the *Data Workbench Controlled Experiments Guide*.

Install the Program Files

Procedure to extract and install the program files for Sensor to the server machine.

1. Log on as the root user or as a user with root authority.
2. Decompress and unpack the installation file using the following command:

```
gunzip installationfilename.tar.gz
tar -xf installationfilename.tar
```
3. Copy the unpacked program files to the directories identified in the following table:

File	Description	Target Directory
libvisual_sciences.so	The collector load module	/usr/local/visual_sciences
J2EECollector.jar	The collector load module libraries	WebSphere /lib directory
txlogd	The transmitter program	/usr/local/bin --OR-- /usr/local/sbin
txlogd.conf	The Sensor configuration file	/etc
trust_ca_cert.pem	The certificate used to validate the digital certificate that Insight Server presents during the connection process	/usr/local/visual_sciences



Note: The installation package contains a spreadsheet file called TestExperiment.xls. This spreadsheet is a tool that architects use to configure a controlled experiment. Sensor itself does not use this file, so it is not necessary to install the file on the machine where Sensor is running (although you may choose to do so). You might want to, instead, copy the file to a location where your architects can access it or simply extract the file from the installation package as needed. For more information about controlled experimentation, see the Insight Controlled Experiments Guide.

Permissions on the Program Files

Incorrect permissions on the program files cause the majority of problems encountered when installing Sensor.

Please make sure that you set the permissions exactly as stated in this section.

By default, the program files in the tar file have the following permissions. Depending on how your system is configured, these settings might be altered (unmasked) when you extract the files.

To reset the permissions to the recommended default settings, use the chmod commands below.



Note: Check that the directories into which you have installed the files permit at least this level of access.

File	Default Permissions	chmod command
libvisual_sciences.so	rwx --x --x	chmod 711
J2EECollector.jar	rw- rw- r--	chmod 664
txlogd	rwx --x --x	chmod 711
txlogd.conf	rw- rw- r--	chmod 664
trust_ca_cert.pem	rw- rw- r--	chmod 664

If you want to use permissions other than the recommended defaults, review the information in Sensor UNIX File Permissions, to be sure you understand how these files are used.

Edit the Sensor Configuration file

The txlogd.conf file contains the configuration parameters for Sensor.

You must edit the file to specify, among other things, the size of the disk queue, the address of the Insight Server, and the ID that will be attached to the data produced by this sensor.

The configuration file contains required parameters and optional parameters.

- Required parameters are settings that you must specify when you install Sensor. Without these settings, Sensor does not run successfully.
- Optional parameters are settings that default to pre-defined values (which you can modify) or enable optional features.

To edit the configuration file

1. Open the /etc/txlogd.conf file in a text editor and set the required parameters as well as any desired optional parameters.
2. Save and close the file.

Start the Transmitter and Create the Disk Queue

Procedure to create the disk queue, after you configure the txlogd.conf file.

1. If the directory in which the disk queue resides does not already exist, create it. Make sure that the directory provides both the collector module and the transmitter program with read/write access to the file.
2. On the computer where Sensor is installed, execute the following command to start the transmitter:

```
/usr/local/bin/txlogd -ic -f /etc/txlogd.conf
```

- The “i” option in this command starts the transmitter in interactive mode. This mode displays transmitter messages on the screen and also enables you to interact with the transmitter using keyboard commands.
 - The “c” option directs the transmitter to create the disk queue.
 - The “f” option specifies the location of the configuration file.
3. Verify that the transmitter has created the disk queue in the location specified in the QueueFile parameter and of the size specified in the QueueSize parameter.
 4. If the queue has not been created correctly, type Ctrl+C to terminate the transmitter, then do the following:
 - a. Examine the txlogd.conf file and verify that the QueueFile and QueueSize parameters are set correctly.
 - b. Check that the device to which the disk queue is assigned is operational and has sufficient space available to hold a file of the size specified in the QueueSize parameter.
 - c. Make any necessary corrections and repeat this procedure.

Add the Collector to the Web Application

For WebSphere servers, the collector operates as a filter in the servlet container.

To add the collector to the web application, add the filter to the web application’s web.xml deployment descriptor and restart the web application.

1. Using a text editor, open the web.xml file for the web server whose events Sensor captures.
2. Add the following <filter> and <filter-mapping> elements to the descriptor file. If you did not install txlogd.conf in the /etc directory, you need to enter the correct path to this file in the <param-value> element.

```
<filter>
  <filter-name>VSCollectorFilter</filter-name>
  <description></description>
  <filter-class>
    com.visualsciences.collector.VSCollectorFilter
  </filter-class>
  <init-param>
    <param-name>configPath</param-name>
    <param-value>C:/VisualSensor/txlogd.conf</param-value>
  </init-param>
</filter>
```

```
</init-param>
</filter>

<filter-mapping>
  <filter-name>VSCollectorFilter</filter-name>
  <url-pattern>/*</url-pattern>
</filter-mapping>
```



Note: These lines are case sensitive. Type them exactly as they appear above.

3. Restart the web application. The collector is loaded with the application and will begin collecting event data and writing it to the disk queue.

Declare the Location of the Collector and Shared Object Files

Procedure to edit the Websphere startup script to declare the location of the J2EECollector.jar and libvisual_sciences.so files.

1. Open the setupCmdLine.sh file in the Websphere /bin directory.
2. After the line that defines the \$WAS_CLASSPATH variable, add the following line:

```
WAS_CLASSPATH= "$WAS_CLASSPATH" : "$WAS_HOME" /lib/J2EECollector.jar
```

3. After the case block that defines the \$WAS_LIBPATH variable, add the following line:

```
WAS_LIBPATH= "$WAS_LIBPATH" : /usr/local/visual_sciences
```

4. Save the setupCmdLine.sh file.

Test the Sensor

Procedure to start the transmitter and verify that it can successfully connect to the Insight Server and transmit event data to it.



Note: To verify that the transmitter can successfully send event data to the Insight Server, make sure that the target Insight Server is installed and running before you begin the following test.

1. If the transmitter is not already running, restart it using the following command:

```
/usr/local/bin/txlogd -i -f /etc/txlogd.conf
```

2. Open a browser (on any machine) and request a page from the web server on which Sensor is running (be sure to select a page that Sensor is monitoring).
3. After you issue the request, check the transmitter's console for messages indicating that it is sending event data to the target Insight Server.
4. If Sensor is not transmitting data successfully, verify that:
 - The target Insight Server is running.
 - The ServerAddress and ServerPort parameters are set correctly in txlogd.conf. If you specified ServerAddress using a server name, try using its numeric IP address instead.
 - The value of the CertName parameter matches the common name that appears on the digital certificate of the target Insight Server exactly.

Add the Transmitter to Your System Startup Script

Information to ensure that the transmitter loads automatically when the web server machine is restarted.

Add the following command (which launches the transmitter) to your system startup script.

```
/usr/local/bin/txlogd -f /etc/txlogd.conf
```

This command starts the transmitter as a daemon. Operating and error messages that the transmitter generates are written to syslog.

Capturing Additional Data

Sensors for all platforms can collect any of the data available in the HTTP request and response headers.

The Sensors for the J2EE Platform provide a mechanism for collecting data that is not available on other platforms. The collector for the J2EE platform (J2EE collector) sits on the application layer, which enables it to collect sensitive data that is available only to the application and should not be exposed through page tagging or in the headers.



Note: While page tags and header modification can hide the data, it is still available to those who examine the source code of a page or look at the headers using browser plug-in tools.

For example, the J2EE collector can be used to capture cost per click (CPC) data for links displayed on a page, sensitive partner information on a page, and many other data points. The J2EE environment makes it easy for you to modify your WEBAPP to capture this custom data using our collector class.

When a Sensor for the J2EE Platform receives a request, it invokes a collector class that imports the appendToLog function. The appendToLog function appends to the initial request the query string parameters specified in the appendToLog function. This results in the URI of the initial request containing additional query string name-value pairs that correspond to the names and values of the data that is being captured. For example, CPC=20 would be appended to the initial request when the value of a particular ad placement or click-through link is 20 cents. Insight Server processes these values into the dataset for analysis. One additional benefit to this collection methodology is that it allows the collection of additional data without creating extra log entries, as might be created using page tagging methodologies.

For more information about processing, see the *Dataset Configuration Guide*.

1. Add the following code to the top of the .jsp page from which you want to capture data:

```
<%@ page import="com.visualsciences.collector.VSCollector" %>
```

2. Use the appendToLog() method of the collector object to append the desired name-value pairs to the requested .jsp page's query string. The following example appends "A=1" and "B=2" to the requested .jsp page's query string for the /index.jsp page:

```
<html>
<body>
    <h1>Hello World</h1>
    <%
        VSCollector collector = new VSCollector(request, response);
        collector.appendToLog("A", "1");
        collector.appendToLog("B", "2");
    %>
</body>
</html>
```

The resulting request URI is /index.jsp?A=1&B=2.

3. Repeat this procedure for each .jsp page from which you want to capture additional data.

Administering Sensor

Information about administrative concerns such as editing the Sensor txlogd.conf file, reviewing Sensor Queue Size, and more.

Editing the Sensor txlogd.conf File

Navigation title:Editing the Sensor txlogd.conf File

The **Sensor** configuration file, `txlogd.conf`, contains parameters that **Sensor** uses to establish a connection with the data workbench server.

These parameters include the server's address, port number, and communication protocol (HTTP or HTTPS). The configuration file also contains log-content filtering options and controlled experiment information. Controlled experiments enable site designers to perform experiments to test content or design changes on a subset of all site visitors.

When changing other `txlogd.conf` parameters, such as the IP address of the data workbench server or the name of the **Sensor**, you may be able to simply replace `txlogd.conf` with the updated version and **Sensor** will pick up the changes automatically. On some systems, you may not be able to edit or replace the file without stopping either the web server or the transmitter process.

To open and edit txlogd.conf

1. Browse to the **Sensor** installation directory and open the `txlogd.conf` file in a text editor.
2. Edit the file as necessary.
3. Save and close the file.

- The location of the controlled experiment configuration file (defined by the `ExpFile` parameter in the `txlogd.conf` file) should be in a directory on the web server that is accessible to your web content developers or controlled by your content management system.
- The value in the `ExpCookieURL` parameter is a virtual page that is used to test websites. A user visiting that page will be given a new tracking ID.

For more information about working with `txlogd.conf`, please contact Adobe Consulting Services.

Modify the Java Library Path

Instructions for adding the `visual_sciences.dll` to the Tomcat java library path.

1. On your Windows server, navigate to Tomcat installation directory. (Tomcat > bin)
2. Under bin folder, run `Tomcat9w.exe` (common daemon service manager).

In the Java tab, under Java options, add a new line:

```
-Djava.library.path=C:\Sensor directory
```

Where C:\Sensor directory is the directory containing the `visual_sciences.dll` file.

Confirming that the Data Collector is Running

Navigation title:Confirming that the Data Collector is Running

Check whether the collector is running using different methods.

Recommended Frequency: Every 5-10 minutes

- [Using Site Test](#) on page 937
- [Checking for a Cookie](#) on page 938

Using Site Test

One way to verify that the collector is running is to enable the Site Test function in the transmitter. When you enable Site Test, the transmitter periodically (every 60 seconds) sends a GET request to the web server on which the collector is running. If Site Test does not get a response from the web server, it writes an error message to syslog and sends an error message to the data workbench server (which is written to the sensor-log file).

If Site Test receives a response from the web server, it looks in the queue file for a packet from the web server. If the packet does not appear (indicating that the collector was not running to capture the event), Site Test writes an error message to syslog and sends an error message to Adobe (which also is written to the sensor-log file).

In the requests that Site Test sends to the web server, Site Test sets the User-Agent value to “**Sensor** Test.” If you do not want these requests to appear in your dataset, add the “**Sensor** Test” User-Agent to the Baseline Robots List.txt file or the Extended Robots List.txt file in the Lookups folder on the data workbench server.

To enable Site Test in the transmitter

1. Locate the txlogd.conf file on the machine where **Sensor** is running and open it in a text editor.
2. In the txlogd.conf file, locate the “SiteTest” line and configure it as shown below. If your txlogd.conf file does not include the “SiteTest” line, simply add the line to the end of the configuration file.

SiteTest http, *serverAddress*, *port*, *resource*

where *serverAddress* is the web server’s name or IP address, *port* is the server’s HTTP listening port, and *resource* is the specific resource that you want Site Test to request when testing the server. Note that *resource* can include a query string.

Example: SiteTest http,localhost,80,/index.jsp

To test multiple web servers, you simply specify multiple SiteTest lines.

Checking for a Cookie

Another way to verify that the collector is running on a web server is to check whether **Sensor** is setting a cookie in the responses that the web server is returning to clients. If the collector is working, the web server returns a “v1st” cookie.

It is possible to rename the cookie. If you have done so, you must look for the name specified, not v1st.

You can perform this check using an automated script or monitoring agent. For a sample script or additional help with this task, please contact Adobe Consulting Services.

Confirming that the Data Transmitter is Running

Navigation title:Confirming that the Data Transmitter is Running

Check whether the transmitter is running by setting up alerts, checking the system status of the **Sensor**, and more.

Recommended Frequency: Every 5-10 minutes

- [Checking the Transmitter Process](#) on page 938
- [Setting Up Administrative Alerts in the Data Workbench Server](#) on page 938
- [Checking the System Status of the Sensor](#) on page 939

Checking the Transmitter Process

One way to verify that the transmitter is running is to check that the **Sensor** transmitter process is running on each web server where a **Sensor** instance is installed. The transmitter process appears as “txlogd” in the web server’s list of processes. You can perform this check using a system monitoring tool.

Setting Up Administrative Alerts in the Data Workbench Server

Another way to verify that the transmitter is running is to set up automated administrative alerts in the data workbench server. When administrative alerts have been configured, the data workbench server generates an email alert when it has received no data from a configured and previously connected **Sensor** within the time frame specified in the **Sensor** Alert Timeout (min) parameter in the data workbench server’s Administrative

`Alerts.cfg` file. For more information about setting up administrative alerts, see the *Server Products Installation and Administration Guide*.

Checking the System Status of the Sensor

Yet another way to verify that the transmitter is running is to manually check the **Servers Manager** in Data Workbench.

To view the Servers Manager

- In Data Workbench, right-click within a workspace, click **Admin**, then under **Manage**, click **Servers**.

If the icon for a **Sensor** is green, the transmitter is running.

For more information about the **Servers Manager**, see the Administrative Interfaces chapter of the *Data Workbench Sensor Guide*.

Monitoring Administrative Events

Navigation title: Monitoring Administrative Events

To detect **Sensor** errors as soon as possible and repair them before they cause major problems or outages, you should regularly monitor your Event logs.

Recommended Frequency: At least hourly

You can monitor these events using the Windows Event Viewer or Unix Syslog file and the `*.sensor-log` files located by default in the `Logs` folder within the **Sensor** installation directory. These files indicate the presence of errors during data collection, especially if a **Sensor** cannot connect to the target data workbench server and starts queuing data.

Monitoring Events on Windows

Sensor logs errors to the Application Log of the Windows Event Viewer with a source of “Adobe.”

Messages are logged as “Information,” “Warning,” or “Error” depending on their severity.

To open the Windows Event Viewer:

- Click **Start > Control Panel > Administrative Tools > Event Viewer**.

Monitoring Events on Unix

Sensor logs errors to the `syslog` daemon.

The `syslog` daemon writes error messages to log files based on the rules you specified in your `syslog.conf` file. Errors are logged with the flags “`LOG_DAEMON`” and either “`LOG_NOTICE`” or “`LOG_ERR`,” depending on severity.

To open the Unix syslog

The Unix `syslog` is typically located in `/var/adm/messages` or `/var/log/messages`.

Browse to the appropriate location and open the `syslog`.

Understanding Message Formats

All Sensor messages contain the string “Sensor” and are numbered to reflect the importance of the message being displayed.

Message Number	Message Meaning	Message String
Ixxx	Informational	Sensor Info #

Message Number	Message Meaning	Message String
2xxx	Warning	Sensor Warning #
3xxx	Configuration Error	Sensor Error #
4xxx	Operational Error	Sensor Error #
5xxx	Internal Error	Sensor Error #



Note: Warnings (2xxx) are not currently in use. These numbers are reserved for future use.

Your network management tool can be set to monitor your messages every 5-10 minutes for errors with the “Sensor” source and alert appropriate personnel about issues that may require intervention. You can choose to monitor the system for only certain types of Event messages, such as the “Sensor Error” string. Alternatively, you can apply different rules to events prefaced with “Sensor Info,” “Sensor Warning,” and “Sensor Error” strings.

Identifying Important Messages

Within your event logs, you should pay special attention to and immediately address any messages regarding queue size.

For example, messages such as “Sensor Info 1012: Adobe disk queue is #% full” need attention.

Responding to Sensor Event Messages

Tables that describe Sensor events and suggested actions for the supported web server platforms.

All Platforms

Event Message	Suggested Action
Sensor Info 1010: Sensor initialized.	No action required.
Sensor Info 1011: Sensor terminated. Total clicks queued ##	No action required.
Sensor Info 1012: Adobe disk queue is #% full	This message is logged each time the disk queue utilization crosses a 10% threshold. If this percentage continues to grow, action should be taken before the queue is full and data is lost. The most likely problem is that the Sensor has stopped communicating with the Insight Server. Contact Adobe ClientCare.
Sensor Info 1013: Sensor configuration changed	No action required. The Sensor detected a change in one of its configuration files and will reload.
Sensor Info 1014: Problem seeding random number generator	Contact Adobe ClientCare.
Sensor Info 1016: Config file file name loaded	No action required. The Sensor successfully loaded the configuration file listed.
Sensor Info 1017: Experiment file file name loaded	No action required. The Sensor successfully loaded the experiment file listed.
Sensor Error 3016: Unable to load config file /mypath/myfile	Confirm that the Sensor configuration file specified in the web server configuration exists and has the required permissions to be read by the web server process.
Sensor 3017: Unable to load controlled experiment configuration file /mypath/myfile	Confirm that the controlled experiment file specified in txlogd.conf exists and that the txlogd process has the necessary permissions to read the file.

Event Message	Suggested Action
Sensor Error 3018: Unable to parse content filter lists. Check txlogd configuration file	Verify the syntax of the ContentFilterInclude and ContentFilterExclude entries in txlogd.conf.
Sensor Error 3023: Failed to create lock (g_lockThrottle)	Contact Adobe ClientCare.
Sensor Error 3024: Failed to create lock (g_lockConfigCheck)	Contact Adobe ClientCare.
Sensor Error 4014: Could not open disk queue.	Verify QueueFile file name in txlogd.conf and if believed to be correct, contact Adobe ClientCare.
Sensor Error 4020: Not a queue file	Verify QueueFile file name in txlogd.conf and if believed to be correct, contact Adobe ClientCare.
Sensor Error 4021: Queue is full	Contact Adobe ClientCare.
Sensor Error 4022: Unable to map memory block of length <x> at offset <y>	Contact Adobe ClientCare.
Sensor Error 5012: Could not create mutex.	Contact Adobe ClientCare.
Sensor Error 5013: Could not acquire mutex.	Contact Adobe ClientCare.
Sensor Error 5030: Spawn fork error.	Contact Adobe ClientCare.
Sensor Error 5031: setsid failed.	Contact Adobe ClientCare.
Sensor Error 5032: Spawn fork error.	Contact Adobe ClientCare.
Sensor Error 5033: chdir failed.	Contact Adobe ClientCare.
Sensor Error 5034: Signal received	Program was likely terminated by an external signal. If the source of this signal cannot be determined, contact Adobe ClientCare.
Sensor Error 5035: Exit called from outside main	Contact Adobe ClientCare.
Sensor Error 5036: txlogd is already running	Another instance of the txlogd daemon is already running. Stop the other instance first if you want to run a new one.

Apache/IBM HTTP Server

Event Message	Suggested Action
Sensor Error 3015: The VisualSciencesConfig directive is missing from httpd.conf.	This is a configuration error. The VisualSciencesConfig directive must be in httpd.conf with a parameter that is the location of txlogd.conf.
Sensor Error 3019: vys-cookie was not called before vys-log. Please contact support.	Contact Adobe ClientCare.
Sensor Error 3025: Sub-request points back to itself	Contact Adobe ClientCare.

AOL Server

Event Message	Suggested Action
Sensor Error 3015: ns/server/[server]/module/[module] section is missing in AOLServer config file.	This is a configuration error. Correct as stated in error.

Event Message	Suggested Action
Sensor Error 3019: vys-cookie was not called before vys-log. Please contact support. Contact Adobe ClientCare.	Please contact support. Contact Adobe ClientCare.
Sensor Error 3020: VisualSciencesConfig missing as first entry in [section] section in AOLServer config file.	This is a configuration error. Correct as stated in error.
Sensor Error 3021: VisualSciencesConfig is missing a value in [section] section in AOLServer config file.	This is a configuration error. Correct as stated in error.

iPlanet and Java System Web Servers

Event Message	Suggested Action
Sensor Error 3011: Init directive required. Such as Init fn=vys-init config-file="/mypath/myfile"	This is a configuration error. The iPlanet init directive is missing.
Sensor Error 3015: config-file is not specified in iPlanet Init directive	This is a configuration error. The path to the configuration file was not supplied in the iPlanet Init directive.
Sensor Error 3019: vys-cookie was not called before vys-log. Please check configuration file	vys-cookie must be specified as the first NameTrans directive for each software virtual server.

Changing the Target Data Workbench Server

Navigation title: Changing the Target Insight Server

To change the data workbench server with which a **Sensor** communicates (the target server), you must edit the `txlogd.conf` file on each of the web servers on which **Sensor** is installed.

To change to target data workbench server

1. Stop both the original target server and the new target server.
2. Copy the most current `.vsl` file from the original target server to the new target server. When you restart the new target server in a later step, this causes all new **Sensor** data to be appended to the current, existing `.vsl` file instead of creating a new `.vsl` file. To do so, complete the following steps:
 1. On the original target server, browse to the `\Logs` folder in the data workbench server installation directory.
 2. Copy the most current `.vsl` file in the folder.
 3. On the new target server, browse to the `\Logs` folder in the data workbench server installation directory and paste the `.vsl` file to this folder.
3. On one of the web servers on which **Sensor** is installed, open and edit the `txlogd.conf` file. To do so, complete the following steps:
 1. Browse to the **Sensor** installation directory and open the `txlogd.conf` file in a text editor.
 2. Locate the `ServerAddress` parameter and change it to reflect the address of the new target server.
 3. Save and close the file.
4. Repeat steps 2–3 on all the remaining web servers on which **Sensor** is installed.
5. Restart the original target server (if it is still to be used) and the new target server.
6. Data will begin transmitting to the new target server that you have specified.

Dealing with Offline Sensors

Navigation title:Dealing with Offline Sensors

Plan for two scenarios with regard to the **Sensor** going offline.

Recommended Frequency: As needed

These scenarios are the web server being taken out of rotation from other servers, and the web server going down entirely because of a failure.

The first concept to understand is the “As Of” time in the data workbench server.

Understanding "As Of" Time

Navigation title:Understanding "As Of" Time

A data workbench server becomes aware of a source of data, such as a **Sensor**, when it receives data from that source.

The As Of time is a guarantee that the data workbench server has data for all of the data sources of which it is aware.

Let’s say that we have a set of three **Sensors** that send data to a data workbench server: WEB1, WEB2, and WEB3. As the data workbench server receives and processes the data from these **Sensors**, it automatically comes to expect data from each of these sources. The As Of time indicates the last time that the data workbench server received data from all three of these sources.

In practical terms, the data workbench server only cares about the As Of time and not what might be referred to as “wall time,” or the time from a clock on the wall. The data workbench server knows the time only as the As Of time. This is particularly important for reporting purposes as it guarantees that reports always run based on the As Of time, which ensures that reports with only partial data can never be sent to end users of the system.

The data workbench server uses data that is sent from the transmitter to provide the As Of time, whether it be actual data collected from the website or periodic heartbeats sent by your **Sensors**. These heartbeats serve two purposes:

1. To keep an HTTP/1.1 persistent connection open between the **Sensor** and the data workbench server.
2. To keep the As Of time current in the event that website traffic is not being collected and sent to the data workbench server.

Understanding the Causes

Navigation title:Understanding the Causes

Information about resolving web server problems, such as, if the web server is taken out of rotation, or if the web server fails.

When a Web Server is Taken Out of Rotation

When a web server is taken out of rotation from a pool of servers, but is otherwise connected with the transmitter sending periodic heartbeats, the As Of time is kept current and no intervention on anyone’s part is required.

When a Web Server Fails

When a web server is completely offline due to some catastrophic failure, or is not sending data or heartbeats, the As Of time on the data workbench server stops to guarantee that it represents the last time the data workbench server received data from ALL of the data sources of which it is aware. The system itself continues to process data, which is still available for analysis in Data Workbench, but anything in the data workbench server that is based on the As Of time does not function. For example, the As Of time triggers reporting and is used to create many derived dimensions in the system. When the As Of time has stopped, reporting is not triggered and these particular derived dimensions are not available.

For example, if WEB2 went offline on June 15 and did not send any data for five days, the As Of time would be sometime on June 15. The dimension of Yesterday, for example, would be June 14 even though today's date is June 20.

Solving the Problem

Navigation title:Solving the Problem

When a web server goes offline because of a failure, the solution is a simple one that requires a Data Workbench user with appropriate privileges to open the `Log Processing Mode.cfg` file and add the ID of the **Sensor** (in our example, WEB2) to the "Offline Sources" section.

This section of the file tells the data workbench server that it should no longer expect any data from this source because it is, in fact, offline.



Note: This change does not need to be performed by an Adobe Consultant. Anyone who has appropriate privileges to open the `Log Processing Mode.cfg` file can make this change.

If WEB2 begins to send data again, the data workbench server brings the source back online and adjusts the As Of time to reflect the last time it received data from all of the sources of which it is aware. In other words, new data coming into the system takes precedence over what is written in the `Log Processing Mode.cfg` file.

If WEB2 goes offline again, the As Of time will again stop, and you will need to edit the `Log Processing Mode.cfg` file again even though it might already have WEB2 listed as an offline source. This is an artifact of the design of the product in keeping with the definition of the As Of time: the last time the system has data for all known sources.

When you add more web servers (WEB4, WEB5, WEB6), and they begin sending data to the data workbench server, you do not need to do anything to have the data workbench server recognize the new sources. The system simply becomes aware that it should be expecting data from these new sources, as described above.

Working with Multiple Instances of a Web Server

Navigation title:Working with Multiple Instances of a Web Server

Information about the general configuration of the **Sensor** with one web server instance running on a web server.

Collector

Transmitter

- Listens in on the request and response at the web server level.
- Attaches itself to the web server process.
- On IIS, this is an ISAPI Filter
- Apache, a shared object
- J2EE, a Web App
- Sets v1st cookie (trackingid)

- Sends the data to the Visual Server
- Uses either port 443 or port 80 (by default)
- Can be configured to use other ports

QUEUE
WEB1

Visual Server

Compressed

In this scenario, a single web server instance is writing data to the memory mapped queue file, which is read by the transmitter and sent to the data workbench server.

When **Sensor** is installed on a web server that is running multiple collector instances, you can configure it one of two ways:

- You can have all of the collector modules share one queue file.

When using a single queue file, the management, configuration, and administration is somewhat simplified because the architecture itself is less complex. However, with a single queue file, the entire web server, regardless of the number of instances, is identified as WEB1.

- You can replicate the above architecture multiple times and have each web server instance have a separate queue file.

This enables you to identify each of the web server instances uniquely. In other words, the identification of the web server (and the corresponding SensorID in the **Sensor** configuration) is a function of this configuration.

In any case, the data still has all of the host name information so that you can distinguish between www.client.com, www2.client.com, and so forth. The correct configuration is determined by the analysis goals and whether the analysts need to segment the data based on a specific instance running on a web server.



Note: This type of segmentation is typically used only in operational analysis and does not provide much practical use outside of that area.

Sensor Txlogd.conf File Parameters

Information about the required and optional Sensor txlogd.conf file parameters.



Note: All of the parameters and values in the Txlogd.conf file are case insensitive.

Required Parameters

Navigation title: Required Parameters

Information about the required **Sensor** txlogd.conf parameters.

Required Txlogd.conf Parameters

In this Parameter...	Specify...
SensorID	<p>A character string that uniquely identifies this Sensor.</p> <p>Sensor attaches the SensorID to each event record it sends to the data workbench server. The SensorID enables the event data from this web server to be distinguished from the event data captured by other Sensors.</p> <p>Although a SensorID can consist of any string of characters, by convention, the name of the web server whose events the Sensor is capturing is used. Using the server name as the SensorID makes it easy to determine the source of an event during the analysis stage. It also ensures that the SensorID is unique within the implementation.</p> <p>Example: <code>SensorID web001a</code></p>
ServerAddress	<p>The address of the data workbench server to which this Sensor sends event data.</p> <p> Note:</p> <p>When working in a clustered environment, Sensor should be configured to access the master data workbench server to avoid synchronization issues. In Data Workbench you can view information about the processing data workbench servers in your cluster using the Related Servers menu item in the Servers Manager. For more information about the Servers Manager, see the <i>Data Workbench Sensor Guide</i>.</p> <p>If your web server can resolve server names through DNS, you can specify the server's address by name. If not, you must specify the server's numeric IP address.</p> <p>Example: <code>ServerAddress 10.1.0.7</code> or <code>ServerAddress vserver01.mycompany.com</code></p>
SSL	<p>Whether Sensor communicates with data workbench server using HTTP or HTTPS. Set to "on" for HTTPS or "off" for HTTP.</p> <p>Example: <code>SSL on</code></p>
ServerPort	<p>The port on which the data workbench server listens for event data.</p> <p>Example: <code>ServerPort 443</code></p>
CertName	<p>Required only if the SSL parameter is set to "on."</p> <p>The common name of the data workbench server to which this Sensor sends event data.</p> <p>The value you specify must exactly match the common name that appears on the data workbench server license certificate.</p> <p>Example: <code>CertName vserver01.mycompany.com</code></p>
CertPath	<p>Required only if the SSL parameter is set to "on."</p> <p>The directory in which the certificate authority (<code>trust_ca_cert.pem</code>) file is located</p> <p>Examples: <code>CertPath /usr/local/visualsensor</code></p>

In this Parameter...	Specify...
	CertPath C:\VisualSensor
QueueFile	<p>Not needed for Sensor installations on Microsoft Windows 2000 or 2003 Server machines running Internet Information Service (IIS) versions 5.x or 6.x.</p> <p>The fully qualified name of the disk queue file.</p> <p>Although you can assign any name to this file, by convention, the queue file is named VisualSensor.dat.</p> <p>For Sensor installations on Unix, you can place this file anywhere. On Windows running a Java web server, you must place this file in the same directory as the transmitter. For all other web servers, this file should reside in the /var/queue directory.</p> <p>Example: QueueFile /var/queue/VisualSensor.dat</p> <p> Note: Make sure that the device to which you assign this file has enough free space to accommodate a queue of the size you need.</p>
QueueSize	<p>An integer that represents the size of the disk queue file in MB.</p> <p>For Sensor installations on Microsoft Windows, the queue file itself is created in the same directory as the transmitter and is named Diskq2000.log.</p> <p>The following example sets the queue size to 200 MB:</p> <p>QueueSize 200</p> <p>The following example sets the queue size to 2 GB:</p> <p>QueueSize 2000</p>

Optional Parameters

Navigation title:Optional Parameters

Information about optional **Sensor** txlogd.conf file parameters.

Optional Txlogd.conf Parameters for All Web Servers

Parameter	Description
AddressFilter	<p>Enables you to filter specified IP addresses.</p> <p>When you filter a particular address, a “packet” is not logged. This feature eliminates internal or monitored agents before log processing, thereby increasing the speed of log processing and reducing data storage requirements. You can use wildcards when specifying addresses.</p> <p>Example: AddressFilter 10.0.0.000</p>
ContentFilterInclude ContentFilterExclude	<p>Specify whether to include or exclude certain types of content from logging.</p> <p>The parameter values are prefix-matched against the content type of the reply.</p> <p>For example, “image/” matches all image content types while “image/gif” matches only that type exactly. When multiple matches occur for a given content-type, the most specific match is used. Therefore, you could put “image/gif” in the ContentFilterInclude parameter and “image/” in the ContentFilterExclude parameter and “image/gif” replies are allowed but all other image types are filtered out.</p> <p>Example: ContentFilterInclude *</p> <p>Example: ContentFilterExclude image/,text/css,application/x-javascript</p>
DebugLogPath	<p>Set this parameter only when working with Adobe Consulting Services.</p> <p>Enables debug logging for the web module and transmitter.</p>

Parameter	Description
	<p>You use this parameter when the Sensor is not working correctly. After this parameter is set, you must create an empty file at the specified path location and give “write” rights to it for all users. For example (inside a unix shell on the web server):</p> <ul style="list-style-type: none"> • % cd /var/log • % touch vslog.txt • % chmod a+w vslog.txt <p>You should enable debug logging for only a short period of time, after which the log file should be sent to Adobe Consulting Services to be analyzed.</p> <p>Example: <code>DebugLogPath /var/log/vslog.txt</code></p> <p>Adobe recommends that this parameter first be set in a test environment to determine the effect on your system.</p>
DisableField	<p>Disables the specified field</p> <p>Users can eliminate fields they do not use or do not want to store. If the field takes string values, disabling it passes a blank string. If the field takes numeric values, disabling it passes the number zero (0). You can disable the following fields:</p> <ul style="list-style-type: none"> • sc-status • x-new-visitor • x-trackingid • sc-bytes • c-ip • cs-method • cs-uri-stem • cs-uri-query • s-dns • cs(referrer) • cs(cookie) • cs(user-agent) • cs(useragent) • sc(content-type) • x-experiment <p>Example: <code>DisableField x-trackingid</code></p>
ExpFile	<p>Path to the Controlled Experiment configuration file.</p> <p>Example: <code>ExpFile C:\VisualSensor\experiment.txt</code></p>
ExpCookieURL	<p>Resource that, when requested, causes a new tracking ID to be generated and the user to be placed in an experiment group.</p> <p> Note: This resource does not have to physically exist on the web server.</p> <p>Example: <code>ExpCookieURL /setcookie.htm</code></p>
ExpPartialMatch	<p>If you want to enable your controlled experiments to remap your entire site or an entire subdirectory of your site to another location, set this parameter to “on.” The default is “off.”</p> <p>Example: <code>ExpPartialMatch off</code></p> <p> Note: Be very careful when setting this parameter to “on.”</p>
LogAllNewUsers	Determines whether every new user’s first click is logged even if the user requests a document type that is filtered out by the ContentFilterExclude parameter.

Parameter	Description
	<p>The default is “no.”</p> <p>Typically image files are filtered out by the ContentFilterExclude parameter. If LogAllNewUsers is set to “yes” and the very first document a new user gets from the server is an image, that request is logged. If the LogAllNewUsers parameter is either set to “no” or not set at all (and assuming that images are filtered out by the ContentFilterExclude parameter) and the very first document a new user gets from the server is an image, that request is not logged.</p> <p>Example: <code>LogAllNewUsers no</code></p>
MaxPageLoadTime	<p>The amount of time in seconds that the transmitter waits to send the next batch of packets.</p> <p>The default is 15.</p> <p>Example: <code>MaxPageLoadTime 15</code></p> <p> Note: Do not change this parameter value without first contacting Adobe Consulting Services.</p>
PrivacyID	<p>Enables you to disable visitor tracking, which can be used to comply with opt-out policies.</p> <p>When enabled, Sensor does not log a “packet” for any visitor whose V1st cookie is set to the PrivacyID specified. Because no information is logged for those visitors, no information about those visitors is sent to the data workbench server for processing.</p> <p>To enable this feature, you must complete the following steps:</p> <ol style="list-style-type: none"> 1. PrivacyID must be defined with a value of 0 (zero) in the <code>txlogd.conf</code> file for the Sensor. <p>Example: <code>PrivacyID 0</code></p> <ol style="list-style-type: none"> 2. Website owners must write code to set visitor (V1st) cookies such that the cookie ID value matches the PrivacyID value defined “<code>txlogd.conf</code>.”
SiteTest	<p>Location to which the transmitter (txlogd) periodically sends requests to see if the website is operating correctly.</p> <p>Note that the location is specified in the following format, not as a URL: <code>http,serverAddress,port,/resource</code></p> <p>where <i>serverAddress</i> is server name or IP address, <i>port</i> is server’s HTTP listening port, and <i>resource</i> is the specific resource to request (can include a query string).</p> <p>You can specify multiple SiteTest lines.</p> <p>Example: <code>SiteTest http,localhost,80,/test.html</code></p> <p> Note: Only http is supported at this time.</p>
TrackingCookie	<p>The name of the cookie set on the visitor’s browser.</p> <p>The default is “v1st.”</p> <p>Example: <code>TrackingCookie v1st</code></p>
VerifyCertName	<p>Indicates whether to validate the server against CertName parameter</p> <p>The default is “on.”</p> <p>Example: <code>VerifyCertName on</code></p>

Optional Txlogd.conf Parameters for IIS Web Servers

Parameter	Description
IISCaptureBytesSent	<p>Set this parameter only when working with Adobe Consulting Services.</p> <p>Tells the IIS Sensor which of two possible logging “hooks” should be used to log a packet.</p> <p>Use this parameter when the IIS Sensor is not logging packets correctly. This parameter would be set to “off” if the default logging hook was not working correctly. The default is “on.”</p> <p>Example: <code>IISCaptureBytesSent on</code></p>
IISUseAlternateHandler	<p>Set this parameter only when working with Adobe Consulting Services.</p> <p>Tells the Sensor which of two possible “hooks” should be used to set the v1st cookie.</p> <p>You use this parameter when the IIS Sensor is not setting the v1st cookie correctly. This parameter would be set to “yes” if the default hook was not correctly setting the v1st cookie. The default is “no.”</p> <p>Example: <code>IISUseAlternateHandler no</code></p>
NewUserCacheControl CacheControl	<p>By default, Sensor sends cache control response headers on each request. When the cache control feature is enabled Sensor sends an Expires header with a value of Thu, 01 Dec 1994 16:00:00 GMT to the browser.</p> <p>You can modify the response strings as desired by editing the following two lines in the txlogd.conf file:</p> <pre>NewUserCacheControl <string1> CacheControl <string2></pre> <p>Example:</p> <pre>NewUserCacheControl no-cache=Set-Cookie CacheControl private,max-age=0,must-revalidate</pre> <p>To disable the sending of cache control response headers, type a hyphen for each line as shown below:</p> <pre>NewUserCacheControl - CacheControl -</pre>

Optional Txlogd.conf Parameters for Apache 1.3, Apache 2, and Apache 2.2 Web Servers

In this Parameter...	Specify...
ApacheUseAlternateHandler	<p>Set this parameter only when working with Adobe Consulting Services.</p> <p>Tells the Sensor which of two possible “hooks” should be used to set the v1st cookie.</p> <p>You use this parameter when the Apache Sensor is not setting the v1st cookie correctly. This parameter would be set to “yes” if the default hook was not correctly setting the v1st cookie. The default is “no.”</p> <p>Example: <code>ApacheUseAlternateHandler no</code></p>
ApacheUseBothHandlers	<p>Set this parameter only when working with Adobe Consulting Services.</p> <p>Instructs the Sensor to try setting the v1st cookie in both hooks.</p> <p>You use this parameter when the Apache Sensor is not setting the v1st cookie correctly. The default is “yes.”</p> <p>If set to “yes” and the v1st cookie is not properly set in the first hook, the second hook is used. If set to “no,” you would set the ApacheUseAlternateHandler parameter to indicate which hook to use to set the v1st cookie.</p>

In this Parameter...	Specify...
	<p>Example: <code>ApacheUseBothHandlers yes</code></p>
NewUserCacheControl CacheControl	<p>By default, Sensor sends cache control response headers on each request. When the cache control feature is enabled Sensor sends an Expires header with a value of Thu, 01 Dec 1994 16:00:00 GMT to the browser.</p> <p>You can modify the response strings as desired by editing the following two lines in the <code>txlogd.conf</code> file:</p> <pre>NewUserCacheControl <string1> CacheControl <string2></pre> <p>Example:</p> <pre>NewUserCacheControl no-cache=Set-Cookie CacheControl private,max-age=0,must-revalidate</pre> <p>To disable the sending of cache control response headers, type a hyphen for each line as shown below:</p> <pre>NewUserCacheControl - CacheControl -</pre> <p> Note: Adobe recommends that you do not disable this feature.</p>

Optional Txlogd.conf Parameters for IHS 2 Web Servers

In this Parameter...	Specify...
ApacheUseAlternateHandler	<p>Set this parameter only when working with Adobe Consulting Services.</p> <p>Tells the Sensor which of two possible “hooks” should be used to set the v1st cookie.</p> <p>You use this parameter when the Apache Sensor is not setting the v1st cookie correctly. This parameter would be set to “yes” if the default hook was not correctly setting the v1st cookie. The default is “no.”</p> <p>Example: <code>ApacheUseAlternateHandler no</code></p>
ApacheUseBothHandlers	<p>Set this parameter only when working with Adobe Consulting Services.</p> <p>Instructs the Sensor to try setting the v1st cookie in both hooks.</p> <p>You use this parameter when the Apache Sensor is not setting the v1st cookie correctly. The default is “yes.”</p> <p>If set to “yes” and the v1st cookie is not properly set in the first hook, the second hook is used. If set to “no,” you would set the <code>ApacheUseAlternateHandler</code> parameter to indicate which hook to use to set the v1st cookie.</p> <p>Example: <code>ApacheUseBothHandlers yes</code></p>

Event Data Record Fields

Navigation title: Event Data Record Fields

Log (.vs1) files contain the fields of event data that are collected from servers by **Sensors** and used by the data workbench server in the dataset construction process.

The names of the fields generally follow the naming convention for the W3C extended log file format. Many of the fields have prefixes that indicate the source of the information contained in the field:

- “cs” indicates communication from the client to the server
- “sc” indicates communication from the server to the client
- “s” indicates information from the server
- “c” indicates information from the client
- “x” indicates information that is created by an Adobe product

Baseline Event Data Record Fields

Navigation title:Baseline Event Data Record Fields

Information about baseline event data record fields as recorded by **Sensor**.

Field	Description
c-ip	The IP address of the client as included in the request made to the server. Example: 207.68.146.68
cs(cookie)	The cookies sent by the client with the request. Example: v1st=42FDF66DE610CF36; ASPSESSIONIDQCATDAQC=GPIBKEIBFBFIPL0JMKCAAEPM;
cs(referrer)	The HTTP referrer string sent by the client to the server with the request. Example: http://www.mysite.net/cgi-bin/websearch?qry If you are using page tags, the cs(referrer) is the full URL of the document containing the tag image, including HTTP or HTTPS. Also, you can configure Apache (1.3, 2.0, and 2.2) and IIS sensors to capture the port that is used for the request, which can identify HTTP vs. HTTPS requests.
cs(user-agent)	The string sent by the client with its request to the server that indicates what type of user agent the client is. Example: Mozilla/5.0 (Windows; U; Windows NT 5.1; en-US; rv:1.7) Gecko/20040707 Firefox/0.9.2
cs-method	The method type of the HTTP request Example: GET Reference: http://www.w3.org/TR/2000/NOTE-shoplogfileformat-20001115/#field_method
cs-uri-query	The query string portion of URI (stem + query string = URI) This is preceded by a question mark (?) and may contain one or more name-value pairs separated by ampersands (&). Example: page=homepage
cs-uri-stem	The stem portion of URI (stem + query string = URI) The stem is the actual or logical path to the requested resource on the server. Example: /index.asp
sc(content-type)	The content type of the resource being requested by the client as reported by the server. Examples: text/html, image/png, image/gif, video/mpeg
sc-bytes	The number of bytes of data sent from the server to the client in response to the request. Example: 4996
sc-status	The status code returned to the client by the server. Example: 200 Reference: http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html

Field	Description
s-dns	The fully qualified domain name or IP address of the host of the requested resource. Example: www.omniture.com
x-experiment	The list of all the controlled experiment names and groups that the client is a member of at the time of the request. Example: Home_Exp.Group_1,Registration_Exp.Group_2
x-timestamp	The date and time (GMT) at which the request was received by the server. The time is expressed as the number of 100 nanoseconds since January 1, 1600. Example: 12771098932000000 would be the x-timestamp value for 11:28:52.000000 on Tuesday, September 13, 2005.
x-trackingid	The 64-bit, hexadecimal value of the unique browser identifier found in a persistent cookie as set by a Sensor and provided by the client with a request to a server. Example: 42FDF66DE610CF36

The data workbench server can derive a number of variables from the baseline event data record fields. For more information, see the *Dataset Configuration Guide*.

Extensible Fields

Navigation title:Extensible Fields

Sensor, when used on a server, can collect fields of event data from any valid HTTP request or response header or variable available to it through the server's API.

To collect such fields of data, you must specify the desired header fields or variables in the `txlogd.conf` configuration file for **Sensor**.

- [Request Headers](#) on page 953
- [Server Variables](#) on page 953

Request Headers

Following is the syntax for specifying a request header field to be collected (for example, Host, Accept-Encoding, Keep-Alive, and so on) in `txlogd.conf`:

```
LogHeader RequestHeaderName
```

The collected data is recorded by **Sensor** to a field named “cs(RequestHeaderName)” in the `.vs1` files created by the data workbench server. For example, to collect the specific request header value from the request header “Host,” you would type “LogHeader Host” in `txlogd.conf`. The data is recorded to the field “cs(Host)” in the event data record.

Server Variables

Sensor can collect fields of data from response headers or API-accessible server variables using `SpecialLogFile` entries that you include in the `txlogd.conf` file. You can also use “`SpecialLogFile`” entries in addition to or instead of “`LogHeader`” entries to collect request headers. See [Request Headers](#) on page 953. The request headers option remains available for backward compatibility.

Following is the syntax for specifying a “`SpecialLogFile`” in `txlogd.conf`:

```
SpecialLogFile cs(log field) = serverVariable stage
```

The following table includes descriptions of the components of a “`SpecialLogFile`” entry.

SpecialLogField Entries

Component	Description
cs(log field)	The name of the field to which the collected data is recorded in the event data record and the .vsl files created by the data workbench server.
serverVariable	Any server variable that is available to Sensor through the server's API Example: response.p3p
stage	Either vys_log or vys_cookie Specifying the stage requires that you know which server variables are available for vys_log and vys_cookie. Example: For the serverVariable response.p3p, you would enter vys_log.

For help configuring **Sensor** to collect extensible event data record fields, contact Adobe Consulting Services.

Sample Configuration Files

Sample files for the configuration files.

Magnus.conf File Sample

Example of a complete magnus.conf file for Netscape Enterprise Server, iPlanet, Sun ONE, and Sun Java System web server installation and configuration.

```
#ServerRoot /usr/iplanet/servers/https-nova.visualsciences.net
ServerID https-nova.visualsciences.net
ServerName nova.visualsciences.net
ErrorLog /usr/iplanet/servers/https-nova.visualsciences.net/logs/errors
PidLog /usr/iplanet/servers/https-nova.visualsciences.net/logs/pid
User nobody
MtaHost localhost
DNS off
Security off
ClientLanguage en
AdminLanguage en
DefaultLanguage en
RqThrottle 128
StackSize 131072
CGIWaitPid on
TempDir /tmp/https-nova.visualsciences.net-ab0d7966
Init fn=flex-init access="$accesslog" format.access="%Ses->client.ip% - %Req->vars.auth-user% [%SYSDATE%] \"%Req->reqpb.clf-request%\" %Req->srvhdrs.clf-status% %Req->srvhdrs.content-length%"
Init fn=load-types mime-types=mime.types
Init fn="load-modules"
shlib="/usr/iplanet/servers/bin/https/lib/libNSServletPlugin.so"
func="NSServletEarlyInit,NSServletLateInit,NSServletNameTrans,NSServletService"
shlib_flags="(global|now)"
Init fn="NSServletEarlyInit" EarlyInit=yes
Init fn="NSServletLateInit" LateInit=yes
Init fn="load-modules"
shlib="/usr/iplanet/servers/visual_sciences/saf_visual_sciences.so"
func="vys-cookie,vys-log,vys-init,vys-content-type"
Init fn="vys-init" config-file="/etc/txlogd.conf"
```

Obj.conf File Sample

Example of a complete obj.conf file for Netscape Enterprise Server, iPlanet, Sun ONE, and Sun Java System web server installation and configuration.

```
<Object name=default>
  NameTrans fn="vys-cookie"
  NameTrans fn="NSServletNameTrans" name="servlet"
  NameTrans fn="pxf2dir" from="/servlet"
  dir="/usr/iplanet/servers/docs/servlet" name="ServletByExt"
  NameTrans fn=pxf2dir from=/mc-icons dir=/usr/iplanet/servers/ns-icons"
  name="es-internal"
  NameTrans fn="pxf2dir" from="/manual"
  dir="/usr/iplanet/servers/manual/https" name="es-internal"
  NameTrans fn=document-root root="$docroot"
  PathCheck fn=unix-uri-clean
  PathCheck fn="check-acl" acl="default"
  PathCheck fn=find-pathinfo
  PathCheck fn=find-index index-names="index.html,home.html"
  ObjectType fn=type-by-extension
  ObjectType fn=force-type type=text/plain
  ObjectType fn="vys-content-type"
  Service type="magnus-internal/jsp" fn="NSServletService"
  Service method=(GET|HEAD) type=magnus-internal/imagemap fn=imagemap
  Service method=(GET|HEAD) type=magnus-internal/directory fn=index-common

  Service method=(GET|HEAD|POST) type=*~magnus-internal/* fn=send-file
  AddLog fn=flex-log name="access"
  AddLog fn="vys-log"
</Object>

<Object name=cgi>
  ObjectType fn=force-type type=magnus-internal/cgi
  Service fn=send-cgi user="$user" group="$group" chroot="$chroot" dir="$dir"
  nice="$nice"
</Object>

<Object name="servlet">
  ObjectType fn=force-type type=text/html
  Service fn="NSServletService"
</Object>

<Object name="jsp092">
  ObjectType fn="type-by-extension"
  ObjectType fn="change-type" type="magnus-internal/jsp092"
  if-type="magnus-internal/jsp"
  Service fn="NSServletService" type="magnus-internal/jsp092"
</Object>

<Object name="ServletByExt">
  ObjectType fn=force-type type=magnus-internal/servlet
  Service type="magnus-internal/servlet" fn="NSServletService"
</Object>

<Object name="es-internal">
  PathCheck fn="check-acl" acl="es-internal"
</Object>
```

Txlogd.conf File Samples

Examples of txlogd.conf files for the supported web server/OS combinations.

IIS on Microsoft Windows Server 2000 or Later

Sample of the txlogd.conf file required for Sensor for Internet Information Service (IIS) 5.x or 6.x running under Microsoft Windows Server 2000 or later.

```
# sample txlogd configuration file for IIS

# Unique identifier associated with this sensor, sent to the Server
SensorID SENSOR1

# Address of the Server.
ServerAddress 10.1.0.7

# Encrypt Sensor to Server communication. Defaults to 'on'.
SSL on

# Port used to connect to the Server. Defaults to 443 or 80, depending on
SSL being
# 'on' or 'off'.
ServerPort 443

# The common name (CN) of the Server as set in the Server's certificate.
CertName vs1.domain.com

# Queue file size in MB.
QueueSize 2000

# Path to directory that contains .pem certificate files. This parameter
must point to the directory in
# which the Sensor's trusted_ca_cert.pem file resides.
CertPath C:\VisualSensor\

# The amount of time in seconds the transmitter will wait to send the next
batch of packets.
# Defaults to 15. Do not change without consulting Adobe.
MaxPageLoadTime 15

# The name of the cookie set on the visitor's browser. Defaults to 'vlst'.
TrackingCookie vlst

# Whether or not to validate the server against CertName. Defaults to 'on'.

#VerifyCertName on

# Include or exclude certain content types from logging. These will be
prefix matched against
# the content type of the reply.
# For example, "image/" will match all image content types while "image/gif"
# will match only that type
# exactly. When there are multiple matches for a given content-type, the
most specific
# match will be used. Therefore, you could put "image/gif" in the
ContentFilterInclude and "image/" in
# the ContentFilterExclude and image/gif replies will be allowed but all
other image types will be
# filtered out.
ContentFilterInclude *
ContentFilterExclude image/,text/css,application/x-javascript

# For IIS only. Used to work around certain problems. Defaults to 'off'.
# Only set to 'on' on advice from tech support.
```

```
IISCaptureBytesSent off

# IIS only. Used to work around certain problems. Defaults to 'no'.
# Only set to 'yes' on advice from tech support.
IISUseAlternateHandler no

# Path to the Controlled Experiment configuration file.
ExpFile C:\VisualSensor\experiment.txt

# Resource that, when requested, will cause a new tracking ID to be generated
# and the user to be placed
# an experiment group. Note that this resource does not have to physically
# exist on the Web server.
ExpCookieURL /setcookie.htm

# If set to 'on', URLs visited need only match part of the URL listed in
# the experiment file to be
# considered a match. Defaults to 'off'.
ExpPartialMatch on

# The transmitter (txlogd) will periodically send requests to this location
# to see if the site is
# operating correctly. Note that the location is specified in the following
# format, not as a URL:
#     http, <serverAddress>, <port>, /<resource>
# where: <serverAddress> is server name or IP address, <port> is server's
# HTTP listening port,
# <resource> is the specific resource to request (can include a query
# string).
# Only http is supported at this time. Multiple SiteTest lines can be
# specified.
SiteTest http,localhost,80,/test.html

# Additional HTTP headers to log. Fields will be saved as cs(header-name)
# in Server. Any valid
# HTTP header can be used (this is not a complete list).
LogHeader Accept
LogHeader Accept-Encoding
LogHeader Connection
LogHeader Host
LogHeader Keep-Alive
LogHeader Referrer

AddressFilter 10.2.1.89
# test machine
AddressFilter 192.168.*./*
# internal IP

DisableField s-dns
DisableField cs(cookie)

PrivacyID 0x11111111

# Send this Cache-Control response header to all new visitors
NewUserCacheControl no-cache=Set-Cookie

# Send this Cache-Control response to returning visitors
CacheControl private,max-age=0,must-revalidate
```

Lotus Sametime on Windows Server 2000 or Later

Sample of the txlogd.conf file required for Sensor for Lotus Sametime running under Microsoft Windows Server 2000 or later.

```
# sample txlogd configuration file for Lotus Sametime

# Unique identifier associated with this sensor, sent to the Server.
SensorID SENSOR1

# Address of the Server.
ServerAddress 10.1.0.7

# Encrypt Sensor to Server communication. Defaults to 'on'.
SSL on

# Port used to connect to the Server. Defaults to 443 or 80, depending on
SSL being 'on'
# or 'off'.
ServerPort 443

# The common name (CN) of the Server as set in the Server's certificate.
CertName vsl.domain.com

# Path to directory that contains .pem certificate files.
CertPath C:\VisualSensor

# Full path to the Sensor disk queue on the web server.
# The queue file must be in the same directory as the Transmitter.
QueueFile C:\VisualSensor\diskq2000.log

# Queue file size in MB.
QueueSize 2000

# The amount of time in seconds the transmitter will wait to send the next
batch of packets.
# Defaults to 15.
#MaxPageLoadTime 15

# The name of the cookie set on the visitor's browser. Defaults to 'vlst'.
TrackingCookie vlst

# Whether or not to validate the server against CertName. Defaults to 'on'.
#VerifyCertName on

# Include or exclude certain content types from logging. These will be
prefix matched against the
# content type of the reply. For example, "image/" will match all image
content types while "image/gif"
# will match only that type exactly. When there are multiple matches for a
given content-type, the most
# specific match will be used. Therefore, you could put "image/gif" in the
ContentFilterInclude and
# "image/" in the ContentFilterExclude and image/gif replies will be allowed
but all other image types
# will be filtered out.
ContentFilterInclude *
ContentFilterExclude image/,text/css,application/x-javascript

# Used to work around certain problems. Defaults to 'no'. Only set to 'yes'
on advice from tech
```

```

# support. Apache 1.3, 2.x, IHS 1.3, and IHS 2.x only.
ApacheUseAlternateHandler no

# Path to the Controlled Experiment configuration file.
ExpFile C:\VisualSensor\experiment.txt

# If the user visits this page then a new tracking ID is generated and the
# user is placed in an
# experiment group.
ExpCookieURL /setcookie.html

# If set to 'on', URLs visited need only match part of the URL listed in
# the experiment file to be
# considered a match. Defaults to 'off'.
ExpPartialMatch on

# The transmitter (txlogd) will periodically send requests to this location
# to see if the site is
# operating correctly. Note that the location is specified in the following
# format, not as a URL:
#     http, <serverAddress>, <port>, /<resource>
# where: <serverAddress> is server name or IP address, <port> is server's
# HTTP listening port,
# <resource> is the specific resource to request (can include a query
# string).
# Only http is supported at this time. Multiple SiteTest lines can be
# specified.
SiteTest http,localhost,80,/test.html

```

Lotus Domino on Windows Server 2000 or Later

Sample of the txlogd.conf file required for Sensor for Lotus Domino running under Microsoft Windows Server 2000 or later.

```

# sample txlogd configuration file for Lotus Domino

# Unique identifier associated with this sensor, sent to the Server.
SensorID SENSOR1

# Address of the Server.
ServerAddress 10.1.0.7

# Encrypt Sensor to Server communication. Defaults to 'on'.
SSL on

# Port used to connect to the Server. Defaults to 443 or 80, depending on
# SSL being 'on'
# or 'off'.
ServerPort 443

# The common name (CN) of the Server as set in the Server's certificate.
CertName vs1.domain.com

# Path to directory that contains .pem certificate files.
CertPath C:\VisualSensor

# Full path to the Sensor disk queue on the web server.
# The queue file must be in the same directory as the Transmitter.
QueueFile C:\VisualSensor\diskq2000.log

# Queue file size in MB.

```

```
QueueSize 2000

# The amount of time in seconds the transmitter will wait to send the next
batch of packets.
# Defaults to 15.
#MaxPageLoadTime 15

# The name of the cookie set on the visitor's browser. Defaults to 'v1st'.
TrackingCookie v1st

# Whether or not to validate the server against CertName. Defaults to 'on'.
#VerifyCertName on

# Include or exclude certain content types from logging. These will be
prefix matched against the
# content type of the reply. For example, "image/" will match all image
content types while "image/gif"
# will match only that type exactly. When there are multiple matches for a
given content-type, the most
# specific match will be used. Therefore, you could put "image/gif" in the
ContentFilterInclude and
# "image/" in the ContentFilterExclude and image/gif replies will be allowed
but all other image types
# will be filtered out.
ContentFilterInclude *
ContentFilterExclude image/,text/css,application/x-javascript

# Used to work around certain problems. Defaults to 'no'. Only set to 'yes'
on advice from tech
# support. Apache 1.3, 2.x, IHS 1.3, and IHS 2.x only.
ApacheUseAlternateHandler no

# Path to the Controlled Experiment configuration file.
ExpFile C:\VisualSensor\experiment.txt

# If the user visits this page then a new tracking ID is generated and the
user is placed in an
# experiment group.
ExpCookieURL /setcookie.html

# If set to 'on', URLs visited need only match part of the URL listed in
the experiment file to be
# considered a match. Defaults to 'off'.
ExpPartialMatch on

# The transmitter (txlogd) will periodically send requests to this location
to see if the site is
# operating correctly. Note that the location is specified in the following
format, not as a URL:
#      http, <serverAddress>, <port>, /<resource>
# where: <serverAddress> is server name or IP address, <port> is server's
HTTP listening port,
# <resource> is the specific resource to request (can include a query
string).
# Only http is supported at this time. Multiple SiteTest lines can be
specified.
SiteTest http,localhost,80,/test.html

# Lotus Domino specific additional fields to log (complete list).
SpecialLogField cs(dn)=vars.X509SubjectName vys_log
SpecialLogField cs(cert)=vars.certificate vys_log
```

IBM HTTP Server on IBM AIX 5.2

Sample of the txlogd.conf file required for Sensor for IBM HTTP Server 2.0.x running under IBM AIX 5.2.

```
# sample txlogd configuration file for IHS 2.0.x

# Unique identifier associated with this sensor, sent to the Server.
SensorID SENSOR1

# Address of the Server.
ServerAddress 10.1.0.7

# Encrypt Sensor to Server communication. Defaults to 'on'.
SSL on

# Port used to connect to the Server. Defaults to 443 or 80, depending on
SSL being 'on'
# or 'off'.
ServerPort 443

# The common name (CN) of the Server as set in the Server's certificate.
CertName vsl.domain.com

# Path to directory that contains .pem certificate files.
CertPath /usr/local/visualsensor

# Full path to the Sensor disk queue on the web server.
QueueFile /var/queue/VisualSensor.dat

# Queue file size in MB.
QueueSize 2000

# The amount of time in seconds the transmitter will wait to send the next
batch of packets.
# Defaults to 15.
MaxPageLoadTime 15

# The name of the cookie set on the visitor's browser. Defaults to 'vlst'.
TrackingCookie vlst

# Whether or not to validate the server against CertName. Defaults to 'on'.
#VerifyCertName on

# Include or exclude certain content types from logging. These will be
prefix matched against the
# content type of the reply. For example, "image/" will match all image
content types while "image/gif"
# will match only that type exactly. When there are multiple matches for a
given content-type, the most
# specific match will be used. Therefore, you could put "image/gif" in the
ContentFilterInclude and
# "image/" in the ContentFilterExclude and image/gif replies will be allowed
but all other image types
# will be filtered out.
ContentFilterInclude *
ContentFilterExclude image/,text/css,application/x-javascript

# Used to work around certain problems. Defaults to 'no'. Only set to 'yes'
on advice from tech
# support. Apache 1.3, 2.x, IHS 1.3, and IHS 2.x only.
ApacheUseAlternateHandler no
```

```

# Path to the Controlled Experiment configuration file.
ExpFile /home/experiment.txt

# If the user visits this page then a new tracking ID is generated and the
# user is placed in an
# experiment group.
ExpCookieURL /setcookie.html

# If set to 'on', URLs visited need only match part of the URL listed in
the experiment file to be
# considered a match. Defaults to 'off'.
ExpPartialMatch on

# The transmitter (txlogd) will periodically send requests to this location
to see if the site is
# operating correctly. Note that the location is specified in the following
format, not as a URL:
#           http, <serverAddress>, <port>, /<resource>
# where: <serverAddress> is server name or IP address, <port> is server's
HTTP listening port,
# <resource> is the specific resource to request (can include a query
string).
# Only http is supported at this time. Multiple SiteTest lines can be
specified.
SiteTest http,localhost,80,/test.html

# Additional HTTP headers to log. Fields will be saved as cs(header-name)
in Server. Any valid
# HTTP header can be used (this is not a complete list).
LogHeader Accept
LogHeader Accept-Encoding
LogHeader Connection
LogHeader Host
LogHeader Keep-Alive
LogHeader Referrer

# IHS 2.0.x specific additional fields to log (complete list).
SpecialLogFile x = request.allowed vys_log
SpecialLogFile x = request.ap_auth_type vys_log
SpecialLogFile x = request.query vys_log
SpecialLogFile x = request.bytes_sent vys_log
SpecialLogFile x = request.canonical_filename vys_log
SpecialLogFile x = request.chunked vys_log
SpecialLogFile x = request.clength vys_log
SpecialLogFile x = request.content_encoding vys_log
SpecialLogFile x = request.content_type vys_log
SpecialLogFile x = request.eos_sent vys_log
SpecialLogFile x = request.expecting_100 vys_log
SpecialLogFile x = request.filename vys_log
SpecialLogFile x = request.handler vys_log
SpecialLogFile x = request.header_only vys_log
SpecialLogFile x = request.hostname vys_log
SpecialLogFile x = request.method vys_log
SpecialLogFile x = request.method_number vys_log
SpecialLogFile x = request.mtime vys_log
SpecialLogFile x = request.no_cache vys_log
SpecialLogFile x = request.no_local_copy vys_log
SpecialLogFile x = request.path_info vys_log
SpecialLogFile x = request.proto_num vys_log
SpecialLogFile x = request.protocol vys_log

```

```

SpecialLogField x = request.proxyreq vys_log
SpecialLogField x = request.range vys_log
SpecialLogField x = request.read_body vys_log
SpecialLogField x = request.read_chunked vys_log
SpecialLogField x = request.read_length vys_log
SpecialLogField x = request.remaining vys_log
SpecialLogField x = request.request_time vys_log
SpecialLogField x = request.sent_bodyct vys_log
SpecialLogField x = request.status vys_log
SpecialLogField x = request.status_line vys_log
SpecialLogField x = request.clf-request vys_log
SpecialLogField x = request.unparsed_uri vys_log
SpecialLogField x = request.uri vys_log
SpecialLogField x = request.used_path_info vys_log
SpecialLogField x = request.user vys_log
SpecialLogField x = request.vlist_validator vys_log
SpecialLogField x = headers.Refferrer vys_log
SpecialLogField x = headers.User-Agent vys_log
SpecialLogField x = headers.Cookie vys_log
SpecialLogField x = headers_out.Cookie vys_log

AddressFilter 10.2.1.89
# test machine
AddressFilter 192.168.*. *
# internal IP

DisableField s-dns
DisableField cs(cookie)

PrivacyID 0x11111111

```

Apache Server 1.3.x on RedHat Linux, Sun Solaris, FreeBSD, or Mac OSX

Sample of the txlogd.conf file required for Sensor for Apache Server 1.3.x running under RedHat Linux 7.x, Sun Solaris 2.6 or later, FreeBSD 4 or later, or Mac OS X PowerPC.

```

# sample txlogd configuration file for Apache 1.3

# Unique identifier associated with this sensor, sent to the Server.
SensorID SENSOR1

# Address of the Server.
ServerAddress 10.1.0.7

# Encrypt Sensor to Server communication. Defaults to 'on'.
SSL on

# Port used to connect to the Server. Defaults to 443 or 80, depending on
SSL being 'on'
# or 'off'.
ServerPort 443

# The common name (CN) of the Server as set in the Server's certificate.
CertName vs1.domain.com

# Path to directory that contains .pem certificate files.
CertPath /usr/local/visualsensor

# Full path to the Sensor disk queue on the web server.
QueueFile /var/queue/VisualSensor.dat

```

```
# Queue file size in MB.
QueueSize 2000

# The amount of time in seconds the transmitter will wait to send the next
# batch of packets.
# Defaults to 15.
#MaxPageLoadTime 15

# The name of the cookie set on the visitor's browser. Defaults to 'vlst'.
TrackingCookie vlst

# Whether or not to validate the server against CertName. Defaults to 'on'.
#VerifyCertName on

# Include or exclude certain content types from logging. These will be
prefix matched against the
# content type of the reply. For example, "image/" will match all image
content types while "image/gif"
# will match only that type exactly. When there are multiple matches for a
given content-type, the most
# specific match will be used. Therefore, you could put "image/gif" in the
ContentFilterInclude and
# "image/" in the ContentFilterExclude and image/gif replies will be allowed
but all other image types
# will be filtered out.
ContentFilterInclude *
ContentFilterExclude image/,text/css,application/x-javascript

# Used to work around certain problems. Defaults to 'no'. Only set to 'yes'
# on advice from tech
# support. Apache 1.3, 2.x, IHS 1.3, and IHS 2.x only.
ApacheUseAlternateHandler no

# Path to the Controlled Experiment configuration file.
ExpFile /home/experiment.txt

# If the user visits this page then a new tracking ID is generated and the
# user is placed in an
# experiment group.
ExpCookieURL /setcookie.html

# If set to 'on', URLs visited need only match part of the URL listed in
the experiment file to be
# considered a match. Defaults to 'off'.
ExpPartialMatch on

# The transmitter (txlogd) will periodically send requests to this location
# to see if the site is
# operating correctly. Note that the location is specified in the following
format, not as a URL:
#      http, <serverAddress>, <port>, /<resource>
# where: <serverAddress> is server name or IP address, <port> is server's
HTTP listening port,
# <resource> is the specific resource to request (can include a query
string).
# Only http is supported at this time. Multiple SiteTest lines can be
specified.
SiteTest http,localhost,80,/test.html

# Additional HTTP headers to log. Fields will be saved as cs(header-name)
in Server. Any valid
```

```
# HTTP header can be used (this is not a complete list).
LogHeader Accept
LogHeader Accept-Encoding
LogHeader Connection
LogHeader Host
LogHeader Keep-Alive
LogHeader Referrer

# Apache 1.3 specific additional fields to log (complete list).
SpecialLogFile x = request.allowed vys_log
SpecialLogFile x = request.boundary vys_log
SpecialLogFile x = request.byterange vys_log
SpecialLogFile x = request.bytes_sent vys_log
SpecialLogFile x = request.case_preserved_filename vys_log
SpecialLogFile x = request.chunked vys_log
SpecialLogFile x = request.clength vys_log
SpecialLogFile x = request.clf-request vys_log
SpecialLogFile x = request.content_encoding vys_log
SpecialLogFile x = request.content_type vys_log
SpecialLogFile x = request.expecting_100 vys_log
SpecialLogFile x = request.filename vys_log
SpecialLogFile x = request.handler vys_log
SpecialLogFile x = request.header_only vys_log
SpecialLogFile x = request.hostname vys_log
SpecialLogFile x = request.method vys_log
SpecialLogFile x = request.method_number vys_log
SpecialLogFile x = request.mtime vys_log
SpecialLogFile x = request.no_cache vys_log
SpecialLogFile x = request.no_local_copy vys_log
SpecialLogFile x = request.path_info vys_log
SpecialLogFile x = request.proto_num vys_log
SpecialLogFile x = request.protocol vys_log
SpecialLogFile x = request.query vys_log
SpecialLogFile x = request.range vys_log
SpecialLogFile x = request.read_body vys_log
SpecialLogFile x = request.read_chunked vys_log
SpecialLogFile x = request.read_length vys_log
SpecialLogFile x = request.remaining vys_log
SpecialLogFile x = request.request_time vys_log
SpecialLogFile x = request.sent_bodyct vys_log
SpecialLogFile x = request.status vys_log
SpecialLogFile x = request.status_line vys_log
SpecialLogFile x = request.unparsed_uri vys_log
SpecialLogFile x = request.uri vys_log
SpecialLogFile x = request.vlist_validator vys_log
SpecialLogFile x = headers.Referrer vys_log
SpecialLogFile x = headers.User-Agent vys_log
SpecialLogFile x = headers.Cookie vys_log
SpecialLogFile x = headers_out.Cookie vys_log

AddressFilter 10.2.1.89
# test machine
AddressFilter 192.168.*. *
# internal IP

DisableField s-dns
DisableField cs(cookie)

PrivacyID 0x11111111

# Send this Cache-Control response header to all new visitors
```

```
NewUserCacheControl no-cache=Set-Cookie  
  
# Send this Cache-Control response to returning visitors  
CacheControl private,max-age=0,must-revalidate
```

Apache Server 1.3.x on Windows 2000 Server or Later

Sample of the txlogd.conf file required for Sensor for Apache Server 1.3.x running under Microsoft Windows Server 2000 or later.

```
# sample txlogd configuration file for Apache 1.3  
  
# Unique identifier associated with this sensor, sent to the Server.  
SensorID SENSOR1  
  
# Address of the Server.  
ServerAddress 10.1.0.7  
  
# Encrypt Sensor to Server communication. Defaults to 'on'.  
SSL on  
  
# Port used to connect to the Server. Defaults to 443 or 80, depending on  
SSL being 'on'  
# or 'off'.  
ServerPort 443  
  
# The common name (CN) of the Server as set in the Server's certificate.  
CertName vs1.domain.com  
  
# Path to directory that contains .pem certificate files.  
CertPath C:\VisualSensor  
  
# Full path to the Sensor disk queue on the web server.  
# The queue file must be in the same directory as the Transmitter.  
QueueFile C:\VisualSensor\diskq2000.log  
  
# Queue file size in MB.  
QueueSize 2000  
  
# The amount of time in seconds the transmitter will wait to send the next  
batch of packets.  
# Defaults to 15.  
#MaxPageLoadTime 15  
  
# The name of the cookie set on the visitor's browser. Defaults to 'vlst'.  
TrackingCookie vlst  
  
# Whether or not to validate the server against CertName. Defaults to 'on'.  
#VerifyCertName on  
  
# Include or exclude certain content types from logging. These will be  
prefix matched against the  
# content type of the reply. For example, "image/" will match all image  
content types while "image/gif"  
# will match only that type exactly. When there are multiple matches for a  
given content-type, the most  
# specific match will be used. Therefore, you could put "image/gif" in the  
ContentFilterInclude and  
# "image/" in the ContentFilterExclude and image/gif replies will be allowed  
but all other image types  
# will be filtered out.
```

```

ContentFilterInclude *
ContentFilterExclude image/,text/css,application/x-javascript

# Used to work around certain problems. Defaults to 'no'. Only set to 'yes'
# on advice from tech
# support. Apache 1.3, 2.x, IHS 1.3, and IHS 2.x only.
ApacheUseAlternateHandler no

# Path to the Controlled Experiment configuration file.
ExpFile C:\VisualSensor\experiment.txt

# If the user visits this page then a new tracking ID is generated and the
# user is placed in an
# experiment group.
ExpCookieURL /setcookie.html

# If set to 'on', URLs visited need only match part of the URL listed in
the experiment file to be
# considered a match. Defaults to 'off'.
ExpPartialMatch on

# The transmitter (txlogd) will periodically send requests to this location
to see if the site is
# operating correctly. Note that the location is specified in the following
format, not as a URL:
#     http, <serverAddress>, <port>, /<resource>
# where: <serverAddress> is server name or IP address, <port> is server's
HTTP listening port,
# <resource> is the specific resource to request (can include a query
string).
# Only http is supported at this time. Multiple SiteTest lines can be
specified.
SiteTest http,localhost,80,/test.html

# Additional HTTP headers to log. Fields will be saved as cs(header-name)
in Server. Any valid
# HTTP header can be used (this is not a complete list).
LogHeader Accept
LogHeader Accept-Encoding
LogHeader Connection
LogHeader Host
LogHeader Keep-Alive
LogHeader Referrer

# Apache 1.3 specific additional fields to log (complete list).
SpecialLogField x = request.allowed vys_log
SpecialLogField x = request.boundary vys_log
SpecialLogField x = request.byterange vys_log
SpecialLogField x = request.bytes_sent vys_log
SpecialLogField x = request.case_preserved_filename vys_log
SpecialLogField x = request.chunked vys_log
SpecialLogField x = request.clength vys_log
SpecialLogField x = request.clf-request vys_log
SpecialLogField x = request.content_encoding vys_log
SpecialLogField x = request.content_type vys_log
SpecialLogField x = request.expecting_100 vys_log
SpecialLogField x = request.filename vys_log
SpecialLogField x = request.handler vys_log
SpecialLogField x = request.header_only vys_log
SpecialLogField x = request.hostname vys_log
SpecialLogField x = request.method vys_log

```

```

SpecialLogField x = request.method_number vys_log
SpecialLogField x = request.mtime vys_log
SpecialLogField x = request.no_cache vys_log
SpecialLogField x = request.no_local_copy vys_log
SpecialLogField x = request.path_info vys_log
SpecialLogField x = request.proto_num vys_log
SpecialLogField x = request.protocol vys_log
SpecialLogField x = request.query vys_log
SpecialLogField x = request.range vys_log
SpecialLogField x = request.read_body vys_log
SpecialLogField x = request.read_chunked vys_log
SpecialLogField x = request.read_length vys_log
SpecialLogField x = request.remaining vys_log
SpecialLogField x = request.request_time vys_log
SpecialLogField x = request.sent_bodyct vys_log
SpecialLogField x = request.status vys_log
SpecialLogField x = request.status_line vys_log
SpecialLogField x = request.unparsed_uri vys_log
SpecialLogField x = request.uri vys_log
SpecialLogField x = request.vlist_validator vys_log
SpecialLogField x = headers.Refferrer vys_log
SpecialLogField x = headers.User-Agent vys_log
SpecialLogField x = headers.Cookie vys_log
SpecialLogField x = headers_out.Cookie vys_log

AddressFilter 10.2.1.89
# test machine
AddressFilter 192.168.*.*
# internal IP

DisableField s-dns
DisableField cs(cookie)

PrivacyID 0x11111111

# Send this Cache-Control response header to all new visitors
NewUserCacheControl no-cache=Set-Cookie

# Send this Cache-Control response to returning visitors
CacheControl private,max-age=0,must-revalidate

```

Apache Server 2.0.x or 2.2 on Linux, Sun Solaris, or FreeBSD

Sample of the txlogd.conf file required for Sensor for the following Apache Servers.

```

Apache Server 2.0.40 running under RedHat Linux 7.x or later or Sun Solaris
SPARC 2.6 or later
Apache Server 2.0.42 or later running under RedHat Linux 7.x or later, Sun
Solaris SPARC 2.6 or later, SUSE Linux 9.x or later, or FreeBSD 5.3
Apache Server 2.0.42 or later running under 64-bit versions of RedHat Linux
ES 4 and ES 5
Apache Server 2.2 running under RedHat Linux 7.x or later

# sample txlogd configuration file for Apache 2.0.x or 2.2

# Unique identifier associated with this sensor, sent to the Server.
SensorID SENSOR1

# Address of the Server.
ServerAddress 10.1.0.7

```

```
# Encrypt Sensor to Server communication.  
# Defaults to 'on'.  
SSL on  
  
# Port used to connect to the Server. Defaults to 443 or 80, depending on  
SSL being 'on' or 'off'.  
ServerPort 443  
  
# The common name (CN) of the Server as set in the Server's certificate.  
CertName vsl.domain.com  
  
# Path to directory that contains .pem certificate files.  
CertPath /usr/local/visualsensor  
  
# Full path to the Sensor disk queue on the web server.  
QueueFile /var/queue/VisualSensor.dat  
  
# Queue file size in MB.  
QueueSize 2000  
  
# The amount of time in seconds the transmitter will wait to send the next  
batch of packets. Defaults to 15.  
MaxPageLoadTime 15  
  
# The name of the cookie set on the visitor's browser. Defaults to 'v1st'.  
TrackingCookie v1st  
  
# Whether or not to validate the server against CertName. Defaults to 'on'.  
#VerifyCertName on  
  
# Include or exclude certain content types from logging. These will be  
prefix matched against the  
# content type of the reply. For example, "image/" will match all image  
content types while "image/gif"  
# will match only that type exactly. When there are multiple matches for a  
given content-type, the most  
# specific match will be used. Therefore, you could put "image/gif" in the  
ContentFilterInclude and  
# "image/" in the ContentFilterExclude and image/gif replies will be allowed  
but all other image types will be filtered out.  
ContentFilterInclude *  
ContentFilterExclude image/,text/css,application/x-javascript  
  
# Apache 1.3, 2.x, IHS 1.3, and IHS 2.x only. Used to work around certain  
problems. Defaults to 'no'.  
# Only set to 'yes' on advice from tech support.  
ApacheUseAlternateHandler no  
  
# Path to the Controlled Experiment configuration file.  
ExpFile /home/experiment.txt  
  
# If the user visits this page then a new tracking ID is generated and the  
user is placed in an experiment group.  
ExpCookieURL /setcookie.htm  
  
# If set to 'on', URLs visited need only match part of the URL listed in  
the experiment file to be  
# considered a match. Defaults to 'off'.  
ExpPartialMatch on  
  
# The transmitter (txlogd) will periodically send requests to this location
```

```
to see if the site is
# operating correctly. Note that the location is specified in the following
format, not as a URL:
#           http, <serverAddress>, <port>, /<resource>
# where: <serverAddress> is server name or IP address, <port> is server's
HTTP listening port,
# <resource> is the specific resource to request (can include a query
string).
# Only http is supported at this time. Multiple SiteTest lines can be
specified.
SiteTest http,localhost,80,/test.html

# Additional HTTP headers to log. Fields will be saved as cs(header-name)
in Server. Any valid
# HTTP header can be used (this is not a complete list).
LogHeader Accept
LogHeader Accept-Encoding
LogHeader Connection
LogHeader Host
LogHeader Keep-Alive
LogHeader Referrer

# Apache 2.0.x specific additional fields to log (complete list).
SpecialLogFile x = request.allowed vys_log
SpecialLogFile x = request.ap_auth_type vys_log
SpecialLogFile x = request.query vys_log
SpecialLogFile x = request.bytes_sent vys_log
SpecialLogFile x = request.canonical_filename vys_log
SpecialLogFile x = request.chunked vys_log
SpecialLogFile x = request.clength vys_log
SpecialLogFile x = request.content_encoding vys_log
SpecialLogFile x = request.content_type vys_log
SpecialLogFile x = request.eos_sent vys_log
SpecialLogFile x = request.expecting_100 vys_log
SpecialLogFile x = request.filename vys_log
SpecialLogFile x = request.handler vys_log
SpecialLogFile x = request.header_only vys_log
SpecialLogFile x = request.hostname vys_log
SpecialLogFile x = request.method vys_log
SpecialLogFile x = request.method_number vys_log
SpecialLogFile x = request.mtime vys_log
SpecialLogFile x = request.no_cache vys_log
SpecialLogFile x = request.no_local_copy vys_log
SpecialLogFile x = request.path_info vys_log
SpecialLogFile x = request.proto_num vys_log
SpecialLogFile x = request.protocol vys_log
SpecialLogFile x = request.proxyreq vys_log
SpecialLogFile x = request.range vys_log
SpecialLogFile x = request.read_body vys_log
SpecialLogFile x = request.read_chunked vys_log
SpecialLogFile x = request.read_length vys_log
SpecialLogFile x = request.remaining vys_log
SpecialLogFile x = request.request_time vys_log
SpecialLogFile x = request.sent_bodyct vys_log
SpecialLogFile x = request.status vys_log
SpecialLogFile x = request.status_line vys_log
SpecialLogFile x = request.clf-request vys_log
SpecialLogFile x = request.unparsed_uri vys_log
SpecialLogFile x = request.uri vys_log
SpecialLogFile x = request.used_path_info vys_log
SpecialLogFile x = request.user vys_log
```

```

SpecialLogField x = request.vlist_validator vys_log
SpecialLogField x = headers.Refferrer vys_log
SpecialLogField x = headers.User-Agent vys_log
SpecialLogField x = headers.Cookie vys_log
SpecialLogField x = headers_out.Cookie vys_log

AddressFilter 10.2.1.89
# test machine
AddressFilter 192.168.*.*
# internal IP

DisableField s-dns
DisableField cs(cookie)

PrivacyID 0x11111111

# Send this Cache-Control response header to all new visitors
NewUserCacheControl no-cache=Set-Cookie

# Send this Cache-Control response to returning visitors
CacheControl private,max-age=0,must-revalidate

```

Apache Server 2.0.x or 2.2 on Windows Server 2000 or Later

Sample of the txlogd.conf file required for Sensor for Apache Server 2.0.x or 2.2 running under Microsoft Windows Server 2000 or later.

```

# sample txlogd configuration file for Apache 2.0.x or 2.2

# Unique identifier associated with this sensor, sent to the Server.
SensorID SENSOR1

# Address of the Server.
ServerAddress 10.1.0.7

# Encrypt Sensor to Server communication.
# Defaults to 'on'.
SSL on

# Port used to connect to the Server. Defaults to 443 or 80, depending on
# SSL being 'on'
# or 'off'.
ServerPort 443

# The common name (CN) of the Server as set in the Server's certificate.
CertName vs1.domain.com

# Path to directory that contains .pem certificate files.
CertPath C:\VisualSensor

# Full path to the Sensor disk queue on the web server.
# The queue file must be in the same directory as the Transmitter.
QueueFile C:\VisualSensor\diskq2000.log

# Queue file size in MB.
QueueSize 2000

# The amount of time in seconds the transmitter will wait to send the next
# batch of packets.
# Defaults to 15.
MaxPageLoadTime 15

```

```

# The name of the cookie set on the visitor's browser. Defaults to 'v1st'.
TrackingCookie v1st

# Whether or not to validate the server against CertName. Defaults to 'on'.
#VerifyCertName on

# Include or exclude certain content types from logging. These will be
prefix matched against the
# content type of the reply. For example, "image/" will match all image
content types while "image/gif"
# will match only that type exactly. When there are multiple matches for a
given content-type, the most
# specific match will be used. Therefore, you could put "image/gif" in the
ContentFilterInclude and
# "image/" in the ContentFilterExclude and image/gif replies will be allowed
but all other image types
# will be filtered out.
ContentFilterInclude *
ContentFilterExclude image/,text/css,application/x-javascript

# Apache 1.3, 2.x, IHS 1.3, and IHS 2.x only. Used to work around certain
problems. Defaults to 'no'.
# Only set to 'yes' on advice from tech support.
ApacheUseAlternateHandler no

# Path to the Controlled Experiment configuration file.
ExpFile C:\VisualSensor\experiment.txt

# If the user visits this page then a new tracking ID is generated and the
user is placed in an
# experiment group.
ExpCookieURL /setcookie.htm

# If set to 'on', URLs visited need only match part of the URL listed in
the experiment file to be
# considered a match. Defaults to 'off'.
ExpPartialMatch on

# The transmitter (txlogd) will periodically send requests to this location
to see if the site is
# operating correctly. Note that the location is specified in the following
format, not as a URL:
#      http, <serverAddress>, <port>, /<resource>
# where: <serverAddress> is server name or IP address, <port> is server's
HTTP listening port,
# <resource> is the specific resource to request (can include a query
string).
# Only http is supported at this time. Multiple SiteTest lines can be
specified.
SiteTest http,localhost,80,/test.html

# Additional HTTP headers to log. Fields will be saved as cs(header-name)
in Server. Any valid
# HTTP header can be used (this is not a complete list).
LogHeader Accept
LogHeader Accept-Encoding
LogHeader Connection
LogHeader Host
LogHeader Keep-Alive
LogHeader Referrer

```

```

# Apache 2.0.x specific additional fields to log (complete list).
SpecialLogField x = request.allowed vys_log
SpecialLogField x = request.ap_auth_type vys_log
SpecialLogField x = request.query vys_log
SpecialLogField x = request.bytes_sent vys_log
SpecialLogField x = request.canonical_filename vys_log
SpecialLogField x = request.chunked vys_log
SpecialLogField x = request.clength vys_log
SpecialLogField x = request.content_encoding vys_log
SpecialLogField x = request.content_type vys_log
SpecialLogField x = request.eos_sent vys_log
SpecialLogField x = request.expecting_100 vys_log
SpecialLogField x = request.filename vys_log
SpecialLogField x = request.handler vys_log
SpecialLogField x = request.header_only vys_log
SpecialLogField x = request.hostname vys_log
SpecialLogField x = request.method vys_log
SpecialLogField x = request.method_number vys_log
SpecialLogField x = request.mtime vys_log
SpecialLogField x = request.no_cache vys_log
SpecialLogField x = request.no_local_copy vys_log
SpecialLogField x = request.path_info vys_log
SpecialLogField x = request.proto_num vys_log
SpecialLogField x = request.protocol vys_log
SpecialLogField x = request.proxyreq vys_log
SpecialLogField x = request.range vys_log
SpecialLogField x = request.read_body vys_log
SpecialLogField x = request.read_chunked vys_log
SpecialLogField x = request.read_length vys_log
SpecialLogField x = request.remaining vys_log
SpecialLogField x = request.request_time vys_log
SpecialLogField x = request.sent_bodyct vys_log
SpecialLogField x = request.status vys_log
SpecialLogField x = request.status_line vys_log
SpecialLogField x = request.clf-request vys_log
SpecialLogField x = request.unparsed_uri vys_log
SpecialLogField x = request.uri vys_log
SpecialLogField x = request.used_path_info vys_log
SpecialLogField x = request.user vys_log
SpecialLogField x = request.vlist_validator vys_log
SpecialLogField x = headers.Referrer vys_log
SpecialLogField x = headers.User-Agent vys_log
SpecialLogField x = headers.Cookie vys_log
SpecialLogField x = headers_out.Cookie vys_log

AddressFilter 10.2.1.89
# test machine
AddressFilter 192.168.*. *
# internal IP

DisableField s-dns
DisableField cs(cookie)

PrivacyID 0x11111111

# Send this Cache-Control response header to all new visitors
NewUserCacheControl no-cache=Set-Cookie

```

```
# Send this Cache-Control response to returning visitors
CacheControl private,max-age=0,must-revalidate
```

Sensor Transmitter Command-Line Options

Information about Sensor transmitter start-up commands for UNIX and for Windows.

Start-up Command for UNIX

To start the Sensor transmitter on a UNIX stem, you use the following command:

```
txlogd -f configFileName
```

where configFileName is the fully qualified name of the transmitter configuration file (txlogd.conf).

By default, the transmitter runs as a background process (a daemon) and writes operational messages to syslog.

The following is a complete list of the command-line switches for txlogd:

Switch	Description
-f	Specifies the fully qualified name of the configuration file (txlogd.conf). If you do not specify this switch, the transmitter looks for txlogd.conf in the current directory.
-n	Starts the transmitter in interactive mode (not as a background process).
-i	Starts the transmitter in interactive mode and directs debug output to stdout.
-d	Starts the transmitter as a background process and directs debug output to txlogd-debug.log in the current directory.
-c	Starts the transmitter and creates the disk queue file if it doesn't exist. If the disk queue already exists, this parameter is ignored.
-x	Starts the transmitter and causes it to exit after it transmits the last packet in the disk queue. You can use this option to drain the queue in preparation for an administrative operation such as creating a new queue file.

Start-up Command for Windows

To start Sensor and register it as a service on a Windows system, you use the following command:

```
txlog /regserver
```

Sensor UNIX File Permissions

Information about Sensor UNIX file permissions such as the collector module, the transmitter process, the configuration file, and more.

The Collector Module

Quality	Description
File name	mod_visual_sciences.so (on Apache web servers and IBM HTTP servers) libvisual_sciences.so and J2EECollector.jar (on J2EE application servers) aol_visual_sciences.so (on AOL web servers) saf_visual_sciences.so (on Sun Java web servers)
Default Permissions	rwx r-x r-x (IBM HTTP and Apache 1.3.x)

Quality	Description
	<p>rwx rwx r-x (Apache 2.0.x)</p> <p>rwx --x --x (J2EE, AOL, and Sun Java web servers)</p>
Read by	<p>The web server, which sometimes runs as the root user, but more often runs under a specific user account</p> <p>If the web server runs under a non-root account, the permissions on this file must allow that account to read it.</p>
Runs as	<p>A child process in the web server</p> <p>Child processes run under a user account specified in your web server configuration. On many servers, this is a special account with very limited privileges called “nobody” — but not all web servers use this account. Check your web server’s configuration file to determine what user account is set.</p>
Reads from	<p>txlogd.conf</p> <p>The diskQueue file</p>
Writes to	The diskQueue file

The Transmitter Process

Quality	Description
File name	trust_ca_cert.pem
Default Permissions	<p>rw- r-- r-- (IBM HTTP, AOL, and Sun Java web servers)</p> <p>rw- rw- r-- (Apache web servers and J2EE application servers)</p>
Read by	The transmitter program
Written by	--

The Configuration File

Quality	Description
File name	txlogd.conf
Default Permissions	<p>rw- r-- r-- (IBM HTTP, AOL, and Sun Java web servers)</p> <p>rw- rw- r-- (Apache web servers and J2EE application servers)</p>
Read by	<p>The collector module</p> <p>The transmitter program</p> <p>The administrator responsible for installing, configuring, and maintaining Sensor</p>
Written by	The administrator responsible for installing, configuring, and maintaining Sensor

The Certificate Authority File

Quality	Description
File name	trust_ca_cert.pem
Default Permissions	rw- r-- r-- (IBM HTTP, AOL, and Sun Java web servers) rw- rw- r-- (Apache web servers and J2EE application servers)
Read by	The transmitter program
Written by	--

The Disk Queue

Quality	Description
File name	User defined
Default Permissions	rw- rw- rw- (All server types)
Read by	The collector module The transmitter program
Written by	The collector module The transmitter program

Installing the Monitoring Profile

Directions for installing the data workbench Monitoring Profile.

Installation Steps

1. Configure a new Sensor instance as if it will be used for tagged web page data collection. Be sure the zig.gif file is in the Sensor web server document root. Sensor can be run on the same host as the monitor profiles. (This is not an issue if using a text file for this purpose.)



Note: This Sensor instance must be dedicated to receiving only traffic from the Monitoring Agents. Also, the Sensor can be configured to run on a different port if you are re-purposing a web server for this collection.

2. In the **txlogd.conf** file there is the default line:

```
ContentFilterExclude  
image/,text/css,application/x-javascript,text/javascript
```

For the data workbench Monitoring Profile application (or any "tagged" page implementation), the image type has to be removed in order to collect via a GIF file. The updated line is:

```
ContentFilterExclude text/css,application/x-javascript,text/javascript
```

3. Copy the **insight_monitor.zip/insight_monitor_agent** to a temporary location.

-
4. Update `insight_monitor_agent.cfg` file for your environment. Follow the comments inside the configuration file:

The Monitoring configuration file:

```
# Change to Insight Monitor Sensor name or IP address
Collection Stem = string: http://insightmonitorsensor.foo.com/zig.gif
```

Define where you are collecting all information and provide URL address. This needs to be a dedicated sensor and should receive no traffic except for this application.

```
# If Message Level = 2 the Insight Monitor agent will dump the entire detailed
# status page to the following folder
Detailed Status Dump Dir = string: e:\\insight_monitor\\detailed_status
```

There are paths assuming there is an e: disk. You may want to change this path for your environment.

```
# How many times quick check fails before we consider Insight to really be down
Down Timeout Count = int: 3
```

Sometime when running a Transform profile, data workbench can be unresponsive. This value lets you send an alert if three times in a row the process is unresponsive. This is a way of reducing false positive alerts.

```
# Set to "test" or "staging" if this is not an production host. This is so you can filter
# on environment in the resulting dataset, e.g., seeing only production uptime metrics
Environment = string: production
```

This is where you set the environment and group dimensions. This may be different from host to host.

```
# File to log messages from Insight Monitor agent
Message File = string: e:\\insight_monitor\\insight_monitor.log
```

This is where you can see exactly what the monitor agent is doing by viewing error logs in this path.

```
# Alert on Temp DB internal usage above this percent
Temp DB Space Critical Percentage = int: 95
```

This is to use the temp db internally. It may be alerted when reaching capacity. This is different than physical disk usage.

5. Copy the `insight_monitor_agent` folder to each DPU and FSU host running the data workbench server. The default location as indicated in the configuration file is `e:\\insight_monitor_agent` but you can change this location.
6. Add a Windows scheduled task to invoke the agent every 10 minutes (this period is assumed in the processing rate calculations). The program is `e:\\insight_monitor\\insight_monitor_agent.exe`. The argument is config-file `e:\\insight_monitor\\insight_monitor.cfg`. Start in `e:\\insight_monitor`. The user running executing the task must have permission to read/write `e:\\insight_monitor` and read the Win32 OLE object `rroot\\CIMV2` (required to ascertain the data workbench server service start mode and to check the percentage of space on local disks)
7. Confirm that the VSL file is starting to grow as monitor records accumulate. This will take some time as the traffic volume will be extremely low in a small installation (every 10 minutes the agent sends only one hit for the host-specific data, plus one hit per processing profile).
8. Unzip `insight_monitor.zip\\profiles\\Insight Historic` to a temporary location.
9. Update host name in `profile.cfg`, `dataset\\cluster.cfg`, and the `dataset\\segment export.cfg`.
10. Update the files to the data workbench profiles directory.
11. Update log server and path in `dataset\\log processing.cfg` to the location where the Sensor VSLs are accumulating.
12. [Optional] do the same with the profiles **Insight Profile Status** and **Insight Server Status**. In addition, the status profiles should be reprocessed nightly with a trailing two day window. Add a Windows scheduled

-
- task: The program is e:\insight_monitor\insight_reprocess.exe. The argument is
--profile-path="PATH TO PROFILES\insight profile status"
--start-days-ago=2. Leave start in blank. Add another scheduled task for "insight server status".
insight_reprocess.exe requires read/write access to log processing.cfg to update the start time.
13. In addition, the status profiles should be reprocessed nightly with a trailing two day window. Add a Windows scheduled task: The program is e:\insight_monitor\insight_reprocess.exe. The argument is
--profile-path="PATH TO PROFILES\insight profile status" --start-days-ago=2.
Leave start in blank. Add another scheduled task for "insight server status".
insight_reprocess.exe requires read/write access to **log processing.cfg** to update the start time.
Confirm that each profile is reading the monitor VSLs as they accumulate. Again, this will take some time—probably hours—because of the extremely low volume.

Installation Notes

- **Configuring the Monitoring Profile in a licensed test environment.** The test environment package is included with your implementation of data workbench, allowing you to install and configure the application. If installing on a production FSU or DPU server, then you will need to configure the server to run on a separate port.
- **Deploying a new Sensor specifically for the Monitoring Profile.** You will need to install a new instance of Sensor to the server running the Monitoring Profile. This is in addition to the production instance of Sensor. (There is no additional charge for installing Sensor on either a production or non-production server specifically for the Monitoring Profile.)
- **Disable the monitor agent during data workbench maintenance.** To avoid polluting the uptime and performance metrics, you can set service start mode to manual for the service InsightServer (Omniture Insight Server). A handy PowerShell command is `set-service -name insightserver -startuptype manual`. Set it back to automatic after the maintenance: `set-service -name insightserver -startuptype automatic`. Another option is to temporarily disable the monitor agent scheduled task.
- **The Status profiles need a trailing window** to drop old hosts and profiles as well as old host-profile mappings. However, if the amount of event data is so small that data workbench won't buffer it in, then you may need to extend the size of the window quite a bit to get it to process.
- **The agent collects the overall and oldest as-of time from data workbench detailed status**, which is reported in local host time assuming the event data log time stamps are in UTC (as in VSL files). If the event data timestamps are in a non-UTC time zone the as-of time will be offset in the resulting Insight Profile Status profile. If **all** of your event data time stamps are in the same time zone you can add that offset to *Insight Profile Status\metrics\as of delay minutes.metric*.
- **Two new dimensions were introduced to help the customer group their servers if they are in different states**, such as production, staging, testing servers, and servers in other states. For example if you are looking for "uptime", then you look at servers only in production mode. As a result, the Group dimension is just another way of arbitrarily grouping servers for your needs. For example, in the Monitoring Configuration file you can set, which host your department is servicing, such as Operations, Development, or Marketing.

Workspaces for Monitoring the Data Workbench Server

To successfully identify server health and performance, you can use standard data workbench profiles to monitor the server from the installed agent using current data, or employ profiles of historic data sets to view the impact of performance changes over time.

The data workbench workspaces most commonly used include:

- [Data Workbench Historic workspace](#) on page 979
- [Data Workbench Profile Status workspace](#) on page 982
- [Data Workbench Server Status workspace](#) on page 984

To select a profile, open the drop-down menu from the upper-left corner of the data workbench Client interface.

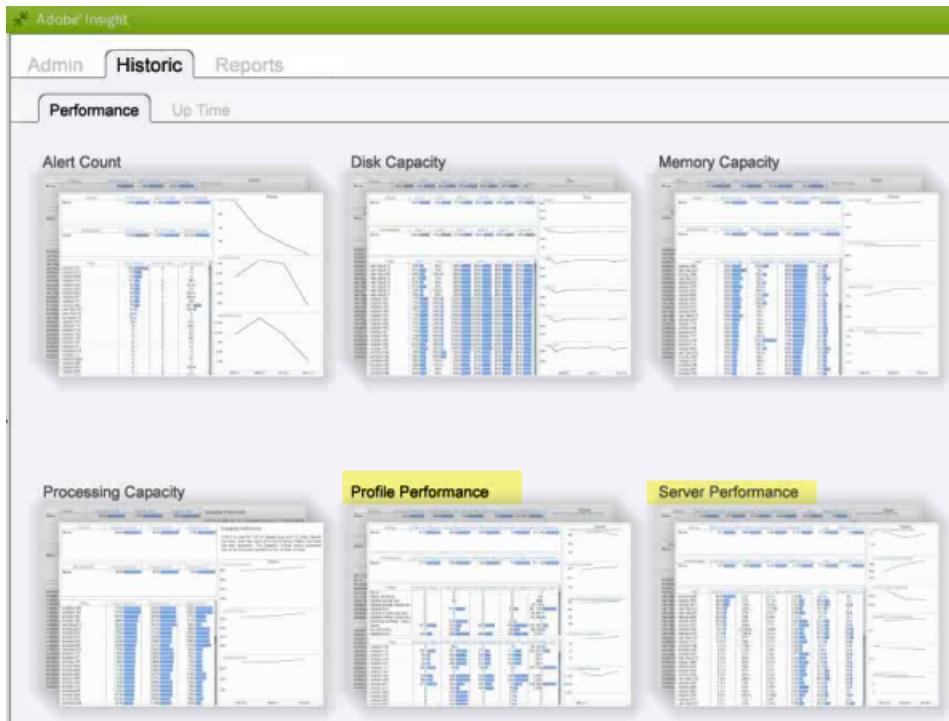
Switch profile to:

Configuration	on sd4dpu325.insight.server.com
✓ Insight Historic	on sd4dpu325.insight.server.com
Insight Profile Status	on sd4dpu325.insight.server.com
Insight Server Status	on sd4dpu325.insight.server.com

Data Workbench Historic workspace

Use the **data workbench Historic profile** to see how configuration, hardware, and other changes impact performance, stability, and server capacity over time.

The Historic profile includes a profile-based *Profile Performance workspace* on page 979 dataset and the server-based *Server Performance workspace* on page 980 dataset under the **Performance** tab. These are the most commonly used datasets viewed for a past perspective of the data workbench server performance. In addition, you can view the *Components workspace* on page 981 and *Components workspace* on page 981 by selecting the **Up Time** tab.



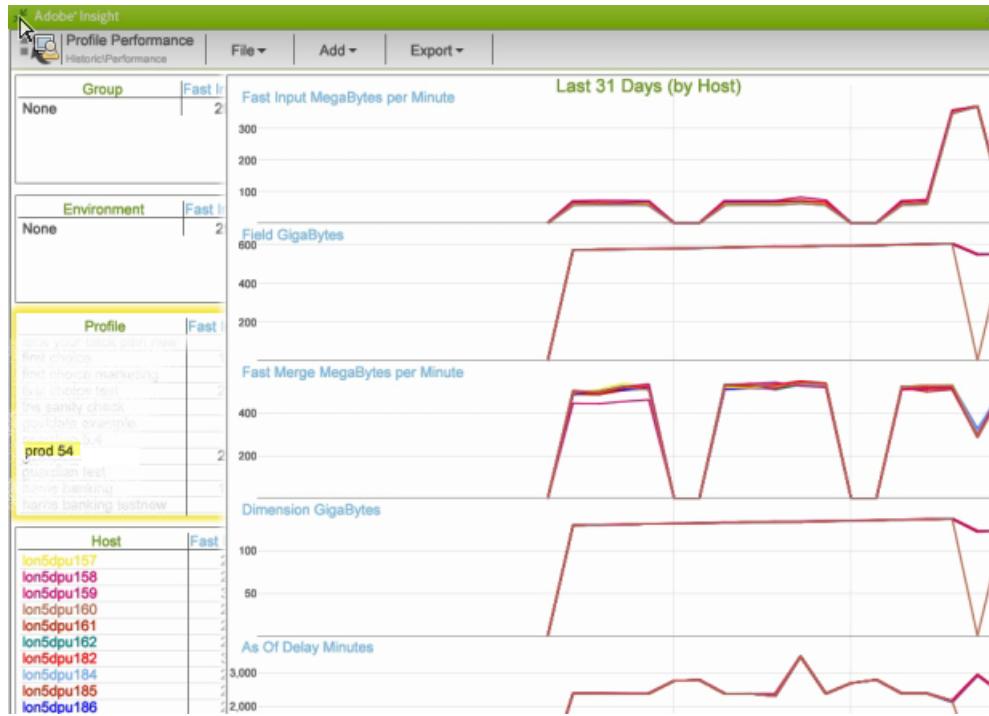
In addition, you can view the *Components workspace* on page 981 and *Components workspace* on page 981 by selecting the **Up Time** tab.

For additional reference information about the dimensions used in the data workbench Historic profile, see *Dimensions in the Data Workbench Historic profile* on page 986

Profile Performance workspace

This dataset includes the following relevant metrics for data workbench monitoring.

- Fast Input MegaBytes per Minute—metrics displaying heavy data input during initial log processing.
- Fast Merge MegaBytes per Minute—metrics displaying transformation.

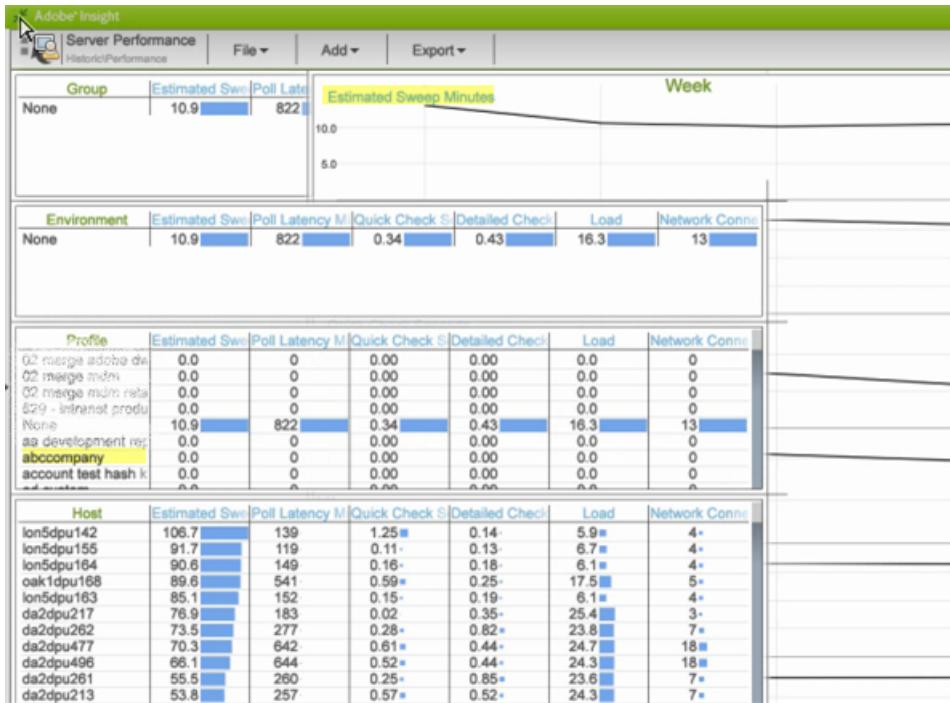


Note: To do a real performance assessment of your profile, look at rate rather than elapsed calendar time. The rate is measured as the changed values between polling every ten minutes.

Server Performance workspace

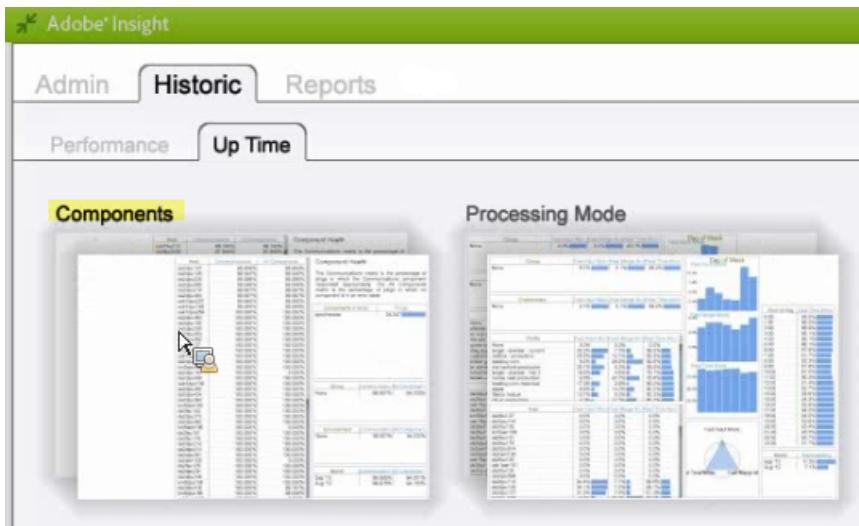
This dataset monitors server metrics beyond the scope of included profiles, and includes the following relevant server metrics for data workbench monitoring.

- Estimated Sweep Minutes — Estimated query resolution time.
- Poll Latency Milliseconds — Indicator of how busy software is by measuring how long it takes to get through a full cycle of servicing every component.



Components workspace

This dataset is located under the Up Time tab.



The Components dataset includes two aspects for component health:

- Communications metric — Did the data workbench server process respond?
- All Components metric — At top of Detailed Status page is a list of components the host is servicing within the data workbench server processes. If any component is in an error state then it is listed under the Components in Error table.

Host	Communications	All Components
lon5sen107	100.000%	0.000%
lon5sen113	100.000%	0.000%
da2sen103	100.000%	0.000%
da2sen110	100.000%	0.000%
da2sen117	100.000%	0.000%
da2dpu463	100.000%	0.000%
da2sen127	100.000%	0.000%
oak1sen102	100.000%	0.000%
da2sen102	100.000%	0.000%
da2sen130	100.000%	0.000%
da2sen109	100.000%	0.000%
da2sen119	100.000%	0.000%
lon5sen106	100.000%	0.000%
da2sen128	100.000%	0.000%
oak1sen101	100.000%	0.000%
da2sen118	100.000%	0.000%
da2dpu144	100.000%	16.093%
da2fsu140	99.578%	28.888%

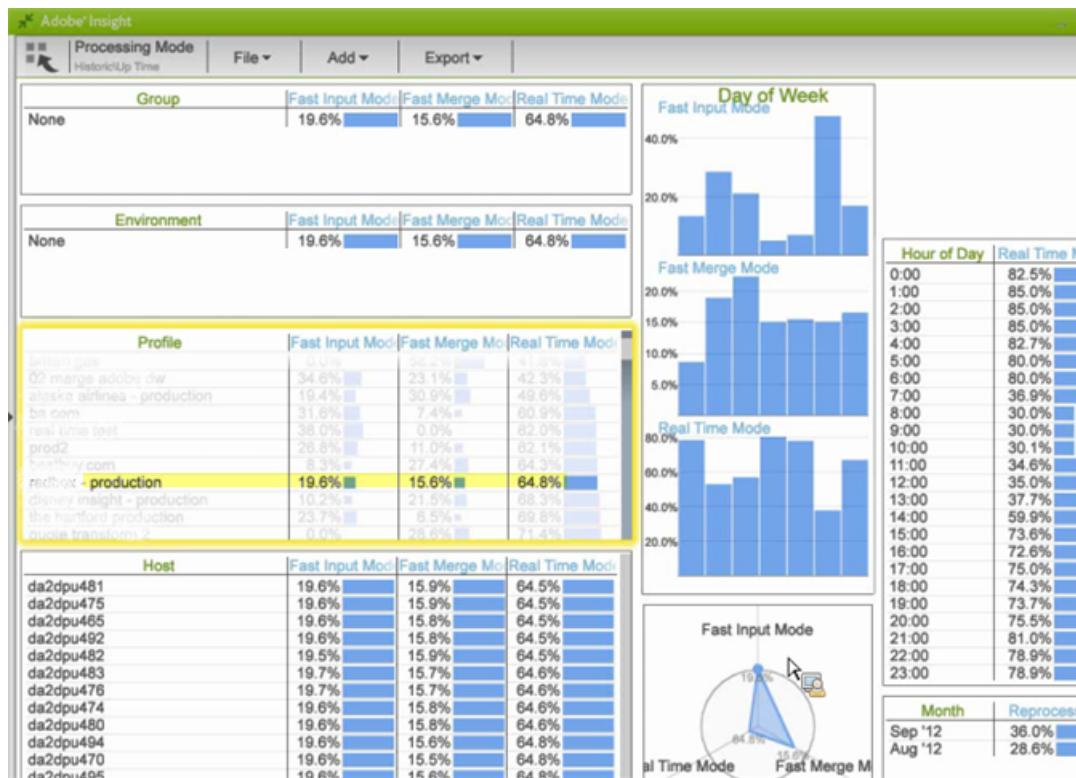
Component Health

The Communications metric is the percentage of pings in which the Communications component responded appropriately. The All Components metric is the percentage of pings in which no component is in an error state.

Components in Error	Pings
adminemail	4,677
diskspacemanager	2,951
synchronize	2,466
replicate	2,079
serverprofilemanager	51

Processing Mode workspace

This workspace is located under the Up Time tab. This workspace lets you observe how much time is taken in fast input, fast merge, and real-time modes.



This dataset provides important server load characteristics, such as identifying data load for

- Day of the week (for example a Fast Input Rate on Tuesday and Wednesday),
- Hour of Day (what percentage of the day is it in Fast Input mode?)

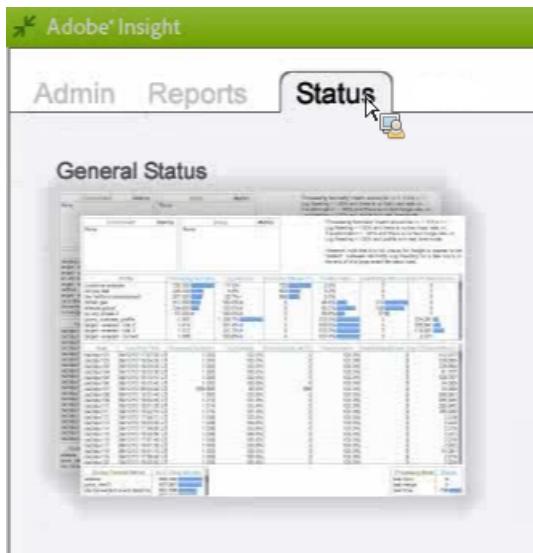
Data Workbench Profile Status workspace

The data workbench Profile Status profile provides current information about the data workbench server health based on the profile rather than server metrics or historical data.

Data Workbench Profile Status

This status profile provides the data workbench server information that is current, but not quite real-time because the agent is polled every ten minutes and reporting always includes this ten-minute latency. More precisely, the datasets generated by this profile provide the latest observation of the server from the agent, which most often has a default polling period of ten minutes.

For additional reference information about the dimensions used in the data workbench Profile Status profile, see [Insight Profile Status profile](#).



This report is more for monitoring operations rather than components or specific traffic fluctuations.

Adobe® Insight

General Status | File ▾ | Add ▾ | Export ▾

Environment	Metrics	Group	Metrics
None		None	
Profile	Processing Normally	Log Reading	Fast Input MB per Min
marketing-marketing	476.000	21.9%	476
bill pay test	403.000	4.8%	403
ba.com	193.250	77.7%	193
bmo-test	140.000	52.4%	140
bmo banking	45.000	99.8%	45
dedupe for wbench	15.000	97.3%	15
memor ba.com	1.000	100.0%	0
bmo - final-test	1.000	100.0%	0
vodafone.uk	1.000	100.0%	0
bmo - production - now	1.000	100.0%	0

Host	Last Ping Time	Processing Normally	Log Reading	Fast Input MB per Min
da2dpu101	09/18/12 18:27:00 UT	1.000	100.0%	0
da2dpu102	09/18/12 18:24:30 UT	1.000	100.0%	0
da2dpu103	09/18/12 18:33:00 UT	1.000	100.0%	0
da2dpu104	09/18/12 01:43:30 UT	1.000	100.0%	0
da2dpu105	09/18/12 18:30:40 UT	1.000	100.0%	0
da2dpu106	09/18/12 18:32:40 UT	1.000	100.0%	0
da2dpu107	09/17/12 04:40:20 UT	404.646	80.7%	404
da2dpu108	09/18/12 18:31:40 UT	1.000	100.0%	0
da2dpu109	09/18/12 18:24:40 UT	1.014	101.4%	0

This gives us a sense of who is in what mode: If we see a high Fast Input rate for a certain profile then that profile is in Fast Input mode.

If the Stalled metric is 1, then the server is stalled. If the value is 0, then the server is not stalled.

Log Reading for large batch loads

Adobe® Insight

General Status | File ▾ | Add ▾ | Export ▾

Environment	Metrics	Group	Metrics			
None		None				
Profile	Processing Normally	Log Reading	Fast Input MB per Min	Transformation	Fast Merge MB per Min	As Of Delay Minutes
processing	0.000	53.2%	0	0.0%	0	0
action-shr2	0.000	0.0%	0	100.0%	0	0
rest - current	0.998	99.8%	0	100.0%	0	3,087
tion	0.998	99.8%	0	100.0%	0	3,309

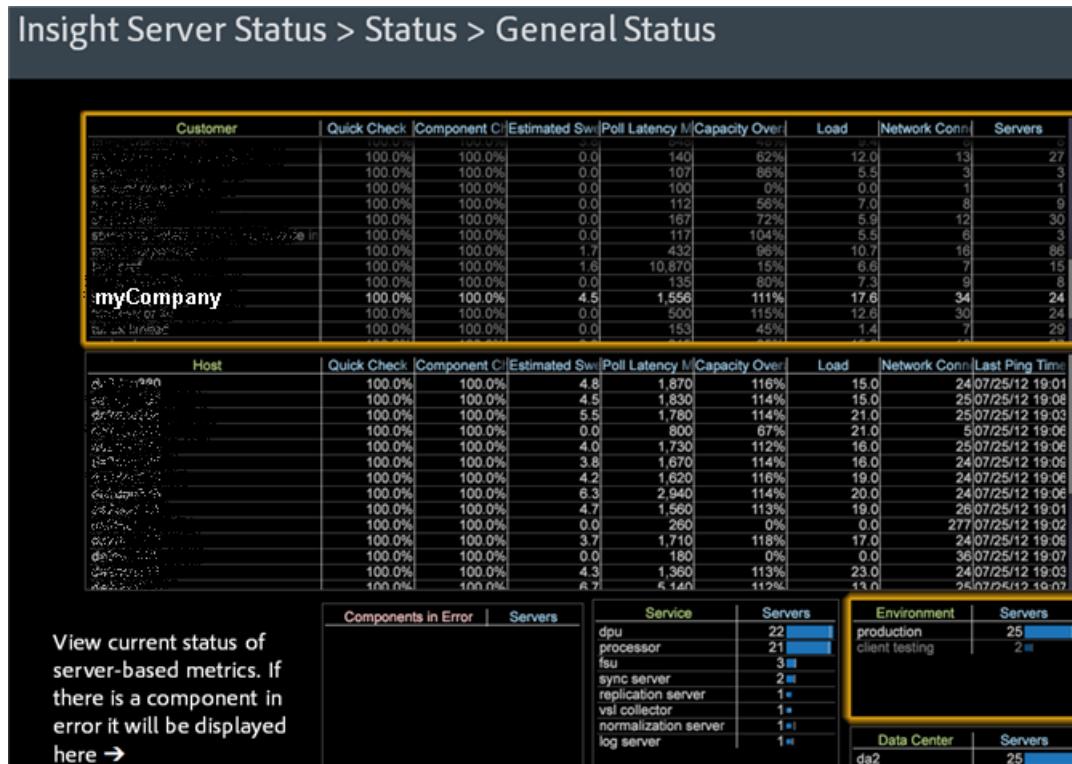
However, note that it is not unusual for Insight to appear to be "stalled" between 99-100% Log Reading for a few hours at the end of a large event file batch load.

Data Workbench Server Status workspace

The data workbench Server Status profile provides current information about data workbench server health based on the server rather than profile metrics or historical data.

General Status

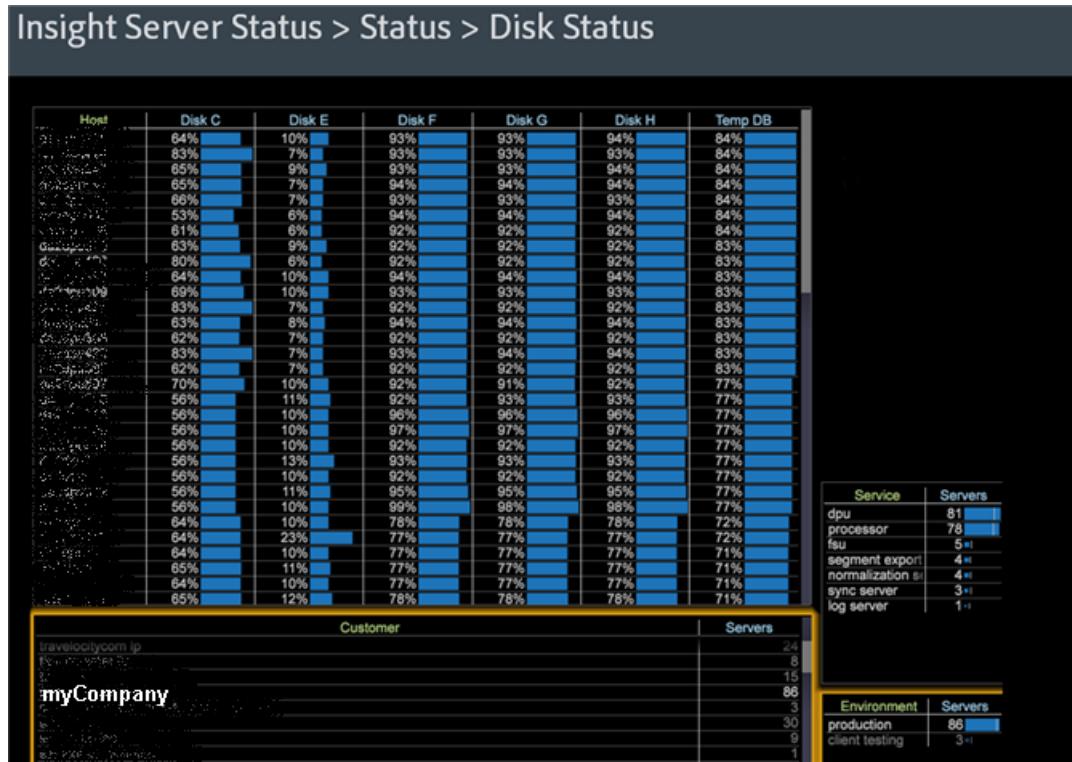
Open the General Status dataset view within the data workbench Server Status profile.



For additional reference information about the dimensions used in the data workbench Server Status profile, see the [Dimensions in the Data Workbench Server Status profile](#) on page 992 profile.

Disk Status

View current disk usage including internal usage of temp.db.



Data Workbench Profile Dimensions and Metrics

This document describes the profiles with their fields, dimensions, and metrics employed by the data workbench Monitoring Profile.

To monitor data workbench servers, data is collected using a script that parses the Detailed Status while also capturing specific server information. Data workbench server information is then passed to a page tag call for the data workbench Sensor to collect and save to a VSL file.

Profiles Employed by the Data Workbench Monitoring Profile

These profiles provide dimensions and metrics that allow you to view server state and performance data:

- [Dimensions in the Data Workbench Profile Status profile](#) on page 990
- [Dimensions in the Data Workbench Server Status profile](#) on page 992
- [Dimensions in the Data Workbench Historic profile](#) on page 986
- [Metrics in the Data Workbench Historical Monitoring profile](#) on page 994

The Status profiles allow you to see how data workbench is currently performing from an operational perspective. The **Profile Status** profile and the **Server Status** profile gather data from the Detailed Status and the data workbench servers. All gathered data is placed into the **cs-uri-query** field for use.

The **Historic profiles** allow you to assess the impact of configuration and hardware changes using historical data. The historical profile may be the most useful because it allows you to assess the impact of configuration and hardware changes over time.

Dimensions in the Data Workbench Historic profile

The following dimensions are available for use in the data workbench Historic profile.

Dimensions in the Data Workbench Historic profile

The following dimensions are available for use in the data workbench Historic profile.

Block	This is the only countable in this configuration, it is the root for all dimensions. A block can be considered a server. It is using the x-trackingid field.  Note: The block ID is a hash of the server name plus host name, so there will be approximately one block per server per profile.
Ping	This is a countable dimension built off of the Block countable. Each row of data in the profile is a ping from the monitoring agent.
Alert Critical	Numeric Dimension built from the cs-uri-query(ad) value. It is built at the Ping level conditioned that cs-uri-query(a) matches "1".
Alert Down	Numeric Dimension built from cs-uri-query(ac) value. It is built at the Ping level, conditioned that cs-uri-query(a) matches "1".
Alert Warning	Numeric Dimension built from cs-uri-query(ae) value. It is built at the Ping level conditioned that cs-uri-query(a) matches "1".
Any Profile Reprocessing	Numeric Dimension built from cs-uri-query(aa) value. It is built at the Ping level conditioned that cs-uri-query(a) matches "1" and cs-uriquery(k) is not empty.
As of Delay Minutes	cs-uri-query(bi) is placed into the x-as-of-delay-minutes field and rounded to the nearest minute. It is built at the Ping level conditioned that cs-uri-query(a) matches "1".
Capacity Row Percentage	Numeric Dimension built from cs-uri-query(r) value. It is built at the Ping level conditioned that cs-uri-query(a) matches "1" and cs-uriquery(k) is not empty.
Capacity Size Percentage	Numeric Dimension built from cs-uri-query(n) value. It is built at the Ping level conditioned that cs-uri-query(a) matches "1" and cs-uriquery(k) is not empty.
Component Check Success	Simple Dimension built from cs-uri-query(v) value. It is built at the Ping level, and conditioned that cs-uri-query(a) matches "1".
Components in Error	cs-uri-query(ao) is split by the ":" delimiter and copied into the x-components-in-error field. Many to Many Dimension built from the x-components-in-error field. Built at the Ping level.
Detailed Check Seconds	Numeric Dimension built from cs-uri-query(l) value. It is built at the Ping level conditioned that cs-uri-query(k) is not empty.
Detailed Check Success	Simple Dimension built from the cs-uri-query(k) value. It is built at the Ping level conditioned that cs-uri-query(a) matches "1".
Dimension GigaBytes	cs-uri-query(bc) is copied into the x-dimension-gigabytes field. The x-dimension-gigabytes field is user for this Simple Dimension, conditioned on cs-uri-query(a) matching "2".
Disk "x" Used Percentage	These Numeric Dimensions are configured from the cs-uri-query(ah, ai, aj, ak, al) values. They are built at the Ping level and conditioned on cs-uri-query(a) matches "1" and cs-uri-query(k) is not empty.
Estimated Sweep Dekaseconds	The x-estimated-sweep-dekaseconds field is used in this Numeric Dimension. This is the estimated sweep time of the servers divided by ten (reduced resolution of sweep measurement to make dimension more reasonably sized).  Note: This dimension is hidden because it is only useful when averaged into a metric.

Fast Input MegaBytes per Minute	The cs-uri-query(bj) value is used for this dimension. The Last Row for a block is used as the value for the dimension. If the dataset is in Fast Input this Numeric Dimension's value will display the MB per minute at which the system is inputting data.
	 Note: This dimension is hidden because it is only useful when averaged into a metric.
Fast Merge MegaBytes per Minute	The cs-uri-query(bk) value is used for this dimension. The Last Row for a Block is used as the value for the dimension. If the dataset is in Fast Merge This Numeric Dimension's value will display the MB per minute at which the system is merging.
	 Note: This dimension is hidden because it is only useful when averaged into a metric.
Field GigaBytes	The cs-uri-query(bg) value is used for this dimension. The value is divided by 1000 and rounded to the nearest whole number. This Numeric Dimension's value will display the amount of space the Fields in the dataset are using.
	 Note: This dimension is hidden because it is only useful when averaged into a metric.
Group	Simple Dimension built at the Ping level from the cs-uri-query(x) value.
Host	The cs-uri-query(b) value is used for this dimension. The Simple dimension's value is the Last Row for a Block.
Last Ping	the x-unixtime field is copied into x-last-ping and divided by 10 to reduce the cardinality. The Numeric Dimension is built at the Block level and uses the x-last-ping field.
Load Average	This is a Numeric dimension using the Last Row for a given Server's cs-uri-query(i) value. It is conditioned on cs-uri-query(k) not being empty. This dimension is used to calculate the average load on the servers in the system being monitored.
	 Note: This dimension is hidden because it is only useful when averaged into a metric.
Log Reading Percentage	the cs-uri-query(be) value is used for this numeric dimension, built at the Ping level. This dimension is used to calculate the percentage of logs being read.
	 Note: This dimension is hidden because it is only useful when averaged into a metric.
Memory Page File Percentage	This is a Numeric dimension using cs-uri-query(o) value, built at the Ping level. It is conditioned on cs-uri-query(k) not being empty, and cs-uri-query(a) matching "1". This dimension is used to calculate the percent of page file memory usage.
	 Note: This dimension is hidden because it is only useful when averaged into a metric.
Memory Physical MegaBytes Total	This is a Numeric dimension using the cs-uri-query(ag) value, built at the Ping level. It is conditioned on cs-uri-query(k) not being empty, and cs-uri-query(a) matching "1".
	 Note: This dimension is hidden because it is only useful when averaged into a metric.

Memory Physical Percentage	This is a Numeric dimension using the cs-uri-query(ag) value, built at the Ping level. It is conditioned on cs-uri-query(k) not being empty, and cs-uri-query(a) matching "1". This dimension is used to calculate the percent of physical memory usage of each Server.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Memory Query Percentage	This is a Numeric dimension using the cs-uri-query(s) value at the Ping level. It is conditioned on cs-uri-query(k) not being empty and cs-uri-query(a) matching "1". This dimension is used to calculate the percent of query memory usage of each Server.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Network Connections	This is a Numeric dimension using the cs-uri-query(q) value built at the Ping level. It is conditioned on cs-uri-query(k) not being empty and cs-uri-query(a) matching "1". This is used to show the number of network connections there are for a given server.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Output Rows	cs-uri-query(bh) value is divided by 100000 and copied into the x-output-rows field to reduce the size of the dimension. X-output-rows is used in a Numeric Dimension built at the Ping level, conditioned that cs-uri-query(a) matches "2".
Ping Type ID	Simple Dimension built using the cs-uri-query(a) value at the Ping level. This shows what kind of Ping was recorded.
Poll Latency Centiseconds	The cs-uri-query(m) value is divided by 10 to reduce dimension size, and copied into the x-poll-latency-centiseconds field. This is a Numeric dimension built at the Ping level, conditioned that cs-uri-query(k) is not empty, and cs-uri-query(a) matches "1". This dimension is used to calculate the poll latency.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Processing Mode ID	The cs-uri-query(bb) value is used for this Simple Dimension, built at the Ping level. It is conditioned that cs-uri-query(bb) is not empty, and that cs-uri-query(a) matches "2". Processing Mode ID allows one to see what mode of processing the system is in (Fast Input, Fast Merge, Real Time).  Note: This dimension is hidden then re-exposed with friendly values in the client-side dimension Processing Mode.
Profile	The cs-uri-query(ba) value is used for this Simple Dimension, it is built at the Ping level. This dimension displays the name(s) of the profile(s) currently being monitored.
Quick Check Seconds	The cs-uri-query(h) value is used for this Numeric Dimension. It is built at the Ping level conditioned that cs-uri-query(a) matches "1".
Quick Check Success	This is a Simple dimension built from the cs-uri-query(g) value built at the Ping level. It is used to calculate the quick check metric.
Real Time Processing Percentage	Numeric Dimension using the cs-uri-query(t) value built at the Ping level. It is conditioned that cs-uri-query(a) matches "1".

Source Furthest Behind	Simple Dimension built from the cs-uri-query(bl) value built at the Ping level. This dimension displays the when the last contact with a data source occurred.
Temp DB Space Percentage	Numeric Dimension built using the cs-uri-query(an) value, built at the Ping level. It is conditioned that cs-uri-query(k) is not empty, and cs-uri-query(a) matches "1". It is used to calculate the percentage of used Temp DB space on a given server.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Transformation Percentage	the cs-uri-query(bf) value is used for this numeric dimension. It is built at the Ping level. This dimension is used to calculate the percentage of complete data transformation.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Data Workbench Version	The cs-uri-query(ab) value is used for this Simple Dimension. It is built at the Ping level and conditioned that cs-uri-query(ab) is not empty, and cs-uri-query(a) matches "1". This displays the version(s) of data workbench server running on each server.
Data Workbench Version Major	The cs-uri-query(ab) value is split and the major release value is copied into the x-insight-version-major field. It is a Simple Dimension built at the Ping level and conditioned that x-insight-version-major is not empty, and cs-uri-query(a) matches "1".

Dimensions in the Data Workbench Profile Status profile

The following dimensions are available for use in the data workbench Profile Status profile.

Block	This is the only countable in this configuration and exists as the root for all dimensions. A block can be considered a server. It is using the x-trackingid field. The block ID is a hash of the server name plus host name, so there will be approximately one block per server per profile.
As of Delay Minutes	cs-uri-query(bi) is placed into the x-as-of-delay-minutes field and rounded to the nearest minute. As of Delay Minutes is a numeric dimension that takes the Last Row from x-as-of-delay-minutes for a block.
Environment	The cs-uri-query(c) value is used for the Environment ID. The Last Row for a Block is used as the value for the dimension. This Simple Dimension will display the Environment in which your Servers are running (provided it is configured properly). This can be set in the insight_monitor_agent.cfg file
Fast Input MegaBytes per Minute	The cs-uri-query(bj) value is used for this dimension. The Last Row for a block is used as the value for the dimension. If the dataset is in Fast Input this Numeric Dimension's value will display the MB per minute at which the system is inputting data.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Fast Merge MegaBytes per Minute	The cs-uri-query(bk) value is used for this dimension. The Last Row for a Block is used as the value for the dimension. If the dataset is in Fast Merge This Numeric Dimension's value will display the MB per minute at which the system is merging.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Field GigaBytes	The cs-uri-query(bg) value is used for this dimension. The value is divided by 1000 and rounded to the nearest whole number. This Numeric Dimension's value will display the amount of space the Fields in the dataset are using.  Note: This dimension is hidden because it is only useful when averaged into a metric.

Host	The cs-uri-query(b) value is used for this dimension. The Simple dimension's value is the Last Row for a Block.
Last Ping	x-last-ping is x-unixtime divide by 10 (to accommodate Numeric dimensions size constraints). Last Ping is the Last Row for a given Block, and it represents the last time the monitoring agent logged the system health.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Log Reading Percentage	the cs-uri-query(be) value is used for this numeric dimension. It is the Last Row for a given Block. This dimension is used to calculate the percentage of logs being read.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Processing Mode ID	The cs-uri-query(bb) value is used for this Simple Dimension. It the Last Row for a given Block. Processing Mode ID allows one to see what mode of processing the system is in (Fast Input, Fast Merge, Real Time).  Note: This dimension is hidden then re-exposed with friendly values in the client-side dimension Processing Mode.
Processing Stalled	The x-processing-stalled field is created through various conditions to indicate whether the profile is currently running or not. It is a simple dimension.  Note: This dimension works best when there are a large number of input logs to fairly distribute amongst the DPUs. If there is, for example, only one large file loaded per day, then data workbench can appear to "stall" for an hour or more resulting in a false positive reading from this dimension.
Profile	The cs-uri-query(ba) value is used for this Simple Dimension. This dimension displays the name(s) of the profiles currently being monitored.
Source Furthest Behind	The last row of cs-uri-query(bl) is copied into the x-source-furthest-behind field. The Simple Dimension uses the Last Row for a given Block. This dimension displays the when the last contact with a data source occurred.
Transformation Percentage	the cs-uri-query(bf) value is used for this numeric dimension. It is the Last Row for a given Block. This dimension is used to calculate the percentage of complete data transformation.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Time Dimensions	Hour, Day, Week, Month, Hour of Day, and Day of Week are all derived from the x-timestamp field.
Group	Grouping word that gives you another way to filter the resulting dataset. Set in the insight_monitor_agent.cfg file.
Metrics	The following lists the metrics included in the data workbench Profile Monitoring Profile and how they are derived.
As Of Delay Minutes	This metric is the sum of the As Of Delay Minutes for each Block, then divided by the Blocks metric.
As Of Delay Seconds	This metric is the sum of the As Of Delay Seconds for each Block, and divided by the total number of Blocks. (As of Delay Seconds Dimension not configured out of the box)
Blocks	Sum one for each Block.
Fast Input MB per Minute	The sum of Fast Input MegaBytes per Minute for each Block divided by the number of Blocks when Fast Input MegaBytes per Minute is greater than zero.
Fast Merge MB per Minute	The sum of Fast Merge MegaBytes per Minute for each Block divided by the number of Blocks when Fast Merge MegaBytes per Minute is greater than zero.

Field GigaBytes	The sum of Field Gigabytes for each Block divided by Blocks metric.
Last Ping Age	The As Of Time minus Last Ping Time.
Last Ping Time	The sum of Last Ping for each Block divided by Blocks, then multiplied by 10.
Log Reading	Where Log Reading Percentage is greater than zero, Log Reading is the sum of Log Reading Percentage for each Block, divided by Blocks metric, all of which is divided by 10.
Processing Stalled	The sum of the Processing Stalled Dimension for each Block, divided by the Blocks metric.
Transformation	The sum of Transformation Percentage for each Block divided by the Blocks metric, when Transformation Percentage is greater than zero, all divided by 10.

Dimensions in the Data Workbench Server Status profile

The following dimensions are available for use in the data workbench Server Status profile.

Server	Built from the x-trackingid field, this countable dimension represents the Servers currently running data workbench.
Agent Version	The cs-uri-query(af) value is used for this Simple Dimension. It is the Last Nonblank value for a Server. This displays the build date and time for the version(s) of the monitoring agent running.
Any Profile Reprocessing	The cs-uri-query(aa) field is used for this Numeric Dimension, it is the value of the Last Row for a given Server, conditional on cs-uri-query(k) is not empty. This dimension is used to indicate if any profiles are Reprocessing.
Capacity Row Percentage	The cs-uri-query(r) field is used for this Numeric Dimension, it is the value of the Last Row for a given Server, conditional on cs-uri-query(k) is not empty.
Capacity Size Percentage	The cs-uri-query(n) field is used for this Numeric Dimension, it is the value of the Last Row for a given Server, conditional on cs-uri-query(k) is not empty.
Common Name	The sc-ur-query(am) field is used for this Simple Dimension, it is the value of the Last Nonblank value for a given Server. It displays the Common Name of the servers being monitored.
Component Check Success	The cs-uri-query(v) field is used for this Simple Dimension, it is the value of the Last Row for a given Server. This dimension checks on the components of the server to verify they are properly functioning.
Components in Error	A Crossrows transformation takes the Last Row value of the cs-uri-query(ao) and copies it into the x-components-in-error field. This Many to Many dimension displays any components in error on servers being monitored.
Environment	The cs-uri-query(c) value is used for the Environment ID. The Last Row for a Block is used as the value for the dimension. This Simple Dimension will display the Environment in which your Servers are running (provided it is configured properly).
	 Note: This dimension is set in insight_monitor_agent.cfg.
Estimated Sweep Dekaseconds	The x-estimated-sweep-dekaseconds field is used in this Numeric Dimension. This is the estimated sweep time of the servers divided by ten (reduced resolution of sweep measurement to make dimension more reasonably sized).
	 Note: This dimension is hidden because it is only useful when averaged into a metric.
Host	The cs-uri-query(b) value is used for this dimension. The Simple dimension's value is the Last Row for a Block.
Last Ping	x-last-ping is x-unixtime divide by 10 (to accommodate Numeric dimensions size constraints). Last Ping is the Last Row for a given Block, and it represents the last time the monitoring agent logged the system health.

Load Average	This is a Numeric dimension using the Last Row for a given Server's cs-uri-query(i) value. It is conditioned on cs-uri-query(k) not being empty. This dimension is used to calculate the average load on the servers in the system being monitored.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Memory Page File Percentage	This is a Numeric dimension using the Last Row for a given Server's cs-uri-query(o) value. It is conditioned on cs-uri-query(k) not being empty. This dimension is used to calculate the percent of page file memory usage.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Memory Physical MegaBytes Total	This is a Numeric dimension using the Last Row for a given Server's cs-uri-query(ag) value. It is conditioned on cs-uri-query(k) not being empty.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Memory Physical Percentage	This is a Numeric dimension using the Last Row for a given Server's cs-uri-query(ag) value. It is conditioned on cs-uri-query(k) not being empty. This dimension is used to calculate the percent of physical memory usage of each Server.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Memory Query Percentage	This is a Numeric dimension using the Last Row for a given Server's cs-uri-query(s) value. It is conditioned on cs-uri-query(k) not being empty. This dimension is used to calculate the percent of query memory usage of each Server.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Network Connections	This is a Numeric dimension using the Last Row for a given Server's cs-uri-query(q) value. It is conditioned on cs-uri-query(k) not being empty. This is used to show the number of network connections there are for a given server.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Poll Latency Centiseconds	This dimension is used to calculate the poll latency. The cs-uri-query(m) value is divided by 10 to reduce dimension size, and copied into the x-poll-latency-centiseconds field. This is a Numeric dimension which takes the Last Row for a given server.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Quick Check Success	This is a Simple dimension built from the cs-uri-query(g) value of the Last Row for a given Server. It is used to calculate the quick check metric.
Temp DB Space Percentage	The last row of the cs-uri-query(an) value is copied into the x-temp-db-space-percentage field. This is a Numeric Dimension that is used to calculate the percentage of used Temp DB space on a given server.  Note: This dimension is hidden because it is only useful when averaged into a metric.
Insight Version	The cs-uri-query(ab) value is used for this Simple Dimension. It is the Last Nonblank value for a Server. This displays the version(s) of data workbench server running on each server.

Group	Grouping word that gives you another way to filter the resulting dataset.  Note: This dimension is set in insight_monitor_agent.cfg.
Metrics	The following lists the metrics included in the data workbench Profile Monitoring Profile and how they are derived.
Capacity Overall	This is the Capacity Size metric times two plus the Capacity Row metric divided by 3.
Capacity Row	This is the sum of the Capacity Row Percentage for each Server divided by Servers metric.
Capacity Size	This is the sum of the Capacity Size Percentage for each Server divided by the Servers metric.
Component Check	This is the number of Servers where Component Check Success equals one, divided by the Server where Service is DPU or FSU, all multiplied by 100 (to make it a percentage).
Disk "x"	The Disk metrics are calculated by taking the sum of their Disk Used Percentage for each Server, divided by the Servers metric.
Estimated Sweep Minutes	This is the sum of the Estimated Sweep Dekaseconds for each Server, divided by the Servers metric where Estimated Sweep Dekaseconds is greater than zero, all divided by 6.
Last Ping Time	The sum of Last Ping for each Block divided by Blocks, then multiplied by 10.
Load	This is the sum of the Load Average for each Server, divided by the Servers metric.
Memory Page File	This is the sum of the Memory Page File Percentage for each Server, divided by the Servers metric.
Memory Physical	This is the sum of the Memory Physical Percentage for each Server, divided by the Servers metric.
Memory Query	This is the sum of the Memory Query Percentage for each Server, divided by the Servers metric.
Network Connections	This is the sum of the Network Connections for each Server divided by the Servers metric.
Poll Latency Milliseconds	This is the sum of the Poll Latency Centiseconds for each Server, divided by the Servers metric, all of which is multiplied by 10.
Quick Check	This is the number of Servers where Quick Check Success equals one, divided by the Servers metric, all of which is multiplied by 100.
Servers	This is the sum of one for each Server, or the total number of monitored Servers.
Temp DB	This is the sum of the Temp DB Space Percentage for each Server, divided by the Servers metric.

Metrics in the Data Workbench Historical Monitoring profile

The following lists the metrics included in the data workbench Historical Monitoring Profile and how they are derived.

Alert Criticals	The sum of the Alert Critical dimension for each Ping.
Alert Downs	The sum of the Alert Down dimension for each Ping.
Alert Warnings	The sum of the Alert Warning dimension for each Ping.
All Components	The count of Pings where Component Check Success equals "1" divided by the Pings metric multiplied by 100.
As Of Delay Minutes	This metric is the sum of the As Of Delay Minutes for each Ping, then divided by the Pings metric.
Blocks	The sum of one for each Block.
Block All	All Blocks.
Capacity Overall	The Capacity Size Metric times 2 plus the Capacity Row metric, divided by 3.
Capacity Row	The sum of the Capacity Row Percentage dimension for each Ping divided by the Pings metric.
Capacity Size	The sum of the Capacity Size Percentage dimension for each Ping, divided by the Pings metric.
Communications	The number of Pings where Quick Check Success matches "1", divided by the Pings metric.

Detailed Check Seconds	The sum of the Detailed Check Seconds dimension for each Ping where the ping type is "server", divided by the Pings metric.
Dimension GigaBytes	The sum of Dimension Gigabytes for each Ping, divided by the Pings metric.
Disk "x"	The Disk metrics are calculated by taking the sum of their Disk Used Percentage for each Ping, divided by the Pings metric.
Estimated Sweep Minutes	This is the sum of the Estimated Sweep Dekaseconds for each Ping, divided by the Pings metric where Estimated Sweep Dekaseconds is greater than zero, all divided by 6.
Fast Input MB per Minute	The sum of Fast Input MegaBytes per Minute for each Ping divided by the number of Pings when Fast Input MegaBytes per Minute is greater than zero.
Fast Input Mode	Pings where Processing Mode dimension is equal to "Fast input" divided by Pings.
Fast Merge MegaBytes per Minute	The sum of Fast Merge Megabytes per Minute for each Ping, divided by Pings metric.
Fast Merge Mode	Pings where Processing Mode equals "fast merge" divided by the Pings metric.
Field GigaBytes	The sum of Field Gigabytes dimension for each Ping divided by Pings metric.
Load	The sum of the Load Average dimension for each Ping, divided by the Pings metric.
Log Reading	The sum of Log Reading Processing dimension for each Ping, divided by the Pings metric, all divided by 10.
Memory Page	The sum of Memory Page File Percentage for each Ping, divided by the Pings metric.
Memory Physical	The sum of the Memory Physical Percentage dimension for each Ping, divided by the Pings metric.
Memory Query	The sum of the Memory Query Percentage for each Ping, divided by the Pings metric.
Memory Total GB	The sum of Memory Physical MegaBytes Total dimension for each Ping, divided by the Pings metric.
Network Connections	This is the sum of the Network Connections for each Ping divided by the Pings metric.
Pings x Capacity Overall	The Pings metric multiplied by the Capacity Overall metric.
Poll Latency Milliseconds	The sum of the Poll Latency Centiseconds dimension for each Ping, divided by the Pings metric, all multiplied by 10.
Query Running	The sum of one for each Ping where Estimated Sweep Dekaseconds is greater than "0", divided by the Pings metric where Ping Type equals "server".
Quick Check Seconds	The sum of Quick Check Seconds for each Ping where Ping Type is equal to "server", divided by the Pings metric.
Output Rows	The sum of Output Rows dimension for each ping divided by the Pings metric, multiplied by 100000.
Real Time Mode	The number of Pings where Processing Mode dimension equals "real time", divided by the Pings metric, all multiplied by 100.
Reprocessing Mode	100 minus the number of Pings where Processing Mode equals "real time" divided by the Pings metric, multiplied by 100.
Stalled	The sum of the Processing Stalled dimension in the Insight Dimensions in the Data Workbench Profile Status profile on page 990 profile.
Temp DB	The sum of Temp DB Space Percentage for each Ping, divided by the Pings metric.
Transformation	The sum of Transformation Percentage for each Ping divided by the Pings metric all divided by 10.

Data Workbench Dashboard User Guide

Navigation title:

The Adobe Data Workbench dashboard is an interactive analytics application that brings web-based analytics and thin-client capabilities to Data Workbench. It lets you easily visualize Data Workbench data and analyze business trends from a browser-based interface customized for your specific needs.

Using the Data Workbench dashboard, business professionals can access web-based views of analytic information that is easy-to-understand and dynamically segmented across critical data feeds. These dashboards are optimized to consolidate business intelligence and analytical data that has been captured and given context through Adobe Data Workbench.

Business Opportunities

Navigation title:Business Opportunities

Executives, analysts, and marketing professionals can use the data workbench dashboard for a variety of tasks.

- Executives can draw meaningful insights with cursory views of easy-to-understand visualizations.
- Data analysts can track progress from enterprise data repositories and analytic systems.
- Marketing professionals can follow and react to fluctuations in channel and web campaigns.
- Your team can utilize timely analysis to improve its performance with informed, strategic decision-making and improved execution across the enterprise.
- Multiple groups within your organization achieve business optimization as they benefit from data access, enhance the analytic processes, and participate in a collaborative environment.

Features

Navigation title:Features

The Adobe data workbench dashboard provides these features.

- Perform on-the-fly analysis of Adobe Data workbench datasets.
- Drill-down on and slice data with a simple mouse-driven interface.
- Create and save multiple analytic dashboards.
- Share analytics dashboards with others.
- Access dashboards created by other users.
- Export data for use in other applications, such as Microsoft® Excel®.

This guide provides information about the functionality and user interface features delivered with the Adobe data workbench dashboard. Because your implementation of the dashboard can be fully customized, your features and functionality may differ from what is documented in this guide.

Glossary

Navigation title:Glossary

The most commonly used terminology and features of Adobe data workbench, including the dashboard.

Profiles

Profiles define a dataset. A profile includes data about a particular subject, as well as the breakdown of that data into user-friendly analytic components. Profiles are created and managed within Adobe data workbench and made accessible through the dashboard application. Profiles are typically designed to fulfill a specific purpose (such as analyzing marketing and website traffic) and define a set of analytic components known as Dimensions, Metrics, and Filters.

Dimensions

Dimensions are categories of like data types. For example, the Days of Week dimension is composed of the following data elements: Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, and Saturday. Specifically, dimensions detail what is being measured. In the dashboard, dimensions are defined within Adobe data workbench

and made accessible through the dashboard application. Each dimension has a unique name within the dashboard interface.

Metrics

Metrics are quantitative objects defined by a quantifiable expression. For example, Page Views per Session is derived from the expression from the count of Page Views divided by the count of Sessions. Metrics answer the question of “how many?” In the dashboard, metrics are defined within Adobe data workbench and made accessible through the dashboard application. Each metric has a unique name within the dashboard interface.

Filters

Filters allow you to segment your data in a specific way. Filters will reduce your data to a subset of the whole result, and are often combined to provide a unique perspective on the data. In the dashboard, filters are first defined within Adobe data workbench and made accessible through the dashboard application. Each filter has a unique name within the dashboard interface.

Getting Started

Navigation title: Getting Started

To begin using the Adobe data workbench dashboard, follow these steps.

1. Install supported browser.

The dashboard is a web-based application that operates using the latest web browser technology. It can be used on any operating system provided that the operating system is equipped with a compatible web browser. Because the dashboard takes advantage of features not supported in older versions of web browsers, more recent browsers are required for an optimal experience when using the application.

While many variations of browsers may work, only the following web browsers are officially supported:

- Google Chrome 15 and above (available at <https://www.google.com/chrome/>)
- Microsoft Internet Explorer 9 and above (available at <http://www.microsoft.com/windows/ie/>)
- Mozilla Firefox 3.6 and above (available at <http://www.mozilla.org/en-US/firefox/all-older.html>)
- Safari 5.1 and above (available at <http://www.apple.com/safari>)

Beyond a compatible web browser, no additional software needs to be downloaded in order to use the dashboard.

2. Request an account.

Contact your administrator to create an account for you. You will be provided with a user name, temporary password, and instructions to access the dashboard.

3. Access the Adobe data workbench dashboard.

The dashboard is a rich internet application that requires only a web browser and appropriate network connection to operate. You only need to ensure that you have a compatible browser available and the ability to connect to the dashboard server in order to access the dashboard.

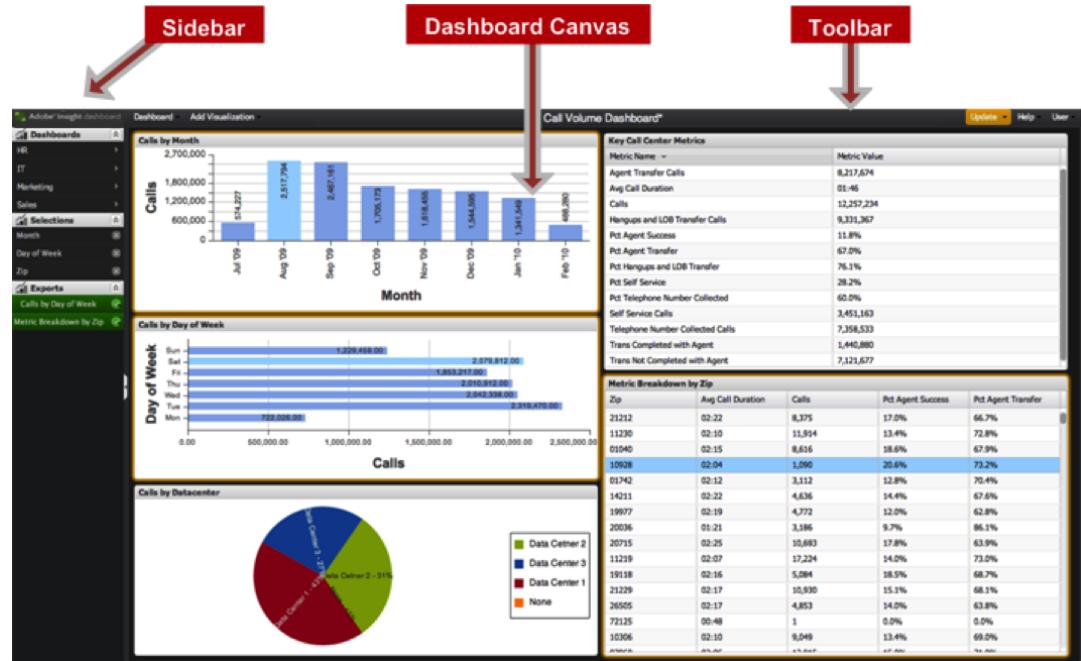
4. Log in to the dashboard.

- a) Open a compatible web browser.
- b) Browse to the URL where the Adobe Data workbench dashboard is hosted. If you are unsure of the URL where dashboard is hosted, please contact your administrator.
- c) Enter your username and password.
- d) Click **Log in**.

Dashboard User Interface

Navigation title: Dashboard User Interface

The data workbench dashboard includes a Sidebar and Toolbar with feature menus and a Dashboard Canvas where you display data visualizations.



Dashboard Canvas

The Dashboard Canvas is the area of the interface used to create and interact with visualizations.

Term	Definition
Visualizations	Rendered data defined graphically based on a set of configured criteria. Multiple visualizations make up a dashboard.

Sidebar

The left sidebar provides quick access and visibility for defined menus:

Menu	Definition
Quick-access Menu	A fly-out menu to quickly open dashboards.
Selections Menu	Lists any current selections within the dashboard, providing additional details and the ability to remove selections at any time.
Exports Menu	Lists any pending or completed exports, providing the ability to retrieve the export when complete.

Toolbar

The top toolbar provides access to these specific menus. It also displays the title of the dashboard currently open.

Menu	Definition
Dashboard Menu	A drop-down menu for opening, creating, saving, and exporting dashboards.

Menu	Definition
Add Visualization Menu	A drop-down menu for adding visualizations to the dashboard.
Dashboard Title	Displays the title of the current dashboard.
Update Menu	Clicking the Update indicator will update the current dashboard, taking into account any new selections and/or configuration changes. The menu provides the ability to adjust the Query-to parameter, along with some advanced functions for administrators.
Help Menu	This drop-down menu is used to access the User's Guide, FAQ, and About section of the application.
User Menu	A drop-down menu for logging out, accessing account settings, and accessing the administration area (for administrators only).

Visualizations

Navigation title:Visualizations

Visualizations are elements added to the dashboard canvas and configured to display various metric and dimension-based data.

All dashboards are composed of one or more visualizations. Each visualization can be created, removed, resized, and reconfigured independently of all other visualizations on the dashboard canvas.

Visualizations are also interactive, allowing users to quickly segment data by making selections on one or more data elements within the visualization. Selections made within one visualization will dynamically apply filters in real-time to other visualizations on the canvas. This renders the same data across all visualizations on the canvas.

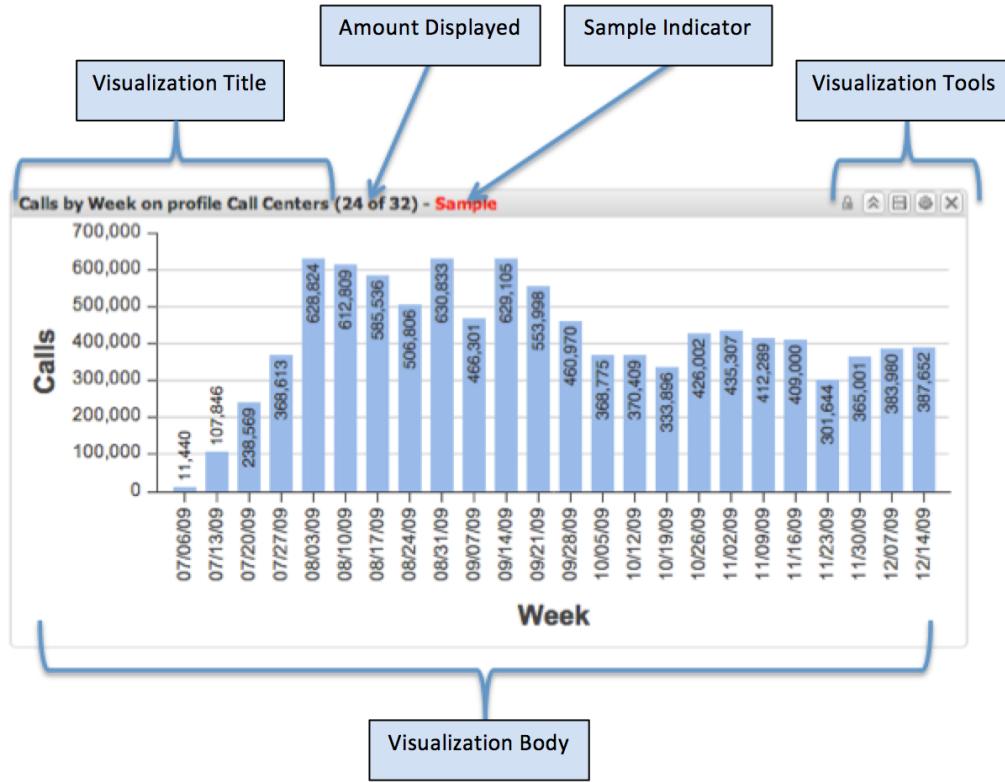
There are eight different types of visualizations. Each one can be added, resized, configured, and removed independently of any other visualization. Visualizations display data defined in data workbench by a data workbench architect.

The eight types of visualizations available include:

- Column Charts
- Bar Charts
- Line Charts
- Tables
- Metric Legends
- Pie Charts
- Scatter Plots
- Rich Text

Visualization User Interface

The header portion of the visualization contains the visualization title and visualization tools, which vary depending on the type and state of the visualization. The body of the visualization contains the content, and is dependent on the type and configuration of the visualization being displayed. The visualization tools only appear when the mouse enters the visualization window. Otherwise they are hidden.



- Visualization Title. Describes this visualization. The title is automatically generated or manually overridden with a custom title.
- Amount Displayed. For the dimension being visualized, displays the amount of data shown versus the total amount available.
- Sample Indicator. Shown when the data visualized is a sample and not a 100% complete query result.
- Visualization Tools. Performs specific operations on the visualizations. Tools available depend on the visualization type, state, and current user permissions.
- Visualization Body. Displays the data of the visualization as configured. This area is interactive and depends on the type of visualization being displayed.

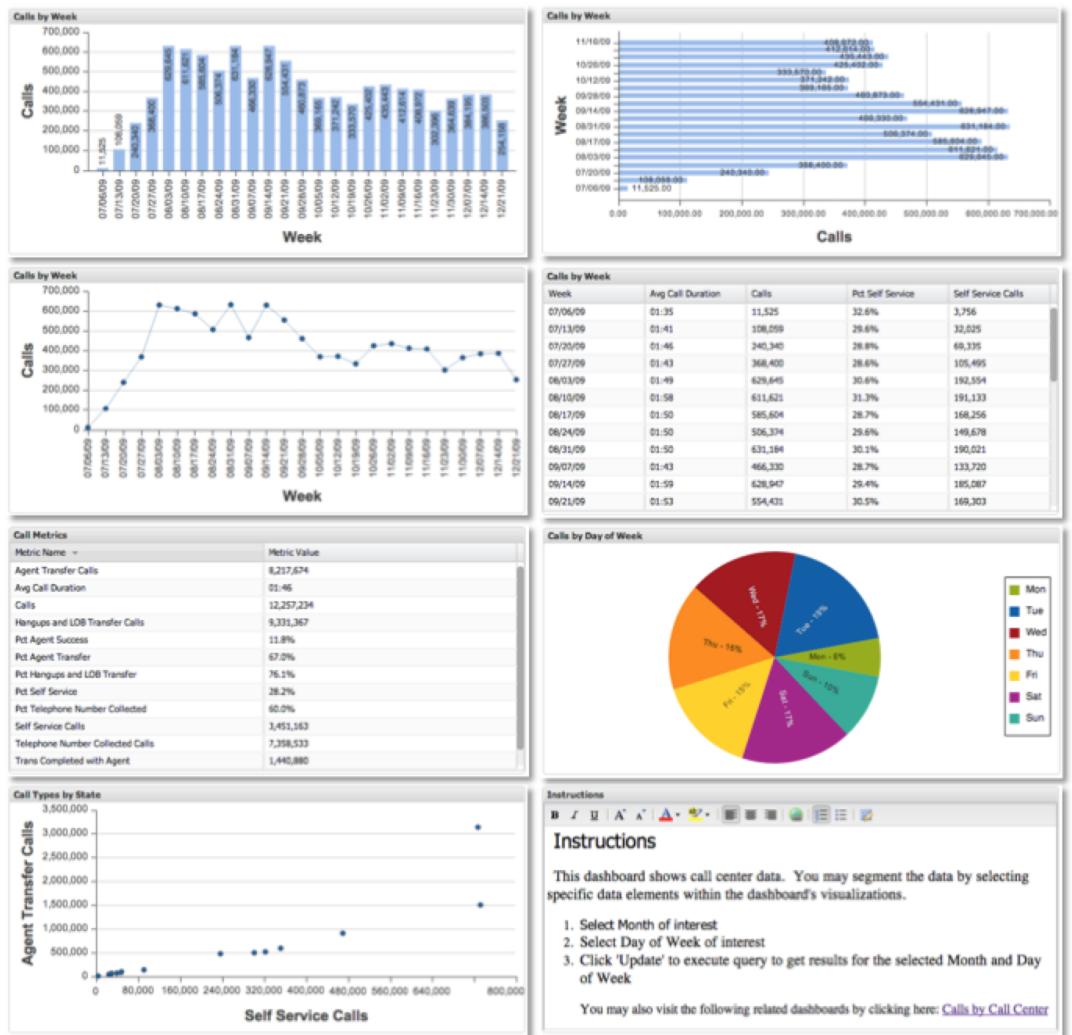
Visualization Types

Navigation title: Visualization Types

Introduces the eight different visualization types.

When selecting a visualization, it is important to carefully consider the dimensions and metrics you are visualizing and how the data can be visualized most effectively. For example, some visualizations have limitations on the quantity of data that can be displayed at once.

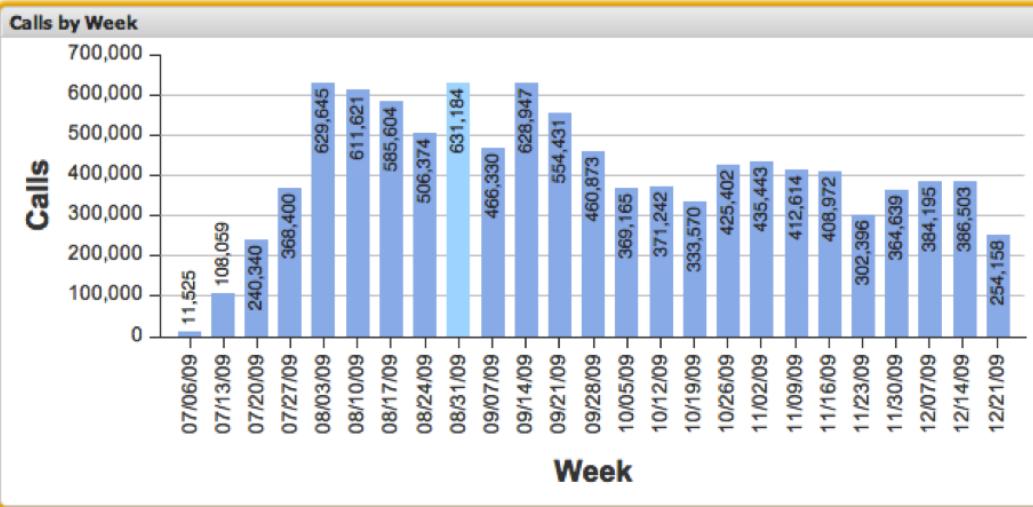
There are eight different visualization types that can be used to visualize and analyze your data: Bar Charts, Column Charts, Line Charts, Pie Charts, Tables, Metric Legends, Scatter Plots, and Rich Text.



Column Charts

Navigation title: Column Charts

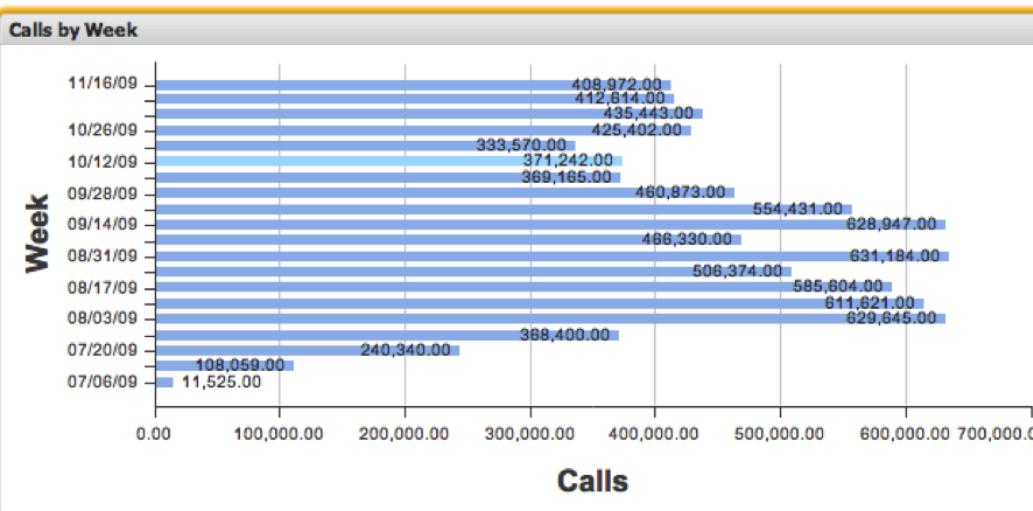
Column chart visualizations allow you to visualize metric and dimension data with each data element represented as a vertical column on the chart. Column charts allow up to 200 unique data elements to be displayed on the chart at once. The dimension values are labeled along the x-axis, while the metric increments are labeled along the y-axis. The column for each dimension value is labeled with its exact metric value.



Bar Charts

Navigation title: Bar Charts

Bar chart visualizations allow you to visualize metric and dimension data with each data element represented as a horizontal bar on the chart. Bar charts allow up to 200 unique data elements to be displayed on the chart at once. The dimension values are labeled along the y-axis, while the metric increments are labeled along the x-axis. The bar for each dimension value is labeled with its exact metric value.

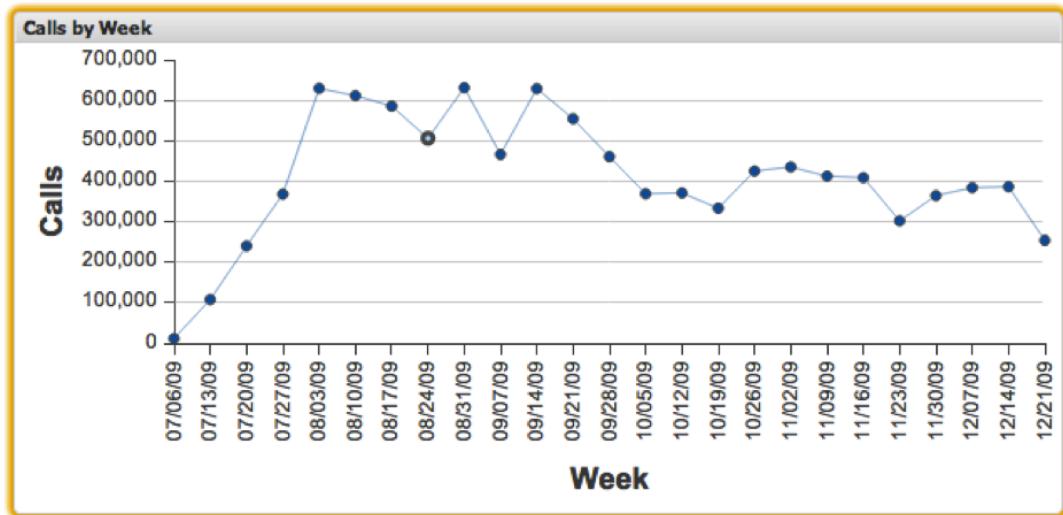


Line Charts

Navigation title: Line Charts

Line chart visualizations allow you to visualize metric and dimension data with each data element represented as a point on an x-y axis, and lines connecting adjacent points. Line charts allow up to 200 unique data elements to be displayed on the chart at once. The dimension values are labeled along the x-axis, while the metric increments are labeled along the y-axis. Placing the mouse over a point in the line chart will display the exact metric value at that point.

Selections can be made in line charts by clicking a particular point. Pressing and holding the Alt key will retain selections as you click additional items. Pressing and holding the Alt key when clicking will also toggle the value of a selection. This is useful for removing just one selection from a group of already selected items.



Tables

Navigation title:Tables

Tables are the most detail-oriented visualizations available in the dashboard application. Tables allow you to visualize metric and dimension data with each data element represented as a row in a table. One column holds the dimension values, while the other columns contain the corresponding metric values. Multiple metrics can be displayed, and there is no restriction on the number of dimension values that can be displayed in a table. All values for a selected dimension that exist in the data workbench data will be displayed in the table.

Column order can be rearranged by clicking-and-dragging one column header to another location in the column header area. Sorting can be performed at any time by clicking on the appropriate column header. Clicking the column header again will reverse the sort order.

Selections can be made in tables by clicking a particular row. Pressing and holding the **Control** key will retain selections as you click additional items. Pressing and holding the **Control** key when clicking will also toggle the value of a selection. This is useful for removing just one selection from a group of already selected items. A range of rows can be selected by first clicking a row to select it, then pressing the **Shift** key while clicking the last row.

Calls by Week				
Week	Avg Call Duration	Calls	Pct Self Service	Self Service Calls
07/06/09	01:35	11,525	32.6%	3,756
07/13/09	01:41	108,059	29.6%	32,025
07/20/09	01:46	240,340	28.8%	69,335
07/27/09	01:43	368,400	28.6%	105,495
08/03/09	01:49	629,645	30.6%	192,554
08/10/09	01:58	611,621	31.3%	191,133
08/17/09	01:50	585,604	28.7%	168,256
08/24/09	01:50	506,374	29.6%	149,678
08/31/09	01:50	631,184	30.1%	190,021
09/07/09	01:43	466,330	28.7%	133,720
09/14/09	01:59	628,947	29.4%	185,087
09/21/09	01:53	554,431	30.5%	169,303

Metric Legends

Navigation title:Metric Legends

Metric legend visualizations allow you to display metric totals in a tabular format. The first column holds the metric name, while the second column holds the metric value. A scrollbar will be displayed in the event that all the rows cannot be displayed on the window at once.

Selections cannot be made within the metric legend since no specific dimension value is being displayed.

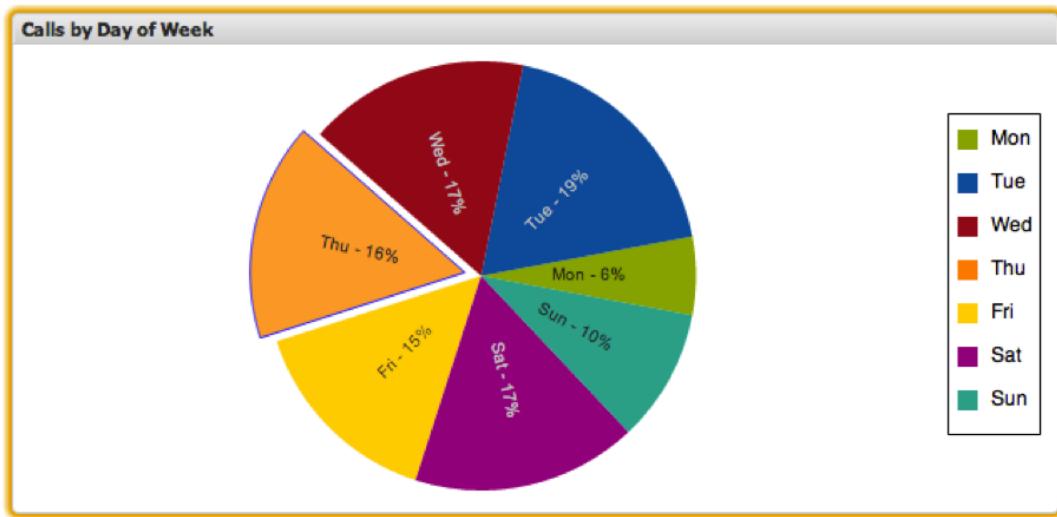
Call Metrics	
Metric Name	Metric Value
Agent Transfer Calls	8,217,674
Avg Call Duration	01:46
Calls	12,257,234
Hangups and LOB Transfer Calls	9,331,367
Pct Agent Success	11.8%
Pct Agent Transfer	67.0%
Pct Hangups and LOB Transfer	76.1%
Pct Self Service	28.2%
Pct Telephone Number Collected	60.0%
Self Service Calls	3,451,163
Telephone Number Collected Calls	7,358,533
Trans Completed with Agent	1,440,880

Pie Chart

Navigation title:Pie Chart

Pie chart visualizations allow you to display relative percentages of the metric values for a limited number of dimensions. A pie chart allows up to 20 unique data elements to be displayed on the chart at once. Each wedge of the pie represents a dimension in the dataset, along with its percentage. Each percentage is calculated relative to the metric values of the other dimensions displayed on the pie chart, which may not necessarily be the entire dataset.

Selections can be made in pie charts by clicking any particular wedge.

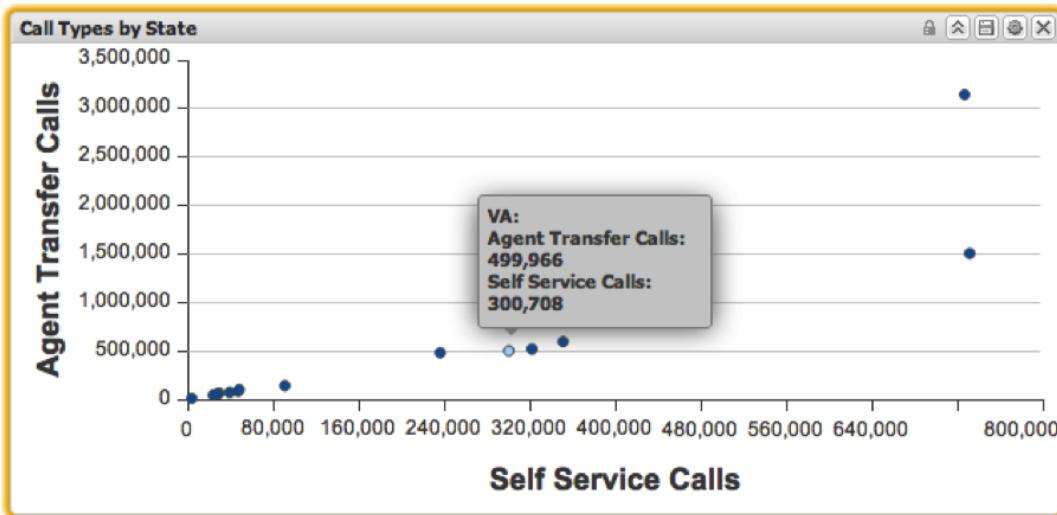


Scatter Plot

Navigation title:Scatter Plot

Scatter plot visualizations allow you to visualize two metrics against a dimension, with each data element represented as a point on an x-y axis. Scatter plots allow up to 200 unique data elements to be displayed on the chart at once. The first metric's increments are labeled along the y-axis, while the second metric's increments are labeled along the x-axis. Placing the mouse over a point in the line chart will display the exact value at that point.

Selections can be made in scatter plots by clicking a particular point. Pressing and holding the **Alt** key will retain selections as you click additional items. Pressing and holding the **Alt** key when clicking will also toggle the value of a selection. This is useful for removing just one selection from a group of already selected items.



Rich text

Navigation title:Rich text

Rich text visualization allows you to embed text-based information within the dashboard. This is useful for providing descriptions, hyperlinks, or instructional information. The text formatting toolbar at the top of the Rich Text window allows you to manipulate the format of text within the window body.

The screenshot shows a 'Rich Text' window with a title bar labeled 'Instructions'. Below the title bar is a toolbar containing various text formatting icons: bold (B), italic (I), underline (U), font size (A up, A down), font color (A red), font style (italic, underline), alignment (center, left, right), and a link icon. The main body of the window contains the following text:

Instructions

This dashboard shows call center data. You may segment the data by selecting specific data elements within the dashboard's visualizations.

1. Select Month of interest
2. Select Day of Week of interest
3. Click 'Update' to execute query to get results for the selected Month and Day of Week

You may also visit the following related dashboards by clicking here: [Calls by Call Center](#)

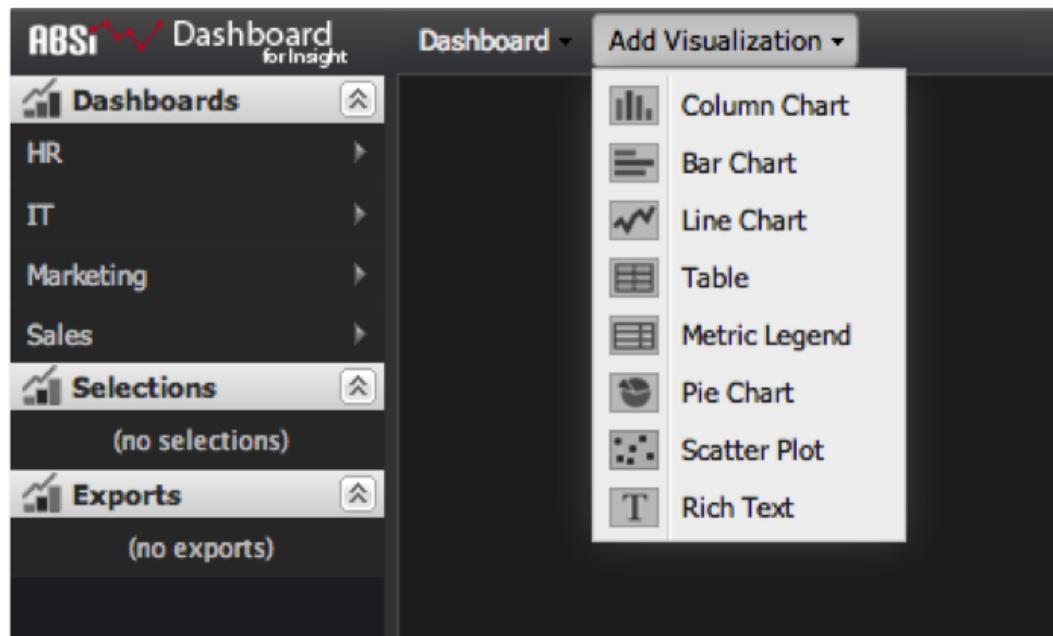
To add text to the rich text window, place the cursor within the body of the rich text window and begin typing. To edit text within the rich text window, use the text formatting toolbar at the top of this window.

Creating Visualizations

Navigation title:Creating Visualizations

Steps to add a visualization to the dashboard canvas.

1. Go to the **Add Visualization** menu at the top of the dashboard interface. You will see a list of available visualizations that can be created.



-
2. Select one of these menu items to create a new empty visualization and add it to the dashboard canvas.

The visualization's configuration dialog opens automatically to allow you to configure the visualization. See [Configuring Visualizations](#) on page 1007 for details on how to configure the visualization.



Note: Cancelling the initial configuration window removes the visualization from the dashboard canvas.

Configuring Visualizations

Navigation title:Configuring Visualizations

Shows how to configure Title, Profile, Dimension, Metric, Filter, Display Top, Sort By, and Time Period.

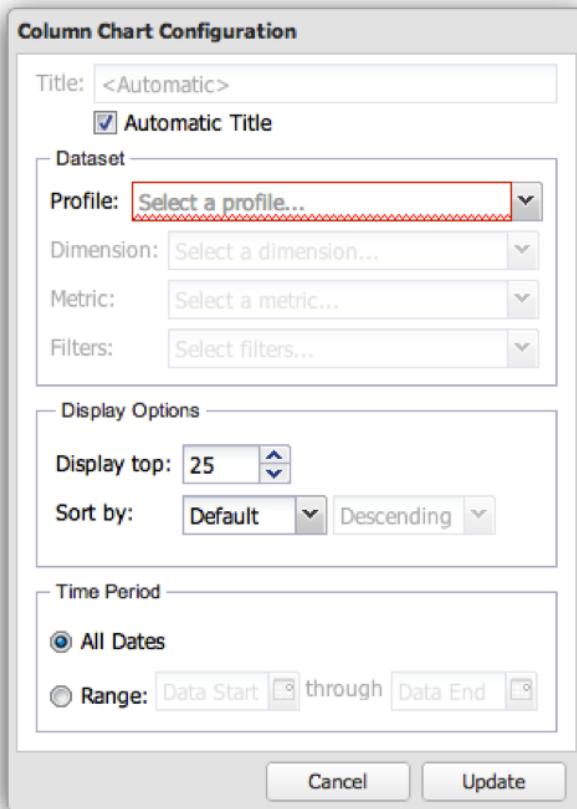
Each visualization on the dashboard canvas has its own configuration. When a visualization is first added to the dashboard canvas, its configuration window will appear automatically. Once configured, the visualization can be modified at any time by clicking the gear icon in the upper right-hand portion of the visualization window.



Note: Configuration options vary slightly depending on the type of visualization being displayed.

Visualization Title

This field allows you to customize the title displayed at the top of the visualization. By default the title is set to **Automatic Title**, which will automatically generate a title for the visualization window. By clearing the **Automatic Title** button, you may place any title in this field. (This field applies to all visualizations.)



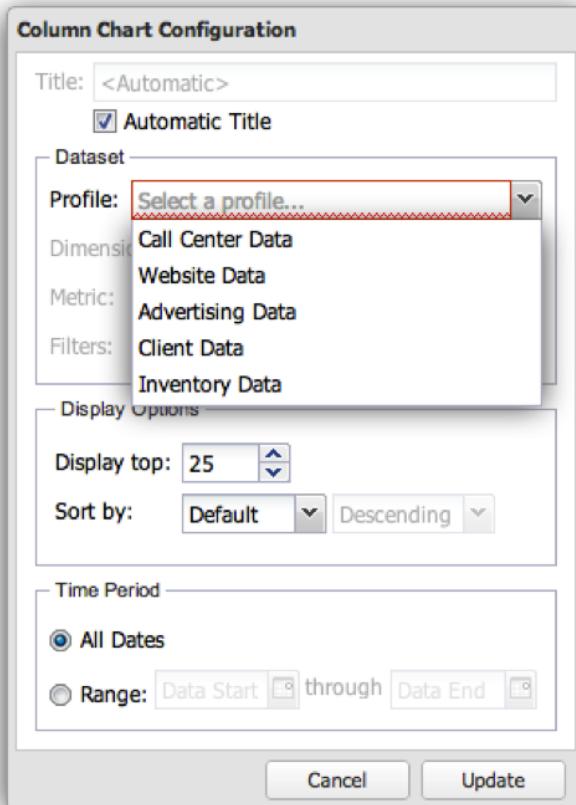
Profile

This field allows you to select which profile you wish to visualize data from. Clicking on the dropdown menu will provide you with a list of profiles for which you have access. (This field is not applicable for Rich Text visualizations.)

Profiles are data sets defined within Data workbench that contain data about a certain domain, along with the dimensions, metrics, and filters that accompany the data. A profile is often designed to fulfill a specific purpose (such as marketing or website traffic).



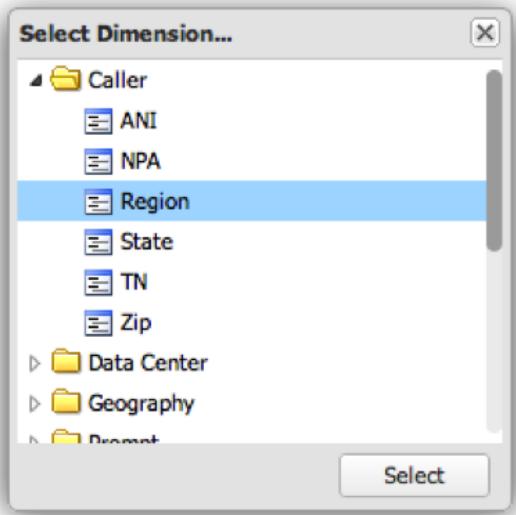
Note: You can see only the profiles for which you have been granted access. For more information see Access Controls.



Dimension

Lets select the dimension you would like to visualize. The list is populated from the list of dimensions available from the profile selected in the Profile field. Click on the desired dimension and then click the Select button. (This field is not applicable for Metric Legends and Rich Text visualizations.)

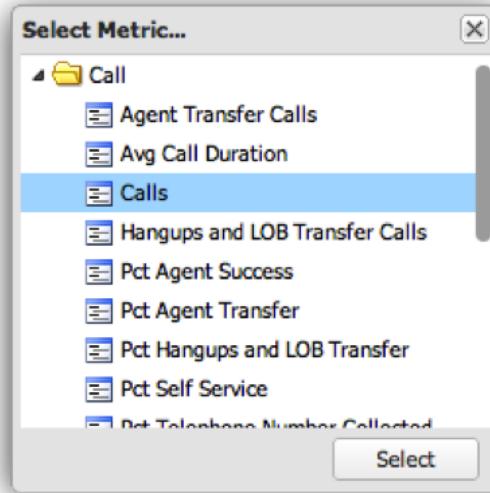
Dimensions are categories of like data types. For example, the Days of Week dimension is composed of the following data elements: Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, and Saturday. Dimensions show what is being measured.



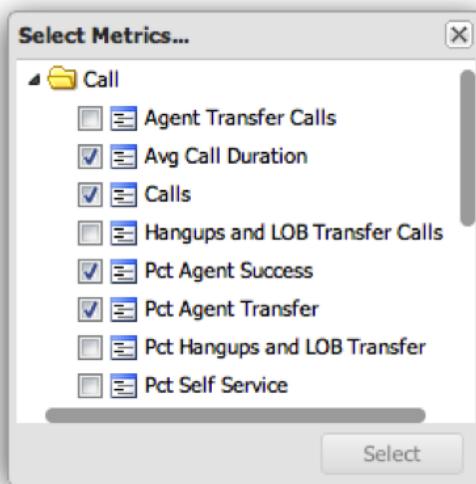
Metric(s)

Lets you select the metrics to visualize. Metrics are quantitative objects and are defined by some quantifiable expression. For example, Page Views per Session is derived from the expression of the count of Page Views divided by the count of Sessions. Metrics answer the question of "how many?"

Single-metric visualizations have a single-metric selection window:



Multi-metric visualizations have a multi-metric selection window:

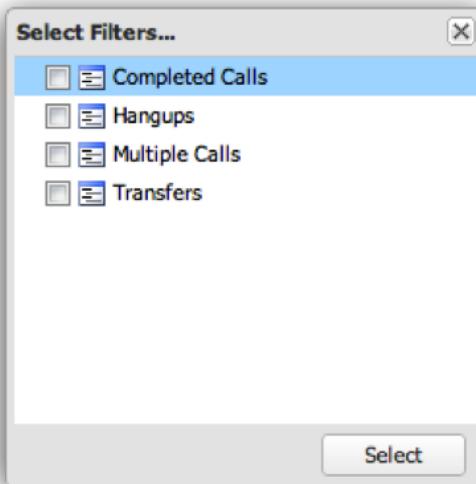


Filters

Select the filters you would like to apply to your visualization. The filter selection window allows you to select multiple filters from the filter list. The list is populated from the list of filters available from the profile selected in the Profile field. Click on the desired filter and then click **Select**.

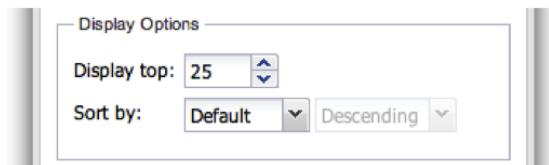


Note: Filters applied here are only applied to their corresponding visualization, not the entire dashboard. This is useful for comparing the results of two different visualizations with different filters applied.



Display Tops

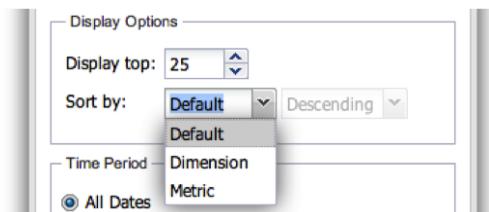
Visualizations in the dashboard are not designed to display the entirety of the data. Rather, they allow you to specify the number of dimension records you would like to display on the visualization. This displays the top number of dimensions depending on the sort-by value given below. (This field is not applicable for Tables, Metric Legends, and Rich Text visualizations.)



Sort By

This allows you to specify how the data should be sorted when it is displayed within the visualization. (This field is not applicable for Tables, Metric Legends, and Rich Text visualizations.) There are multiple sorting options:

- **Default** – Return the data unsorted based on the sort order stored in data workbench. This is the option to use for time-based data such as hour, day, week, or month.
- **Dimension** – Sort the data based on the alphanumeric dimension value.
- **Metric** – Sort the data based on the metric value and is good for quickly visualizing the top dimensions.
- **Descending** – Sort the data in descending order.
- **Ascending** – Sort the data in ascending order.



Time Period

This visualization allows you to specify the desired start and/or end date of the data to display within the visualization.

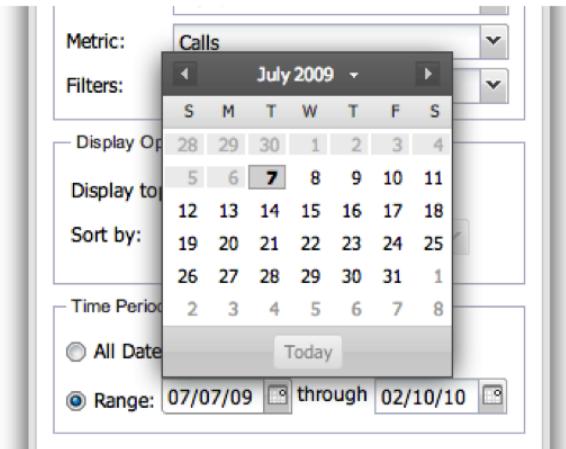
Selecting **All Dates** displays the entire date range available in the profile.

Selecting **Range** displays only the data that falls within a specified range. To enter the date range, you can type in the beginning and/or end date, or use a calendar input by selecting the calendar icon.

(This field is not applicable for Rich Text visualizations.)



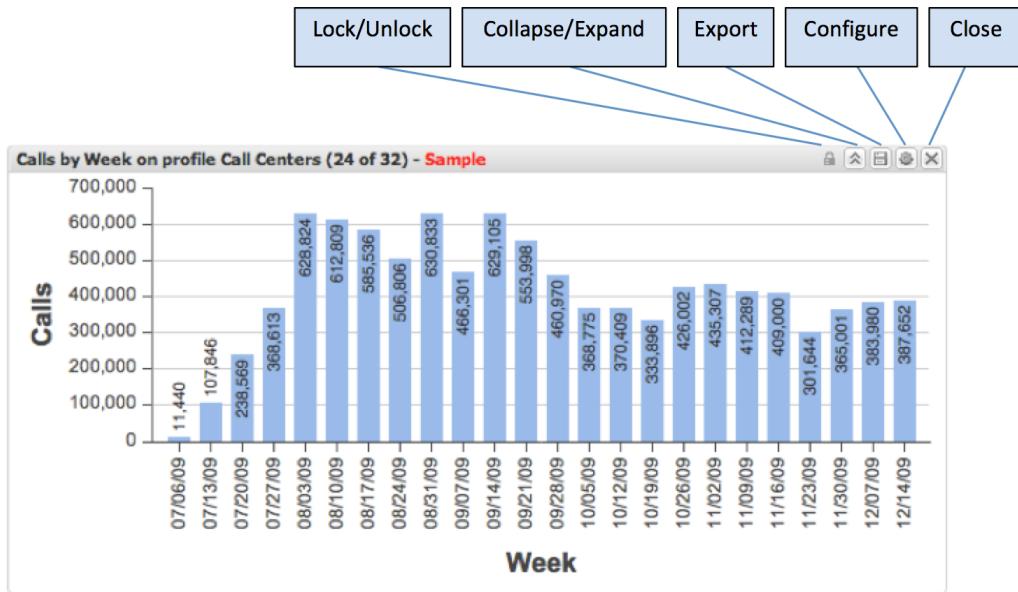
Note: Date ranges applied here are only applied to their corresponding visualization, not the entire dashboard. This is useful for comparing the results of two different visualizations with different date ranges applied.



Manipulating Visualizations

Navigation title: Manipulating Visualizations

The visualization's tool menu allows you to quickly perform certain operations on the visualization. Placing your cursor over any of the icons within the visualization tool menu will provide an explanation of that icon's function.



Locking and Unlocking Visualizations

Navigation title: Locking and Unlocking Visualizations

Locks/unlocks the current visualization.

Visualizations can be locked to preserve selections and help guide analysis for viewers of a particular dashboard. Locks can be toggled by using the lock icon within the visualization's tool menu. For all visualizations except Rich Text and Metric Legends, applying a lock will prevent users from altering selections within that visualization. Any selections that were made prior to locking the visualization will always be applied, and no adjustments to the selections can be made while the visualization is locked.

If no selections were made within the visualization before locking, the lock will serve to prevent users from making any selections within that visualization. Regular users and Administrators can dynamically toggle visualization locks, but Read-only users will always be restricted to any locks that have been placed on visualizations.

Locking visualizations is useful for guiding analysis by pre-selecting items that the entire dashboard should be filtered on. This provides an on-screen visual cue of what's being filtered without allowing it to be altered. Locking is also useful for reducing the potential for confusion by only allowing users to select on visualizations that are relevant to their analysis.

Collapsing and Expanding Visualizations

Navigation title: Collapsing and Expanding Visualizations

Collapses/expands the current visualization.

Visualizations can be collapsed and expanded at any time. Collapsing a visualization will reduce the height of the visualization to just the title and tools portion, hiding the visualization body. Expanding a collapsed visualization will restore its size to its original height.

Collapsing a visualization is particularly useful for preserving a large number of visualizations within a dashboard while taking up minimal space. Users can expand a collapsed visualization to see its contents only when necessary.

Exporting Visualizations

Navigation title: Exporting Visualizations

Exports data from the current visualization.

Data from visualizations can be quickly exported using the export icon at the top-right of the visualization window. Clicking this icon will initiate an export on the server and allow you to download the data in several formats. See [Exporting Data](#) on page 1033 for more information.

Configuring Visualizations

Navigation title: Configuring Visualizations

Visualizations can easily be configured and reconfigured.

Visualizations can be configured and reconfigured by clicking the gear icon at the top-right of the visualization window. This will bring up the visualization's configuration window to allow adjustments to be made to that visualization. For details on configuring visualizations. See [Configuring Visualizations](#) on page 1007.

Removing Visualizations

Navigation title: Removing Visualizations

You can remove a visualization from the dashboard canvas at any point.

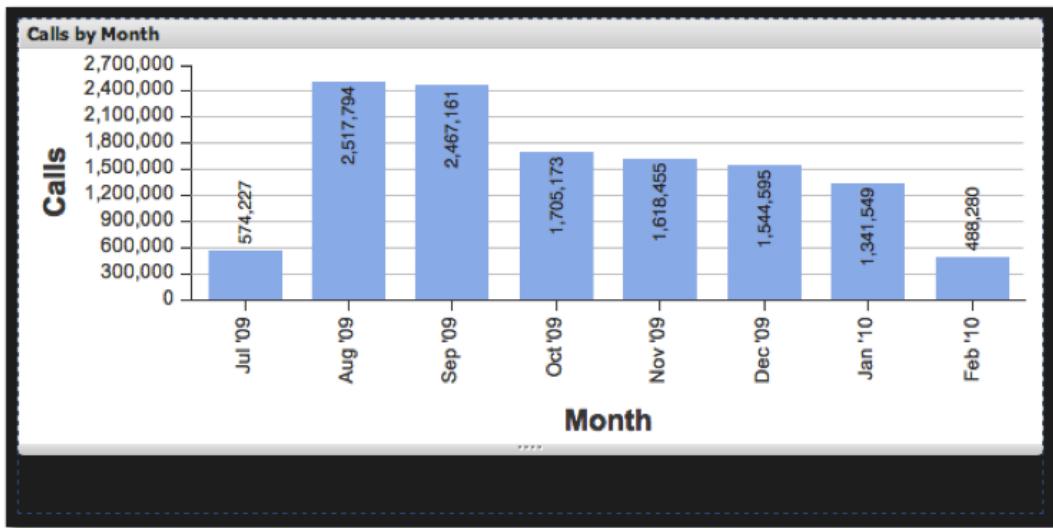
This is done using the **X** icon at the upper right-hand portion of the visualization window. This will remove it from the dashboard canvas, but will not impact any saved versions of dashboards unless the current dashboard is saved. Any selections that were present within the visualization being closed will be removed from the dashboard's selections.

Sizing Visualizations

Navigation title: Sizing Visualizations

In addition to expanding and collapsing a visualization, you can also resize them.

Just use the resize handle at the bottom of the visualization window. By hovering over the bottom edge of a visualization window, you can click-and-drag to make the window taller or shorter on the dashboard canvas. Any other visualizations within the dashboard will automatically be repositioned to fit the new size of the resized visualization. The width of visualization is fixed relative to the width of the browser and cannot be manipulated.

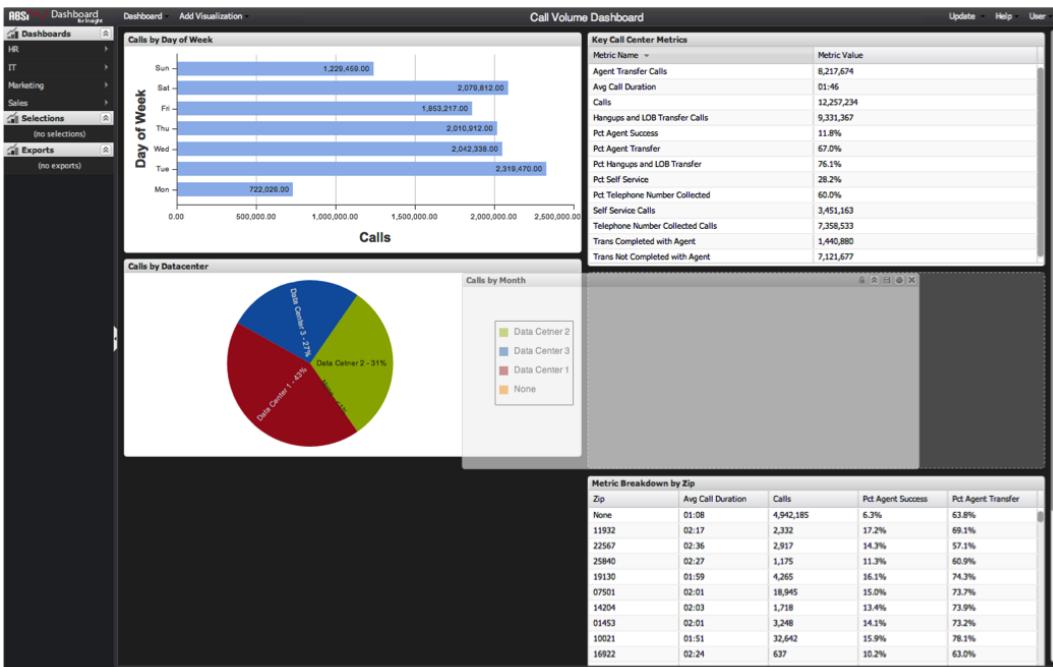


Moving Visualizations

Navigation title:Moving Visualizations

Visualizations can be easily moved and rearranged within the dashboard canvas.

To move a visualization, simply click the visualization's title bar, drag the visualization, and drop it in another location on the dashboard canvas. The dashboard canvas will assist in positioning the visualization by making room for it at the desired location. Existing visualizations will automatically be shifted up or down to make room for the visualization being moved to avoid overlapping.



Dashboards

Navigation title:Dashboards

Dashboards are created to visualize metrics and provide interactive analytic capability with data. By clicking on items within a dashboard, you can quickly and easily segment the data to derive information from your analysis. An unlimited number of dashboards can be created, either for short-term on-the-fly analysis or for long-term dashboard capability. Dashboards can be kept private or easily shared with other dashboard users. You can also export data from any dashboard for use in other tools such as Microsoft® Excel™.

Opening and Viewing Dashboards

Navigation title:Opening and Viewing Dashboards

Dashboards can be opened in one of three ways: through the quick-access menu, through the Dashboard Browser, or by using a dashboard hyperlink. Though dashboards can only be opened one at a time in the dashboards interface, you may create a separate browser tab or window to view multiple dashboards at the same time.

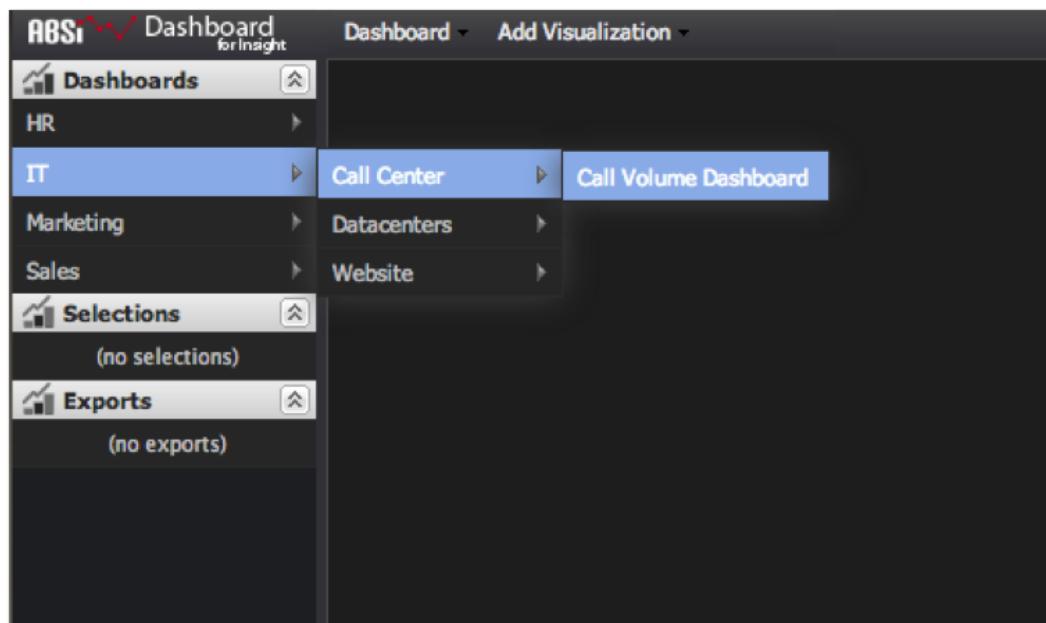
Quick-Access Menu

Navigation title:Quick-Access Menu

One of the easiest methods to open a dashboard is by using the quick-access menu on the left-hand navigation panel of the dashboard interface.

Placing the cursor over a folder menu item (denoted by the right-facing arrow) will expand its contents and allow you to quickly navigate through the dashboard folder hierarchy to find your desired dashboard. Clicking on the dashboard's menu item (denoted by the dashboard title and no right-facing arrow) will open the dashboard for viewing and analysis.

This method is ideal when you know the name and location of the dashboard you would like to open. If you're unsure of the name and/or location of the dashboard, the Dashboard Browser will help you locate and open the dashboard.



Dashboard Browser

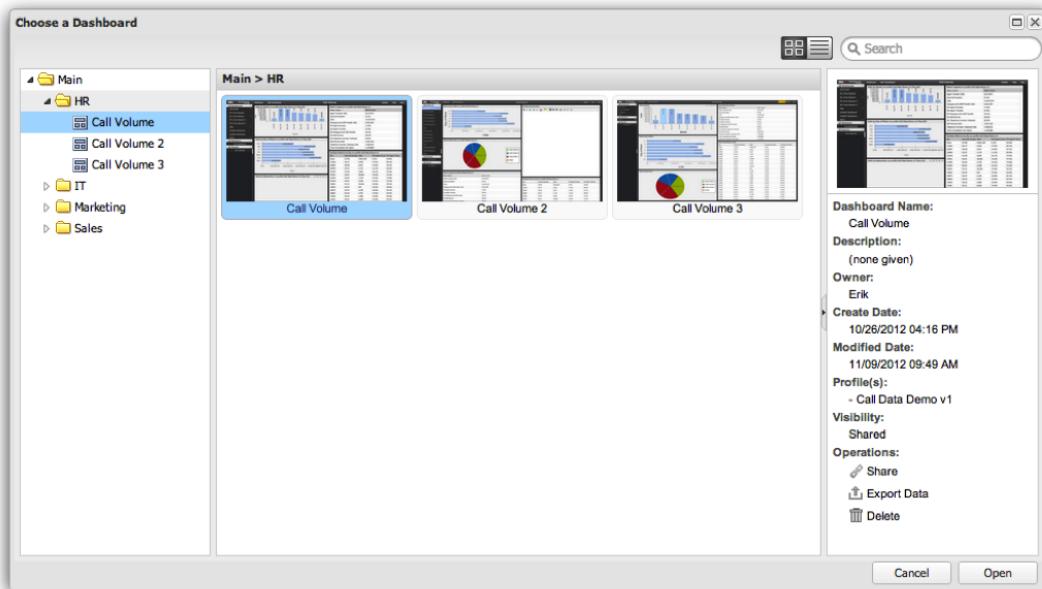
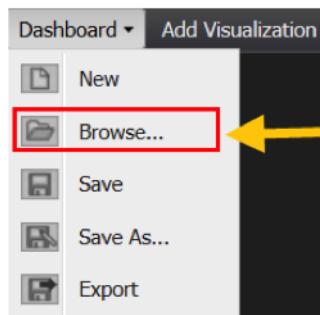
Navigation title:Dashboard Browser

The Dashboard Browser facilitates browsing, searching, and opening dashboards that you and other members of your organization have created and saved.

You can browse and access all dashboards that you have permissions to view (for more information on access controls, see section Access Controls). The Dashboard Browser works well when you need to find a dashboard but are uncertain about which one you need or where it resides. The Dashboard Browser is also useful for getting additional details and performing special functions on a given dashboard.

Opening the Dashboard Browser

When you log into Adobe Data workbench dashboard, the Dashboard Browser will appear by default. You can also access the Dashboard Browser at any time by clicking on the Dashboard menu in the toolbar and selecting **Browse....**



Dashboard Browser Views

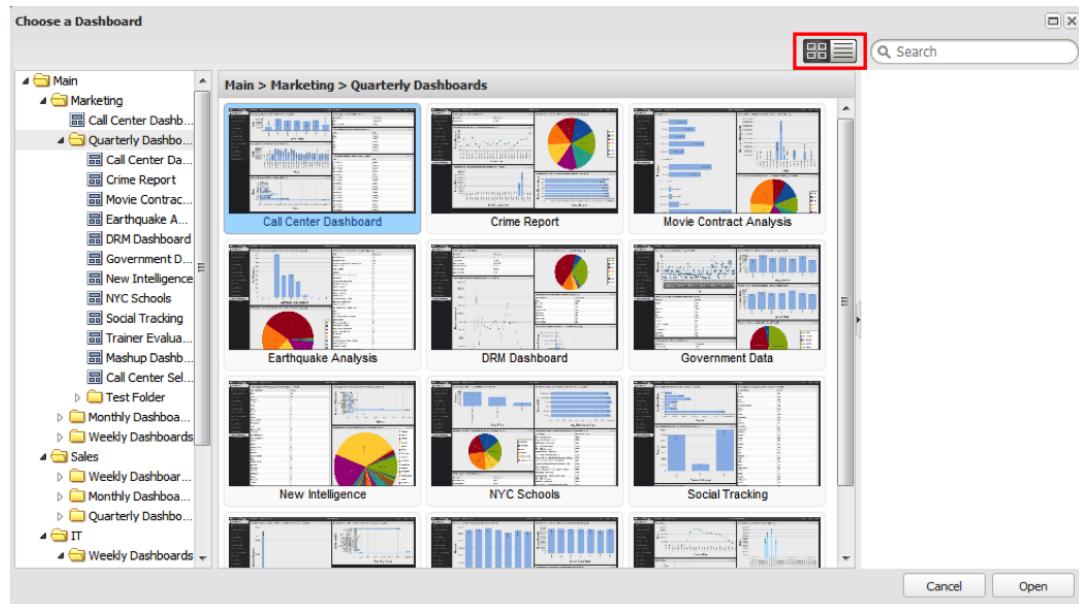
Navigation title: Dashboard Browser Views

Explains the Thumbnail View and the List View.

There are two views that can be used to navigate dashboards within the Dashboard Browser: the Thumbnail View and the List View. The Thumbnail View provides thumbnails of dashboards as you browse, while the List View is a more detail-oriented table-based view of the available dashboards.

Thumbnail View

The Thumbnail View provides a gallery of thumbnails that correspond with the dashboards for which you have access. If a thumbnail displays a default icon, please contact your administrator to have a thumbnail established for that dashboard. The icons to the left of the search bar allow you to toggle between Thumbnail and List Views.



In the thumbnail view, the left-hand navigation panel of the Dashboard Browser displays the folder hierarchy that organizes the storage of dashboards. To view the contents of a folder, click on the folder to show its contents as thumbnails in the center panel of the Dashboard Browser. If there are no dashboards saved in that level of the folder, the message “no dashboards in this location” will display in the center panel of the window.

You can also explore subfolders by clicking on the arrow to the left of the folder of interest. This will expand a list of subfolders and dashboards within the folder you selected. By selecting a folder that contains at least one dashboard, the center panel will display a thumbnail for each of the dashboards located at the level of that selection. The center panel title will also change to indicate the path of the selected folder.

You can then select the dashboard of interest by clicking on it. Once you have clicked on a dashboard thumbnail, the right-hand frame will populate with the details about the dashboard. The dashboard details contain a thumbnail

view of the selected dashboard, its name, a brief description, the owner, creation date, last modification date, profile(s) used to compile the data, visibility controls, and offers a selection of operations.

List View

The List View provides table-based information on the dashboards for which you have access. Each row in the List View's table represents a unique dashboard. Clicking on column headers will allow you to sort the table by that column in either descending or ascending order.

The screenshot shows a software interface titled "Choose a Dashboard". The main title bar says "Main > Marketing > Quarterly Dashboards". Below this is a table with columns: Name, Owner, Create Date, Modify Date, and Visibility. The table lists several dashboards under the "Quarterly Dashboards" folder. One specific dashboard, "Call Center Dashboard", is highlighted with a blue selection bar. The table data is as follows:

Name	Owner	Create Date	Modify Date	Visibility
Main				
Marketing				
Call Center Dashboard	Erik	06/12/12 6:52 PM	06/12/12 6:52 PM	Private
Quarterly Dashboards				
Call Center Dashboard	Erik	06/08/12 6:14 PM	06/08/12 6:30 PM	Private
Crime Report	Erik	06/08/12 6:14 PM	06/08/12 6:30 PM	Private
Movie Contract Analysis	Erik	06/08/12 6:14 PM	06/08/12 6:30 PM	Private
Earthquake Analysis	Erik	06/08/12 6:14 PM	06/08/12 6:30 PM	Private
DRM Dashboard	Erik	06/08/12 6:14 PM	06/08/12 6:30 PM	Private
Government Data	Erik	06/08/12 6:15 PM	06/08/12 6:30 PM	Private

Clicking on a dashboard entry will display the dashboard's details in the right-hand panel of the dashboard browser. To switch between views, re-select the desired view option.

Searching within the Dashboard Browser

Navigation title: Searching within the Dashboard Browser

The Dashboard Browser also allows you to perform searches to find existing dashboards.

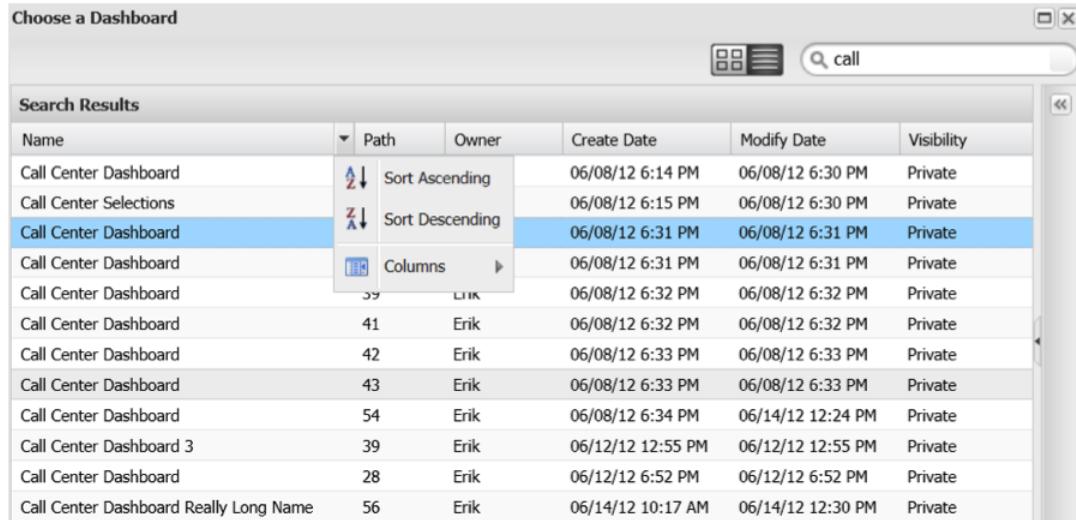
To perform a search, use the search box in the Dashboard Browser's toolbar. Your search results will appear in the Dashboard Browser as you type. You can clear a search at any time by deleting the search text or clicking the X icon in the search box.

The screenshot shows the "Choose a Dashboard" window with a search bar containing "Volume". The search results table shows four dashboards: "Call Volume 2", "Call Volume", "Call Volume 3", and "Call Volume Dashboard". The "Call Volume Dashboard" is selected and highlighted with a blue selection bar. To the right of the table is a detailed sidebar with the following information:

- Dashboard Name: Call Volume Dashboard
- Description: (none given)
- Owner: Gary
- Create Date: 11/05/2012 10:53 AM
- Modified Date: 11/08/2012 03:19 PM
- Profile(s): - Call Data Demo v1
- Visibility: Shared
- Operations:
 - Share
 - Export Data
 - Delete

At the bottom right of the sidebar are "Cancel" and "Open" buttons.

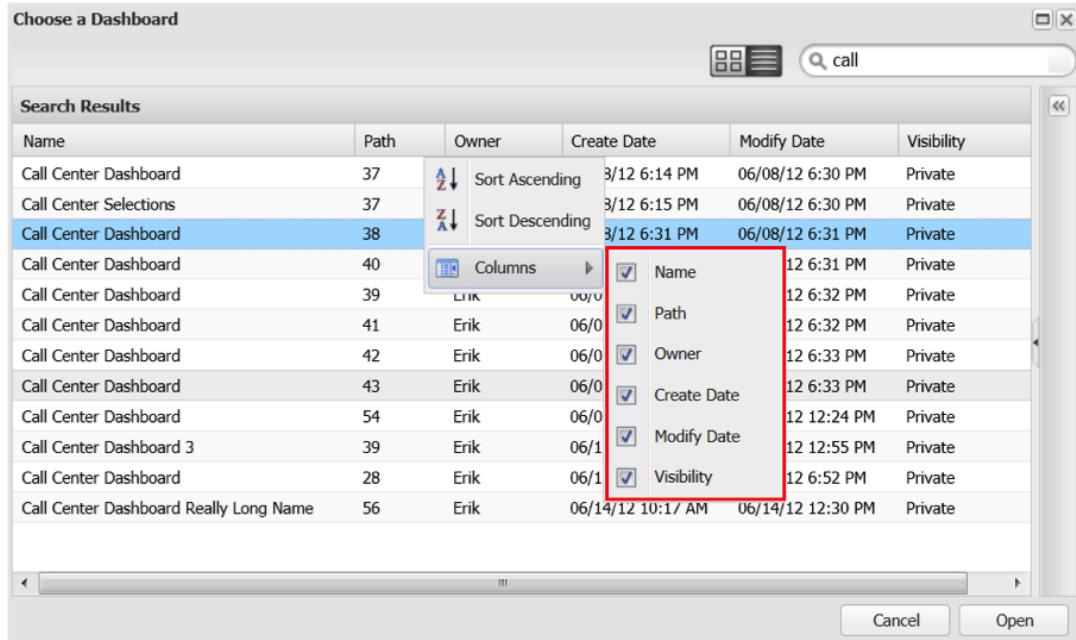
You can also sort your results using the sort functions in each column header. Clicking on a column header will toggle sorting on that column. You can also access the sorting options using the drop-down menu on each column header.



A screenshot of a software interface titled "Choose a Dashboard". Below it is a search bar with the word "call". The main area is titled "Search Results" and contains a table with the following columns: Name, Path, Owner, Create Date, Modify Date, and Visibility. The "Name" column has a dropdown arrow icon with sorting options: "Sort Ascending" (with icons for Z down and A up) and "Sort Descending" (with icons for Z up and A down). The "Path" column has a dropdown arrow icon with the option "Columns" followed by a right-pointing arrow. The table lists several dashboard entries, such as "Call Center Dashboard" and "Call Center Selections", along with their details like creation and modification dates and visibility levels.

Name	Path	Owner	Create Date	Modify Date	Visibility
Call Center Dashboard			06/08/12 6:14 PM	06/08/12 6:30 PM	Private
Call Center Selections			06/08/12 6:15 PM	06/08/12 6:30 PM	Private
Call Center Dashboard			06/08/12 6:31 PM	06/08/12 6:31 PM	Private
Call Center Dashboard			06/08/12 6:31 PM	06/08/12 6:31 PM	Private
Call Center Dashboard			06/08/12 6:32 PM	06/08/12 6:32 PM	Private
Call Center Dashboard	41	Erik	06/08/12 6:32 PM	06/08/12 6:32 PM	Private
Call Center Dashboard	42	Erik	06/08/12 6:33 PM	06/08/12 6:33 PM	Private
Call Center Dashboard	43	Erik	06/08/12 6:33 PM	06/08/12 6:33 PM	Private
Call Center Dashboard	54	Erik	06/08/12 6:34 PM	06/14/12 12:24 PM	Private
Call Center Dashboard 3	39	Erik	06/12/12 12:55 PM	06/12/12 12:55 PM	Private
Call Center Dashboard	28	Erik	06/12/12 6:52 PM	06/12/12 6:52 PM	Private
Call Center Dashboard Really Long Name	56	Erik	06/14/12 10:17 AM	06/14/12 12:30 PM	Private

This **Columns** menu also allows you to modify visible columns in the dashboard list view. Here you can toggle the visibility of columns by selecting or deselecting the column(s) of choice.



A screenshot of the same "Choose a Dashboard" interface, but the "Columns" menu is open over the "Path" column. A red box highlights the "Columns" dropdown and the list of visible columns: Name, Path, Owner, Create Date, Modify Date, and Visibility. Each item in the list has a checkbox next to it, indicating whether it is currently visible. The table data remains the same as in the previous screenshot.

Name	Path	Owner	Create Date	Modify Date	Visibility
Call Center Dashboard	37		06/12 6:14 PM	06/08/12 6:30 PM	Private
Call Center Selections	37		06/12 6:15 PM	06/08/12 6:30 PM	Private
Call Center Dashboard	38		06/12 6:31 PM	06/08/12 6:31 PM	Private
Call Center Dashboard	40		06/08/12 6:31 PM	12 6:31 PM	Private
Call Center Dashboard	39		06/08/12 6:32 PM	12 6:32 PM	Private
Call Center Dashboard	41	Erik	06/08/12 6:33 PM	12 6:33 PM	Private
Call Center Dashboard	42	Erik	06/08/12 6:33 PM	12 6:33 PM	Private
Call Center Dashboard	43	Erik	06/08/12 6:34 PM	12 12:24 PM	Private
Call Center Dashboard	54	Erik	06/12/12 12:55 PM	12 12:55 PM	Private
Call Center Dashboard 3	39	Erik	06/12/12 6:52 PM	12 6:52 PM	Private
Call Center Dashboard	28	Erik	06/14/12 10:17 AM	06/14/12 12:30 PM	Private
Call Center Dashboard Really Long Name	56	Erik			

Details Panel

The Dashboard Details Panel displays a dashboard's detailed information along with a thumbnail version of the dashboard. Information found in the panel includes the dashboard's name, a brief description, the owner and

date of creation, the last date in which it was modified, and the profile(s) that were used to create the dashboard. The dashboard's visibility (whether it is private or shared) is also available in this panel.

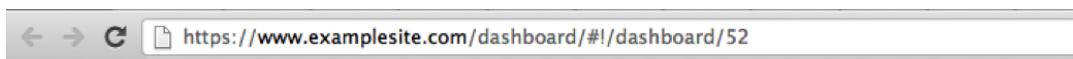
Dashboard Hyperlink

Navigation title: Dashboard Hyperlink

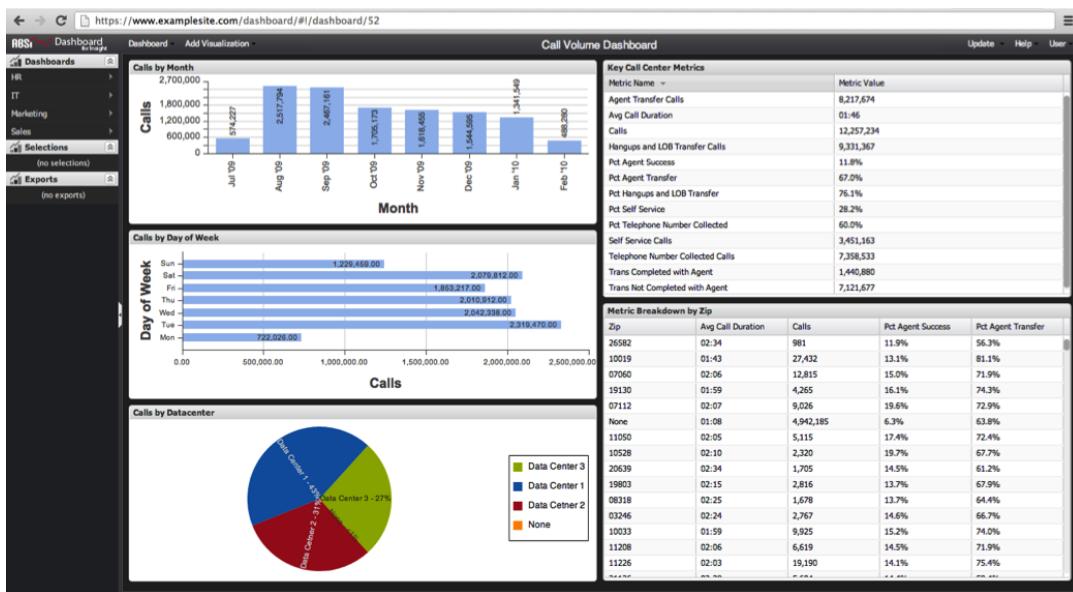
The third way that a dashboard can be opened is by using a dashboard's hyperlink.

Each dashboard has a unique hyperlink that can be used to open the dashboard through the browser's address bar. Dashboard hyperlinks can also be used for bookmarking and sharing via e-mail.

If you have a dashboard's hyperlink, simply enter it into the browser's address bar and navigate to the link. You'll be directed to the dashboard site and prompted to log in (if not already logged in). Once logged in, the dashboard will load in the interface.



Note: You will be prompted with a message if the dashboard no longer exists or you do not have proper permissions to access the dashboard.

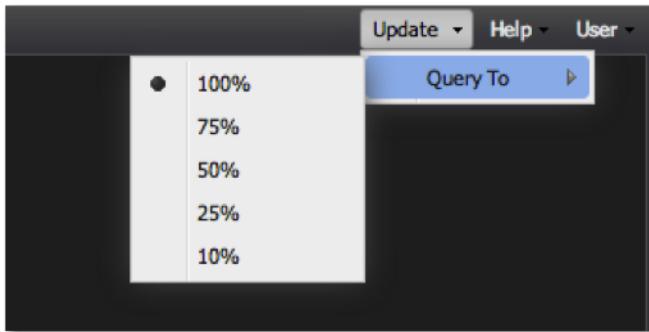


Query-To Parameter

Navigation title: Query-To Parameter

The dashboard allows you to visualize samples of data from Adobe data workbench versus querying your dataset to completion.

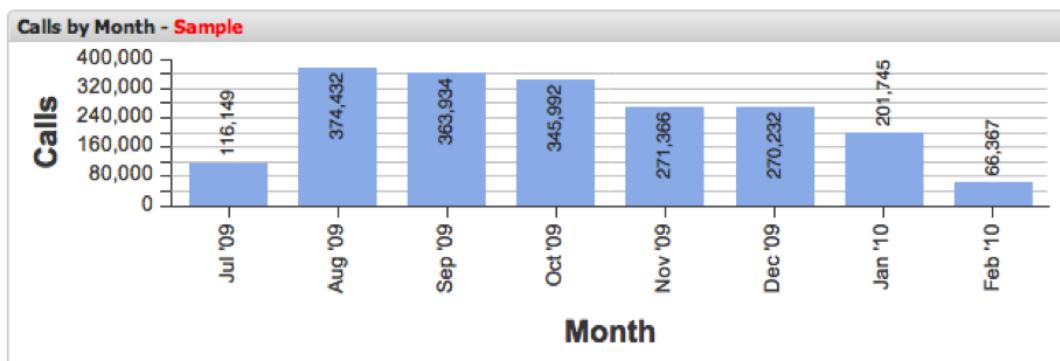
Since sample results are returned quickly, using a lower **Query To** percentage feature facilitates quick dashboard creation and analysis until a full result is returned. The **Query To** parameter can easily be adjusted at any time using the **Query To** menu within the **Update** menu of the toolbar.



Since running queries to 100 percent completion can take several minutes, it is recommended that you adjust the **Query To** parameter to a lower value while building dashboards, or adding and configuring visualizations. It is also recommended to lower this value when fine-tuning your selections within a dashboard until you are sure you are ready to run the query to 100 percent completion.



Note: An indicator will be shown in the header for each visualization that does not have a 100 percent complete query result.



Creating a Dashboard

Navigation title:Creating a Dashboard

Creating a dashboard is recommended even for short-term, ad-hoc analytical needs.



Note: Read-only users cannot create dashboards. This section applies only to regular users and administrators.

Users can decide to create dashboards for several reasons:

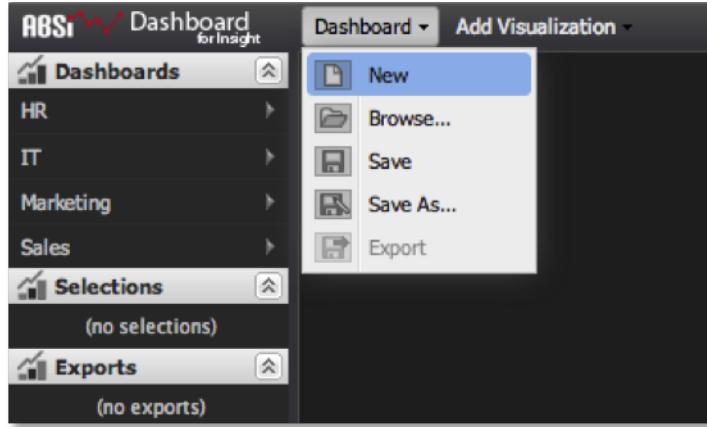
- A new dashboard can be started from scratch for on-the-fly analysis with no intent of reusing or sharing the dashboard.
- A new dashboard can be created for the purpose of performing your own personal analysis that you would like to save and reuse, but not share.
- A new dashboard can be created, saved, and shared for you and the rest of the dashboard user population to access. Whatever the case, each scenario starts at the same point: a blank dashboard canvas.



Note: Before starting to build out your dashboard, it's a good idea to reduce your Query-to percentage to something low, such as 10 percent or 25 percent. This will pull samples of data from Data workbench

much quicker than performing a complete query. Since these sampled results return much more quickly, it provides ideal responsiveness while framing out your dashboard and analysis. Once you're ready to run queries to completion, you can update the query-to parameter to 100 percent. For adjusting query completion, see the [Query-To Parameter](#) on page 1021.

To create a new dashboard, select **New** under the Dashboard menu.



You will be presented with a blank dashboard canvas that is ready to have visualizations added and configured based on your analytic needs. As you work, nothing will be updated on the server until you save.

Next, decide what kind of data you want to display and how you want to display it. It generally helps to start with table visualizations to see the raw data, and then build out other charts to suit. For details on how to add and configure visualizations, see [Creating Visualizations](#) on page 1006. After adding and configuring visualizations to build out the dashboard, you will end up with the following:



From this point you can simply perform your analysis and discard the dashboard, or you can choose to save the dashboard to the server for reuse and/or sharing. For information on how to interact with a dashboard to perform analysis, see the section [Making Selections within the Dashboard](#) on page 1027.

Saving a Dashboard

Navigation title: Saving a Dashboard

1. To save a new dashboard, select either **Save** or **Save As...** from the Dashboard Menu at the top of the dashboard interface. For a new dashboard, both options have the same result.

This launches the **Save New Dashboard** window where you can enter a title, a brief description, and location to store the dashboard.



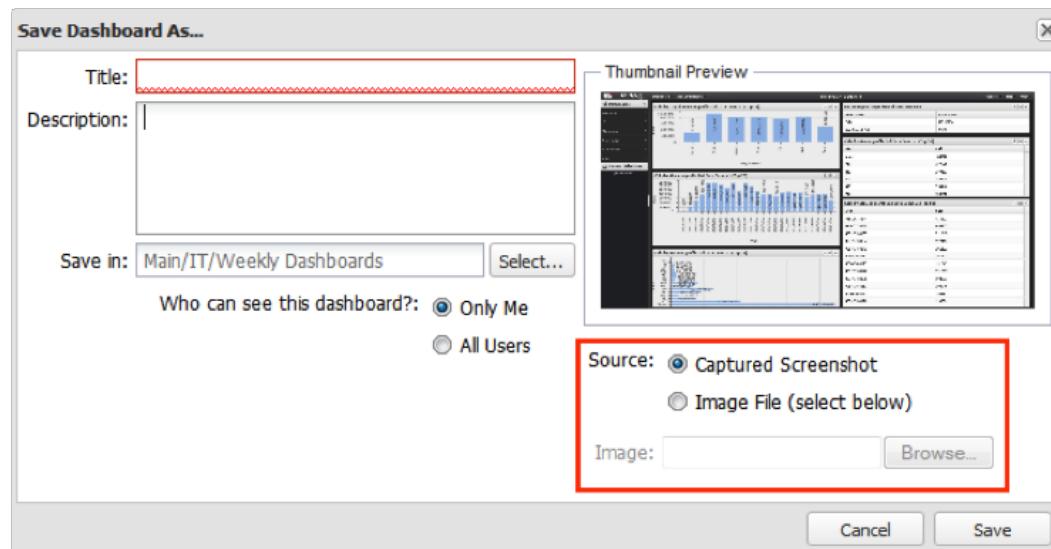
Note: Dashboards can be either private or shared. If a dashboard is labeled **Private** on the **Visibility** field, only you (or a system administrator) can view and edit the dashboard. If a dashboard is labeled **Shared** on the **Visibility** field, every user of the system can view and edit the dashboard.

2. Under the **Source** field, choose an image thumbnail to represent the dashboard. By default, a screenshot of the dashboard will be taken and used as the thumbnail for the dashboard.

This thumbnail will be used if the **Captured Screenshot** option is selected. If you would like to use another image file, select the **Image File** option and choose the image you would like to represent the dashboard.



Note: If your browser does not support screenshot capture, then a default screenshot will be used. In this case you may wish to select an image file to represent the dashboard instead.



3. Click **Save** to save the dashboard to the server. You will be the owner of this dashboard and will have full control over it in the future.

Editing or Updating a Dashboard

Navigation title: Editing or Updating a Dashboard

Existing dashboards can be edited at any time, either for temporary ad hoc purposes or to make permanent changes to be saved to the server.



Note: Only regular users and administrators can edit dashboards. Only the original dashboard creator or an administrator can save changes to the original copy of a dashboard. Otherwise, the only way to edit and save an existing dashboard is to save it as a copy. For more information on saving dashboards, see [Saving a Dashboard](#) on page 1024.

Any changes made to a dashboard will not be changed on the server unless they are saved. To make changes to a dashboard's visualizations, please refer to the chapter on [Visualizations](#) on page 999, which contains information on adding, configuring, and manipulating visualizations.



Note: To simply update a dashboard's title, description, location, or visibility settings, you must be the owner of that dashboard or an administrator.

1. Open the dashboard.
2. Once the dashboard has finished loading, select **Save** from the dashboard drop-down menu.
3. When the **Update Existing Dashboard** window appears, adjust the dashboard title, description, location, and sharing options as needed.
4. Click **Update** to save any updates to the server. See [Saving a Dashboard](#) on page 1024.

Making a Copy of a Dashboard

Navigation title: Making a Copy of a Dashboard

Steps to copy a dashboard.

1. Load the dashboard you want to make a copy of.
2. Select **Save As...** from the dashboard menu at the top of the dashboard interface.
3. When the **Save Dashboard As...** window appears, enter in a new name for the copy of the dashboard.

Other Dashboard Functions

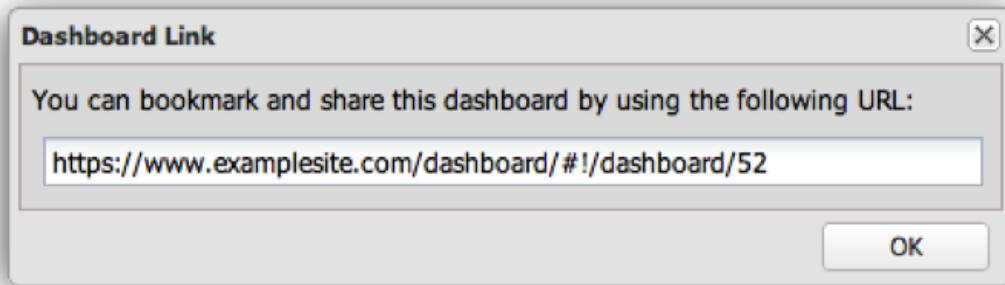
Navigation title: Other Dashboard Functions

Special functions include sharing, exporting, and deleting dashboards and are performed from the details panel of a selected dashboard.

Sharing a Dashboard Hyperlink

The **Share** operation provides a URL that can be used to either bookmark the dashboard or mail a link for another user to access the dashboard. A Dashboard Link window will appear that provides you with the information needed to share the link to the desired dashboard.

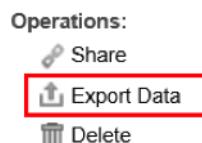




Note: Recipients of the link must also have access to the dashboard in order to view the dashboard.

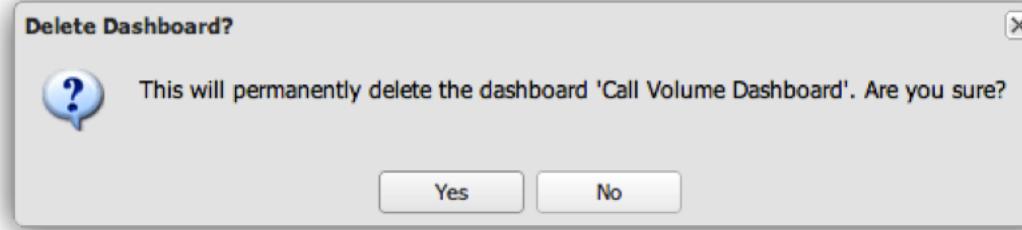
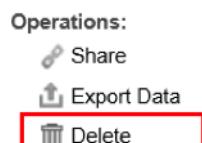
Exporting Dashboard Data

The **Export Data** operation initiates an export of the selected dashboard to be saved as an Excel or CSV (Comma Separated Values) file.



Deleting Dashboards

The **Delete** operation will delete a dashboard. To delete a dashboard, the user must be the owner of the dashboard or have administrator access. Clicking the Delete operation will display a window to confirm that you would like to delete the dashboard.



Dashboard Sharing and Access Controls

Navigation title: Dashboard Sharing and Access Controls

Dashboards can either be private or shared.

If a dashboard is set to **Private**, then only the dashboard creator or a system administrator can view and edit the dashboard.

If a dashboard is set to **Shared**, then every user with proper access can view the dashboard. For users to have proper access, they must have an active account on the dashboard application and have been granted access to each profile that the dashboard contains. Aside from administrators, other users cannot permanently edit dashboards that they did not create. Users can, however, make edits and save the dashboard as a copy.

Making Selections within the Dashboard

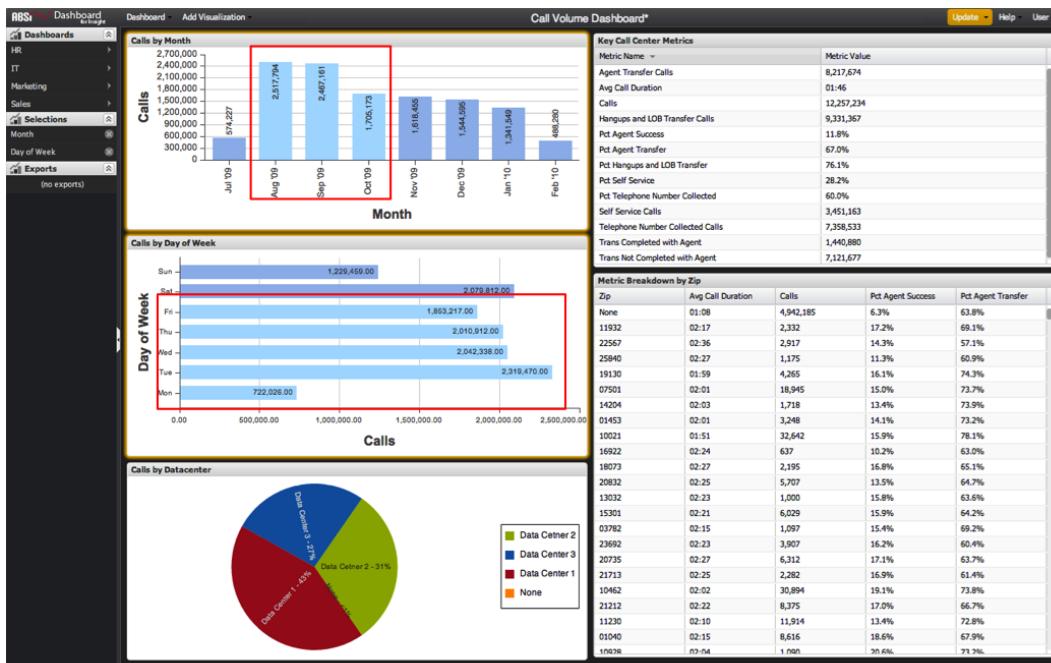
Navigation title: Making Selections within the Dashboard

Data within a dashboard can easily be segmented and explored by the use of selections.

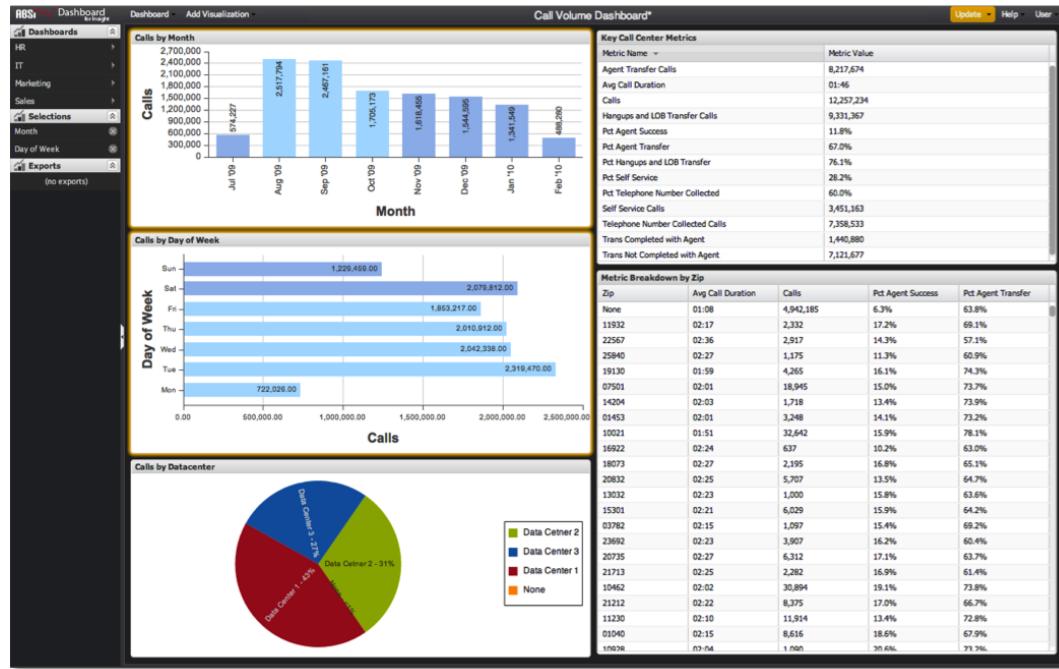
Selections are made by clicking certain elements within visualizations to identify how the data in the dashboard should be segmented. Making a selection in one visualization will segment the data being represented in the other visualizations within the dashboard. Any number of selections can be made, adjusted, or removed at any time, and encourages the user to interact with the data to derive analytical information.

When saving a dashboard, the state of any selections will be preserved at the time of save. Similarly, when a dashboard is loaded, any selections that were preserved during the save will be made effective when the dashboard is loaded.

Selections can be made by clicking on one or more data elements within one of a dashboard's visualizations. Data elements are represented by things like bars in bar charts, columns in column charts, rows in a table, and so on. Selections are highlighted as they are made, and making selections within a visualization will give the visualization an orange border. The exact method for making selections depends on the type of visualization being used.



For each selection that has been made from a visualization, an entry will also appear in the Selections Menu. This entry is listed using the selected dimension name, with one selections entry appearing per visualization.



Note: Visualized data is not automatically updated each time a selection is made. Rather, after you have made your desired selections, you must initiate an update in order to segment the data and update your visualizations.

Locked Selections

Navigation title: Locked Selections

Selections within one or more visualizations on a dashboard can be locked to preserve current selections and prevent further alteration.

Locking visualizations is useful for guiding analysis since certain items can be pre-selected and permanently applied to the entire dashboard. This provides an on-screen visual indicator of what is being filtered without allowing it to be altered.

Locking is also useful for guiding analysis by only allowing users to select on visualizations that are relevant to the analysis that the dashboard is aiming to achieve. For example, in the sample dashboard below, drilling down to the zip code level may be beyond the scope of high-level call volume analysis. In this case it may make sense that the **Metric Breakdown by Zip** table be locked and have selections only allowed on the Month, Day of Week, and Datacenter visualizations.

If you try to make selections on a locked visualization, you will see a message indicating that the visualization is locked. You will also see a gold lock icon in the visualization header when the mouse cursor is hovered over the locked visualization.



As with any selection, locked selections are preserved when a dashboard is saved and remain active when a dashboard is loaded. For more information see [Locking and Unlocking Visualizations](#) on page 1013.

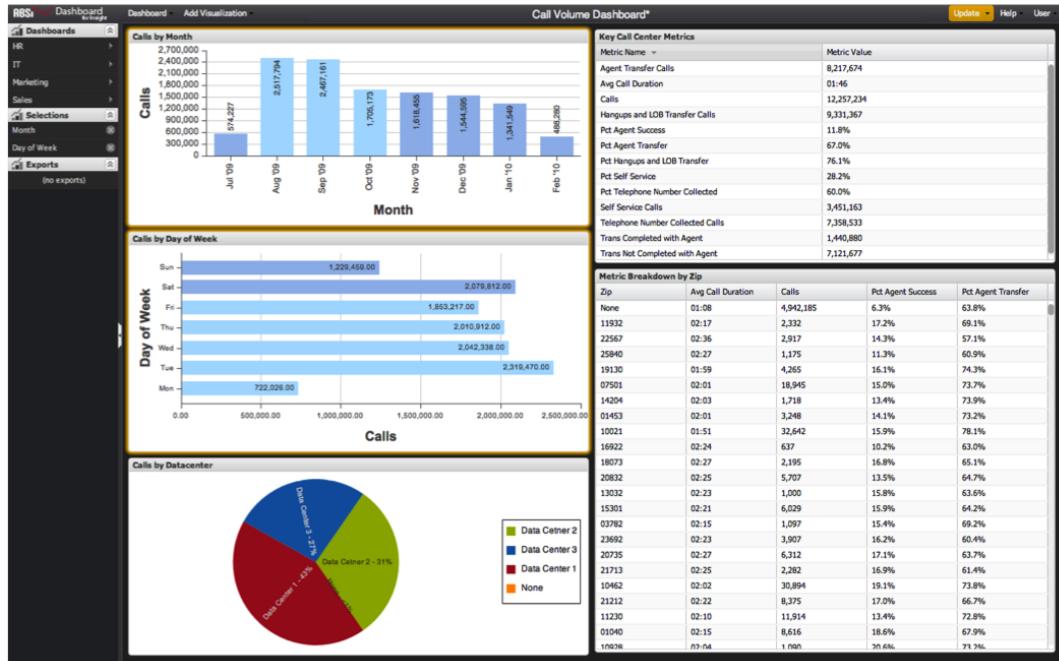
Applying Selections

Navigation title: Applying Selections

Selections are not automatically applied to the dashboard's data results.

Finished making your desired selections and click **Update**.

If the **Update** button is orange, this indicates that you must click it to apply some change in the dashboard's selections. This feature allows you to make multiple selections on the screen and frame out your analytical questions without having a query initiated every time you make a change.

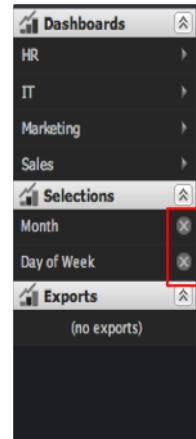


Removing Selections

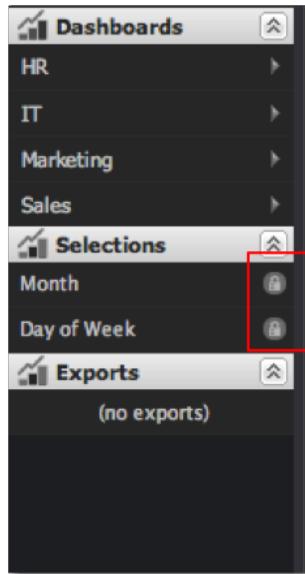
Navigation title: Removing Selections

Selections can be removed at any time.

Just click the X icon inside the corresponding selections item in the Selections Menu.



If a visualization's selections are locked, you will see a small padlock icon replacing the X icon. Locked selections cannot be removed without unlocking the visualization first.



Determining What's Selected

Navigation title:Determining What's Selected

There are two ways to determine which selections have been made within a dashboard.

- First, all selections that have been made on the dashboard will be represented by an item in the **Selections** menu. Each group of selections from a visualization will appear as one item in the list. These will be labeled with the corresponding dimension name (i.e. 'Day', 'Month', etc.). Placing the mouse cursor over this entry will display a popup that identifies exactly which metric values have been selected.



- Another feature to help identify the selections that have been made is visualization highlighting. Any visualization that has a selection will be outlined in orange. Also, placing the mouse cursor over a selection item will highlight its corresponding visualization in bright yellow, showing the user where the selection came from.

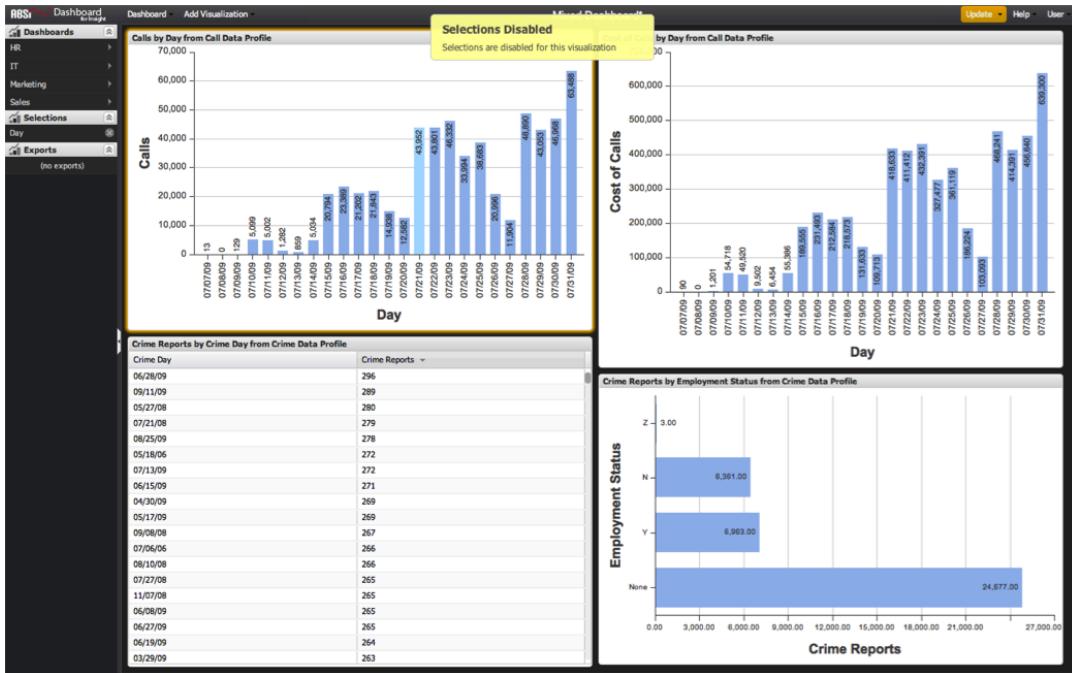
Cross-Profile Selections

Navigation title:Cross-Profile Selections

You can visualize data from multiple profiles in one dashboard.

In some cases, selections from one visualization can also be applied to visualizations from another profile. For example, if you create visualizations from a **Call Center** profile and a **Website Traffic** profile on one dashboard, you could select a target month to have the data in all visualizations simultaneously segmented on that month, despite being entirely different datasets.

When visualizations from multiple profiles exist in a dashboard, you may make a selection in a visualization if that visualization's dimension also exists on all other profiles represented on the screen. However, selections will be disabled if a dimension is not found globally across all other visualizations on the screen, and users will see a **Selections Disabled** message.



Note: Even though dimensions may share the same name across multiple profiles, they may not have the same meaning. It is important to investigate each dimension to determine whether it is appropriate to use it to make selections across multiple profiles.

Exporting Data

Navigation title: Exporting Data

Shows 3 ways of exporting data, and how to retrieve exports.

Data can be exported from the dashboard in three ways.

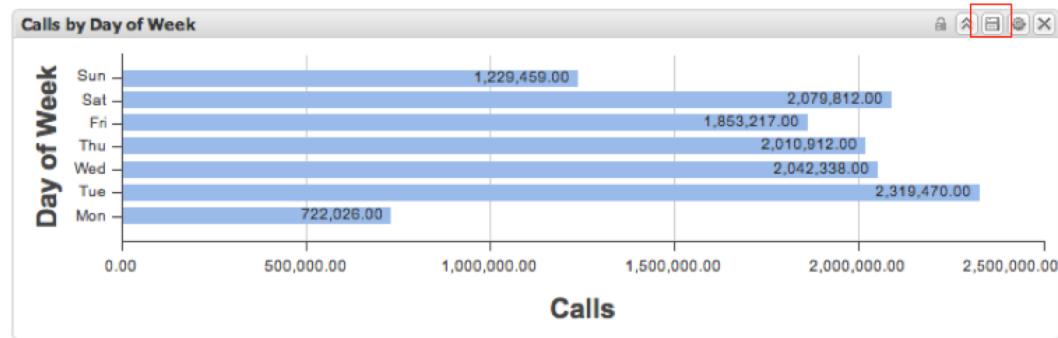
First, you can export data from an individual visualization. Second, you can export your current working dashboard, including any configurations and selections you've made. Third, you can export a saved dashboard without opening it.

Exports are performed using a two-step process. First, exports are queued on the server using one of the three methods above. The status of the export will be displayed in the **Exports** menu as the export is being prepared. Next, when exports are ready, you can download the data in either CSV or Excel format.

Exports may take several minutes, but you can continue to use the application during an export.

Exporting Visualizations

To export data from a visualization, click **Save** in the visualization's tool menu.



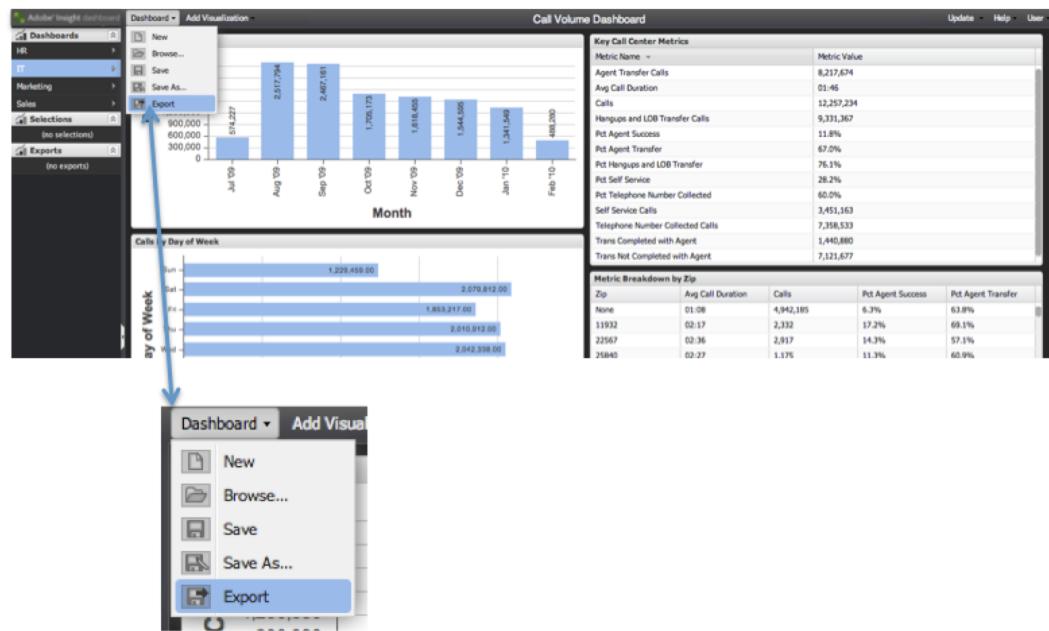
Your export will be initiated on the server and an export indicator will be added to the Exports Menu.

Export Queued

'Calls by Day of Week' queued for export.

Exporting Dashboards

To initiate a data export from a working dashboard, select **Export** from the **Dashboard** menu.



Your export will be initiated on the server and an export indicator will be added to the Exports Menu.

Exporting Saved Dashboards

To initiate a data export from a saved dashboard, use the Dashboard Browser. Within the Dashboard Browser, browse to your desired dashboard and select it so the dashboard's details appear. In the right-hand details panel, under the **Operations** section, select **Export Data**.

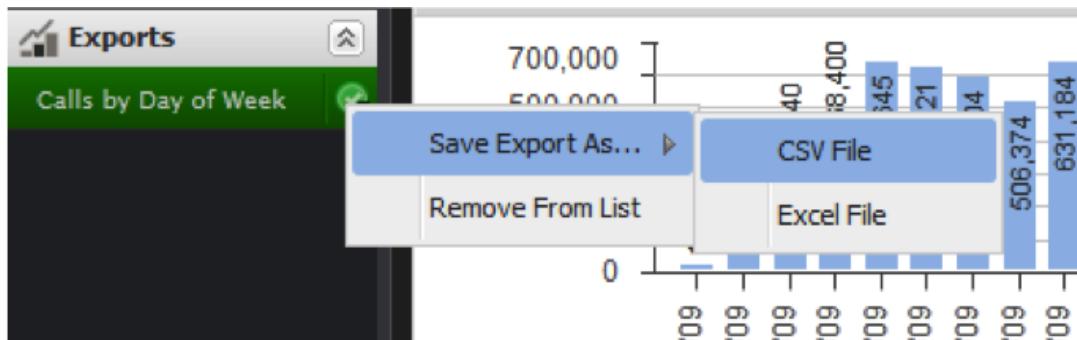
Your export will be initiated on the server and an export indicator will be added to the **Exports** menu.

Retrieving Exports

When an export is complete, a popup notification appears to notify you that the export is ready.



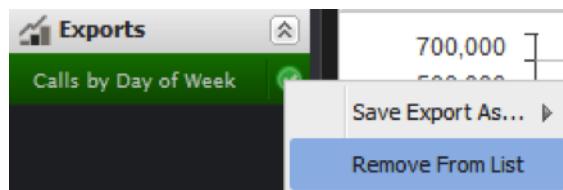
To retrieve the export, use the **Exports** menu. Clicking the green checkmark to the right of the desired export's item will show a drop-down menu. Within this menu, under the **Save Export As...** submenu, select the appropriate menu option to download the export in either CSV or Excel format.



Your browser's file download process now starts.

Exports are not removed automatically, so you can easily download the export in each format. You may remove exports from the **Exports** menu in the left navigation panel. Otherwise, they will automatically be removed when you log out.

To remove an export from the **Export List**, click the checkmark to the right of the export title and select **Remove From List**.



Dashboard Access Controls

Navigation title: Dashboard Access Controls

Dashboards can have two levels of visibility: Private and Shared.

- **Private:** When a dashboard is set to **Private**, it is only accessible by the dashboard owner and administrators. Other users of the system cannot see the existence of the dashboard nor open it.
- **Shared:** When a dashboard is set to **Shared**, every user with proper access can see and open the dashboard. For users to have proper access, they must have been granted access to each profile that the dashboard contains. If a user has not been granted access to one or more profiles within a dashboard, they will not see the existence of the dashboard nor be able to open it.

Only the original owner or an administrator may make changes to a dashboard. This includes making changes to the dashboard content, updating the dashboard name, description, visibility, or deleting a dashboard. Other users may make a copy of the dashboard in order to make changes.

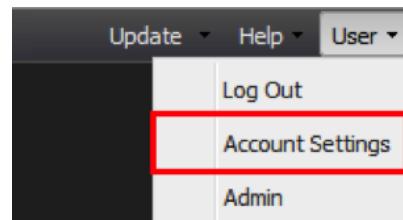
User Account Management

Navigation title: User Account Management

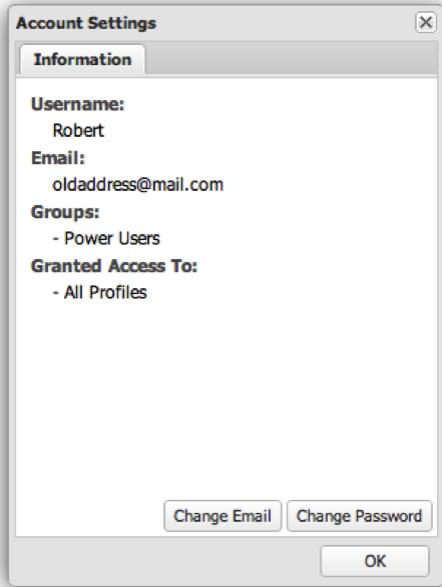
Information on how to access your account settings and change your password.

Accessing your Account Information

To access your account information, select Account Settings from the User menu.



A window will appear that details your account information. This includes your username, e-mail address, group membership, and the data profiles available to you. To verify your user profile has been configured correctly and that you have access to all your data profiles, you can review the profiles listed in the Granted Access To section. If you are missing data profiles, contact your administrator.



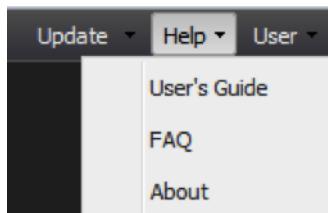
Changing your Password

To change your password, first access your account information. To access your account information, select **Account Settings** from the **User** menu. Click **Change Password**.

Help Menu

Navigation title: Help Menu

Using the **Help** drop-down menu, you can access the dashboard product documentation



- **User's Guide** - Selecting the User's Guide option from the Help drop-down menu directs you to a web link in which you can access this document at any time.
- **FAQ** - The FAQ provides you with a list of frequently asked questions and answers specific to the dashboard. This section serves as a quick reference guide for common user issues.
- **About** - The About portion contains system information, including the version number and legal agreement. It also contains your user-agent string, which is a useful piece of information to provide when seeking technical support.

Additional Support

For further support regarding technical or software performance issues, contact your system administrator. To facilitate the resolution process, be prepared to provide the following information:

- Your Adobe data workbench dashboard user name

-
- Web browser type and version
 - Operating system name and version
 - Dashboard version number
 - Detailed explanation of the issue you are experiencing
 - Screen shots of issue (if possible)

Data Workbench Dashboard Administrator Guide

Navigation title:

Minimum System Requirements

Navigation title:Minimum System Requirements

Lists the minimum configuration requirements.

The Adobe data workbench dashboard is designed to run on a Windows-based server platform with only a few configuration requirements. At a minimum, the dashboard requires:

- Windows Server 2008 (x64 recommended) or 2008 R2 Operating System, with:
 - IIS Web Role enabled
 - Microsoft .NET Framework v4.0
 - Microsoft SQL Server 2008 (the Express Edition is fully compatible.)
- Communication with a properly configured data workbench server running data workbench version 5.45 or above with Query API license

Download Required and Optional Components

Navigation title:Download Required and Optional Components

The dashboard requires several additional Microsoft components in order to operate.

These components are listed below and are downloadable from Microsoft.com. It is recommended that you download these items prior to beginning the installation.

- Microsoft .NET Framework v4.0
- Microsoft SQL Server 2008 Database (Adobe recommends SQL Server 2008 R2 SP1 – Express Edition) with Management Tools
- Microsoft Web Deploy v2.0

Supported Browsers

Navigation title:Supported Browsers

Lists all supported browsers.

The Adobe Insight dashboard is a web-based application that operates using the latest web browser technology. It can be used on any operating system given that the operating system is equipped with a compatible web browser.

Because the dashboard takes advantage of features not supported in older versions of web browsers, more recent browsers are required for an optimal experience when using the application. While many variations of browsers may work, only the following web browsers are officially supported:

- Google Chrome 15 and above (available at <https://www.google.com/chrome/>)
- Microsoft Internet Explorer 9 and above (available at <http://www.microsoft.com/windows/ie/>)

-
- Mozilla Firefox 3.6 and above (available at <http://www.mozilla.org/en-US/firefox/all-older.html>)
 - Safari 5.1 and above (available at <http://www.apple.com/safari>)

Aside from using a compatible web browser, no additional software needs to be downloaded in order to use the dashboard.

Installation Prerequisites

Navigation title: Installation Prerequisites

Before beginning the installation, log on to the web server with an admin account with permissions to configure the server and install new software.



Note: The dashboard system can be deployed in a single tier (IIS and database on the same server) or as a two-tier system. There are additional configuration steps involved in the two-tiered approach that are not covered in this guide.

Enabling IIS

Navigation title: Enabling IIS

The first step is to enable the IIS role on your dashboard server.

1. Under **Administrative Tools**, open the **Server Manager**.
2. Right-click the Roles menu item in the left-hand portion of the **Server Manager** window.
3. Select **Add Roles**.
4. Select **Web Server (IIS)** and continue with the **Add Roles Wizard**. Make sure that the following Role Services are enabled:

Common HTTP Features

Static Content

Default Document

Directory Browsing

HTTP Errors

HTTP Redirection

Application Development

ASP.NET

.NET Extensibility

ISAPI Extensions

ISAPI Filters

Health and Diagnostics

HTTP Logging

Logging Tools

Request Monitor

Tracing

Custom Logging

Security
Basic Authentication
Windows Authentication
URL Authentication
Request Filtering
IP and Domain Restrictions

Management Tools
IIS Management Console
IIS Management Script and Tools
Management Service

- Follow the Wizard to complete the installation.

Installing Required Components

Navigation title:Installing Required Components

Steps to install the required Microsoft components.

- Install Microsoft .NET Framework v4.0.



Note: This must be installed only after installing and configuring the IIS Web Role.

- Follow the wizard and choose defaults where prompted to complete the installation.
- Once installed, verify that the ASP.NET v.4.0 application pool was added within the **Application Pools** listing in IIS.
- Install the Microsoft SQL Server Database.
Use any version of SQL Server 2008 or 2008 R2 (Express is supported) with Management Tools (Adobe recommends SQL Server 2008 R2 SP1 – Express Edition).
- For a generic install without existing SQL Server instances running ahead of time, please select the **Default Instance** option on the **Instance Configuration** screen.
- For the rest of the configuration options, follow the wizard and choose defaults when prompted to complete the installation.
- Install Microsoft Web Deploy v2.0.
For most installations, the **Typical** installation is fine. If, however, you are planning to perform remote deployments, you will need to do a full install (choose **Complete**).

Once all prerequisites are properly installed, you are ready to prepare the data workbench servers to communicate with the dashboard.

Dashboard to Data Workbench Servers Communication

Navigation title:Dashboard to Data Workbench Servers Communication

You must prepare your instance of data workbench to allow the dashboard to communicate with your servers. At a minimum, you must be running data workbench server v. 5.45 and have a valid data workbench Query API license. Each server to be utilized by the dashboard must have a valid Query API license installed.

Once you have met these prerequisites, you must then configure your data workbench instance to grant access to the dashboard via the Query API.

Please refer to the instructions below when preparing your data workbench servers for the Adobe data workbench dashboard.

Verifying Query API Enablement

Navigation title:Verifying Query API Enablement

Each DPU from which the dashboard is to visualize data must have a Query API license.

Below are some instructions on how to validate that Query API is installed and enabled.

1. Find your data workbench server's certificate.
2. Open this certificate in a text editor.
3. Ensure that the line `Product = Query API` exists in the certificate.
4. In the **Trace** folder, open the `InsightServer64.log` in a text editor.
5. In the latest startup log entries, ensure that the line `Enabling Query API (licensed)` appears.
 - If the Query API is not enabled, you will see the entry `Not enabling Query API (not licensed)`.
 - If you do not see any log entries, or suspect that the data workbench server has been restarted since the Query API was added, please restart the data workbench server again and check the log.

If you are unable to validate that Query API is enabled, please contact Adobe ClientCare for assistance.

Configuring Access Control

Navigation title:Configuring Access Control

The dashboard requires certain read-only permissions to be able to access and display your data workbench data. Write privileges are not required.

Configuring access control involves editing your `Access Control.cfg` file on the FSU(s) and adding a new group to grant permission to the data workbench dashboard:

1. Edit the `AccessControl.cfg` file (preferably using the data workbench workstation).
2. Create a new group named *Insight Dashboard Access*.
3. Under this group's members, add the proper CN provided to you for this purpose (Example: CN:INSIGHT-USER01). Contact Adobe Customer Care for this CN.
4. Under this group's read-only access, modify the list to reflect the following:
 - /Profiles/\$
 - /Profiles/
 - /Addresses
 - /Status/
 - /Users/\$
5. Remove any items from the read-write access section, as no write permissions are required for the dashboard.
6. Save the changes to the `AccessControl.cfg` file to the server for permissions to take effect.

Dashboard Deployment

Navigation title:Dashboard Deployment

After installing prerequisite software and implementing the Query API and access control, the next step is to deploy the dashboard system to your server.

Deploying the Dashboard

Navigation title:Deploying the Dashboard

Steps to deploy the dashboard in IIS.

1. Create an installation folder to install the dashboard, such as c:\inetpub\wwwroot\dashboard.
2. Create the dashboard's application pool in IIS.
3. Open the IIS Manager Console.
4. Go to **Application Pools**.
5. Select **Add Application Pool...** in the **Actions** menu to the right.
6. In the **Add Application Pool** form, use the dashboard for the name and select .NET Framework v.4.0.xxxxxx as your .NET Framework Version.
7. Leave other fields as their default and click **OK** to create the application pool.
8. Deploy the dashboard application.
9. Open the IIS Manager Console.
10. Expand the **Default Web Site**, you should see the folder you created in c:\inetpub\wwwroot\dashboard by default.
11. Right-click the folder and select **Convert to Application**.
12. Select the **Application Pool** created in Step 2 (for example, “Dashboard”).
13. Under **Sites**, right-click the web site to which you wish to deploy (for example, “Default Web Site”).
14. Select **Deploy > Import Application**.
15. Browse to and select the dashboard deployment file provided by Adobe.
16. Click **Next** twice to proceed to the Enter Application Information screen.
17. From this screen, you can choose to customize your dashboard deployment.
18. For **Application Path**, enter the folder name previously selected.
19. Under **Disable Automatic Database Upgrade**, enter **False**, since this is a new installation.
20. Customize your connection string, if necessary. For example:

```
Data Source=.;Initial Catalog=thinclientdb;Integrated Security=SSPI;Connect  
Timeout=30;  
Application Name=Insight Dashboard;MultipleActiveResultSets=true;
```
21. Click **Next** and IIS will begin installing the application.
22. Once the installation has completed, you should see no errors in the installation log.

Configuring the SQL Server

Navigation title:Configuring the SQL Server

Before the dashboard can operate, you must allow it to access the SQL Server.

1. Open the SQL Management Studio as an Administrator.
2. Add a new login by right-clicking **Logins** and selecting **New Login**.
3. Enter the full application pool identity name.
By default, the application pool identity is named after the application pool. If you choose dashboard, then the identity will be named IIS AppPool\dashboard.
4. Select **Server Roles** and check the **dbcreator** role.
5. Click **OK** to add the new user.

Initializing the Dashboard

Navigation title:Initializing the Dashboard

The final step is to run the dashboard for the first time to allow it to initialize.

-
1. Open a web browser and enter the URL to the newly deployed site (for example, <http://localhost/dashboard>).
 2. Connecting for the first time will setup the database tables, so you may experience a slight delay.
 3. The initial logon credentials are:
 - **Username:** admin
 - **Password:** password
 4. On your first login, go to **User > Account Settings** and select **Change Password** to change your administrator password.

The dashboard installation is now complete. If you haven't already, use the instructions detailed in the nch section of this guide to configure your communication with data workbench servers and to manage users and groups.



Note: Dashboard error and audit logs can be found in the `logs` directory within the installation path.



Note: If you need to change the application pool identity to a different account, make sure to grant access to the database and give the identity read/write access to the `logs` folder in the installation path.



Note: If you ever need to change the connection string for the database, simply edit the value using the **IIS Management Console**.

Add Dashboard License Key

Navigation title: Add Dashboard License Key

The dashboard product requires a license provided by Adobe ClientCare.

1. Open **SQL Management Studio** as an Administrator.
2. Open the database created by dashboard (for example, `thinclientdb`).
3. Right-click the **Configuration** table and click **Edit Top 200 Rows**.
4. Find the **licenseKey** field and enter the key provided by Adobe ClientCare into the **Value** column.
5. Restart the **Application Pool** in the **IIS Manager Console**.

Managing from the Administration Console

Navigation title: Managing from the Administration Console

Administrative functions of the dashboard application are performed using the **Administration Console**. Using this console, you can manage profile connections, users, user groups, folder hierarchies, and the schema definitions used within the dashboard.

To reach the Administration Console, you must first be logged in to the dashboard as an administrator. Once logged into the dashboard interface, click **User > Admin** to navigate to the Administration Console.

Administration Console

Profile Manager Group Manager Folder Manager

Add Profile Connection

Name: [] Profile DPU Address: [] PBU: https://da3d918f.insight.annilure.com

	User Cert	Cert CN	Enabled	Available
Call Data Demo v1	Yes	Inbound	Yes	
Contracts Analysis v4	Yes	Inbound	Yes	
CrimeDataDC v2	Yes	Inbound	Yes	
Earth Quake v1	Yes	Inbound	Yes	
ES2 v3	Yes	Inbound	Yes	
GeoData Example	Yes	Inbound	Yes	
New Intelligence 3	Yes	Inbound	Yes	
NYC Schools v3	Yes	Inbound	Yes	
SocialTrack-v2	Yes	Inbound	Yes	
Trainer Evaluation v1	Yes	Inbound	Yes	
WITS v1	Yes	Inbound	Yes	

Managing Profile Connections

Navigation title: Managing Profile Connections

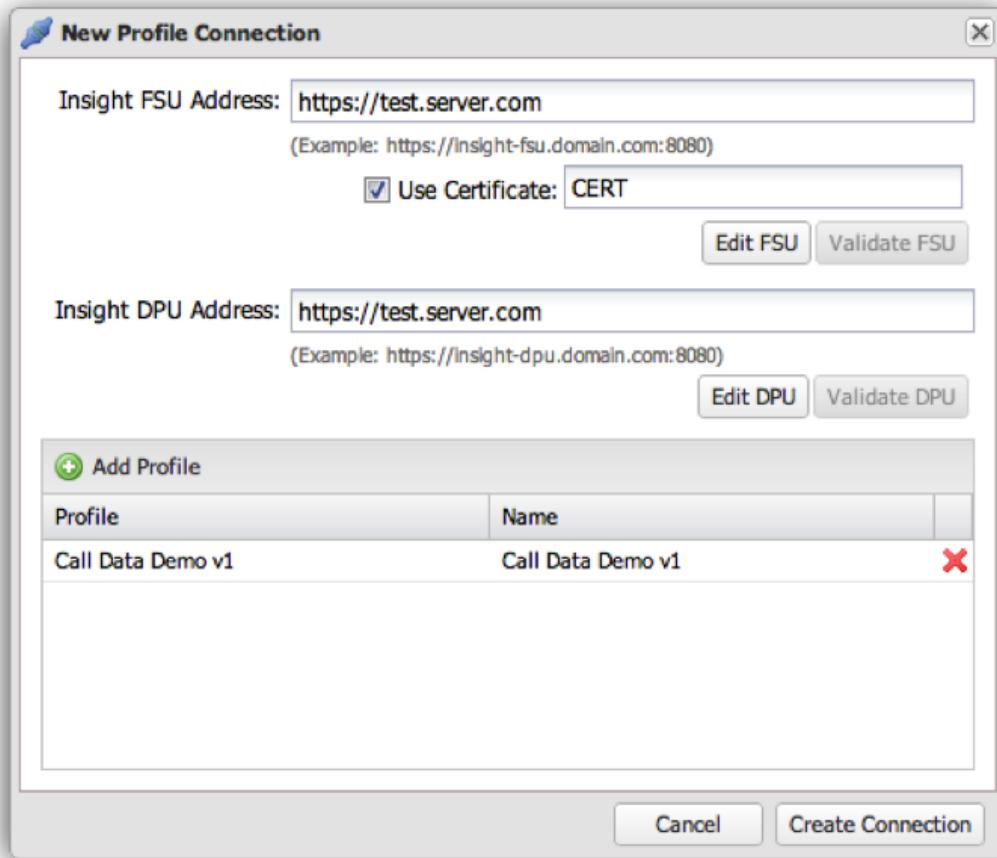
Connections between the data workbench Dashboard Server and data workbench data profiles are managed using the Profile Manager within the Administration Console. Data workbench profiles used within the dashboard do not need to be centrally located, and can be hosted anywhere if there is appropriate connectivity between the dashboard server and the data workbench server(s) hosting the profile(s).

Only administrators have the privilege to create and modify profile connections. To manage profile connections, use the Profiles tab in the Administration Console.

Adding a Profile Connection

Navigation title: Adding a Profile Connection

1. Click **Add Profile Connection** to bring up the **New Profile Connection** window.
2. Using the form below, fill in the necessary fields:



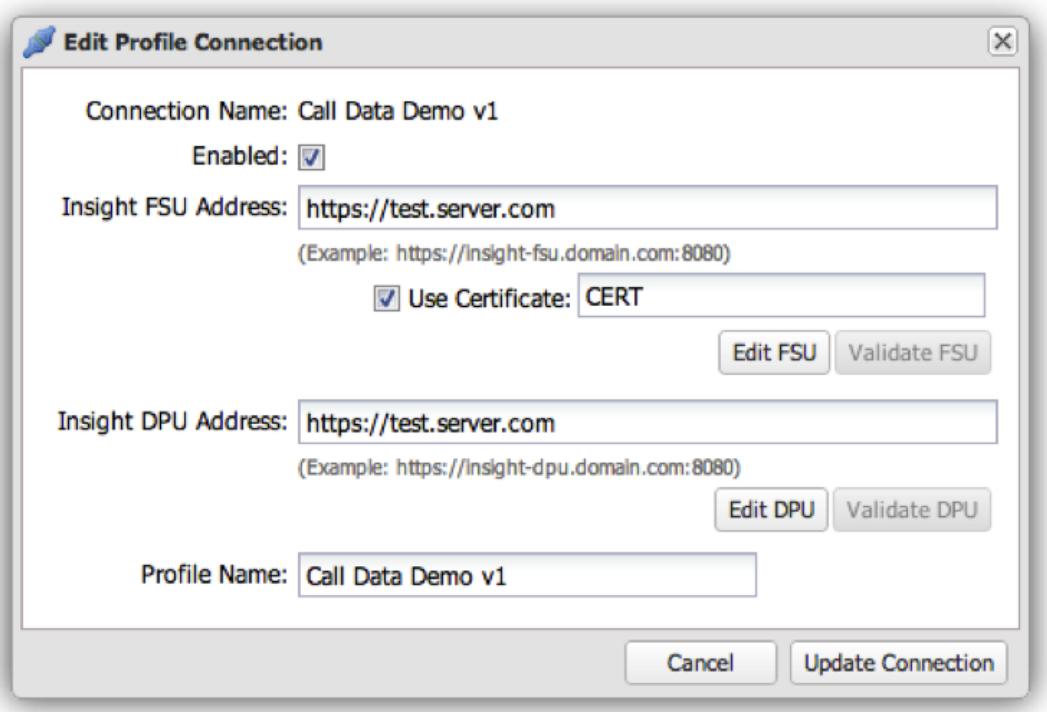
- a) **Insight FSU Address:** Enter the address to the FSU that hosts the profile(s) you would like to add to dashboard.
- b) **Use Certificate:** Click if authentication is required, along with the CN of the certificate (this certificate needs to be added to the dashboard server in advance using the Windows Certificate Manager).

- c) **Insight DPU Address:** Enter the address to the DPU that hosts the profile(s) you would like to add to dashboard.
- d) **Profiles:** If multiple profiles exist on the given FSU/DPU entries above, multiple profiles may be added here by clicking the Add Profile button.
- **Profile:** Enter the name of the profile as it appears in Insight
 - **Name:** Enter the name of the profile as it should appear in dashboard.
3. Once the form has been configured appropriately, click **Create Connection** to add the profile connection to the system.
If the operation was successful, you will see a prompt indicating that the profile connection has been created.

Editing a Profile Connection

Navigation title: Editing a Profile Connection

1. Click the gear icon  at the end of the row for the profile connection you wish to edit.



2. Edit the profile connection's settings and click **Update Connection** to save.

If the operation was successful, you will see a message indicating that profile connection was successfully updated.

Removing a Profile Connection

Navigation title: Removing a Profile Connection

1. Click the X at the end of the row to remove the connection to the profile.
2. A prompt will be displayed asking if you want to delete the profile connection. Select **Yes** to delete the profile connection.



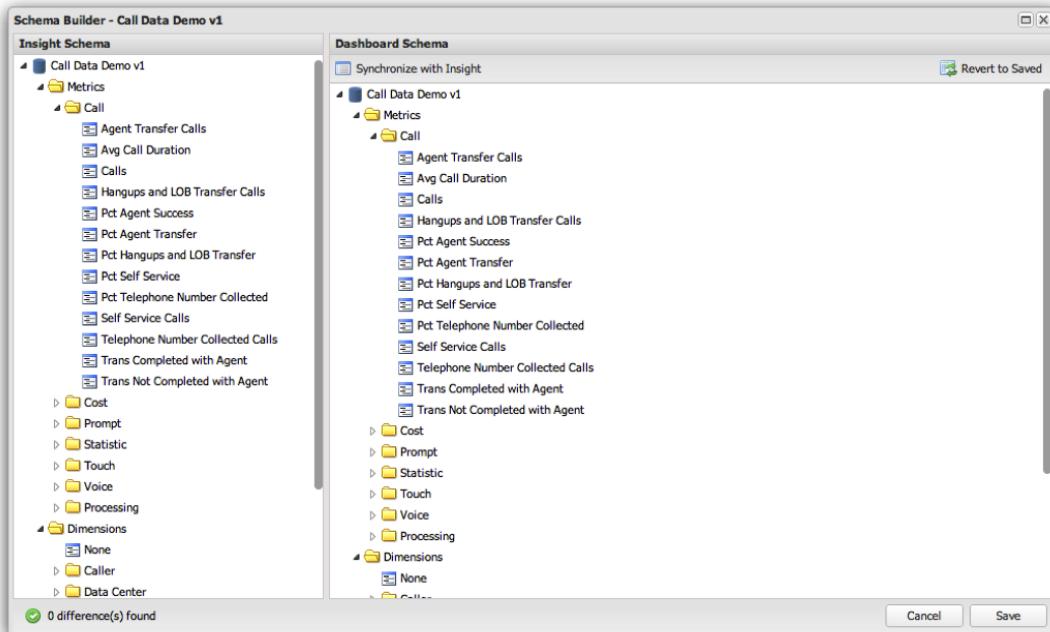
If the operation was successful, you will see a message indicating that the profile connection was successfully removed.

Managing the Schema

Navigation title:Managing the Schema

The schema used within the Adobe Insight dashboard is not retrieved live from Insight. Instead, the dashboard administrator exposes certain items from the Insight schema to end-users of the dashboard system. This allows the administrator to only provide access to the metrics, dimensions, and filters relevant to dashboard users.

Schema management is performed using the **Schema Builder** window in the **Administration Console** and is accessible by clicking the icon to the right of a profile connection's row.



Initializing and Updating a Profile's Schema Definition

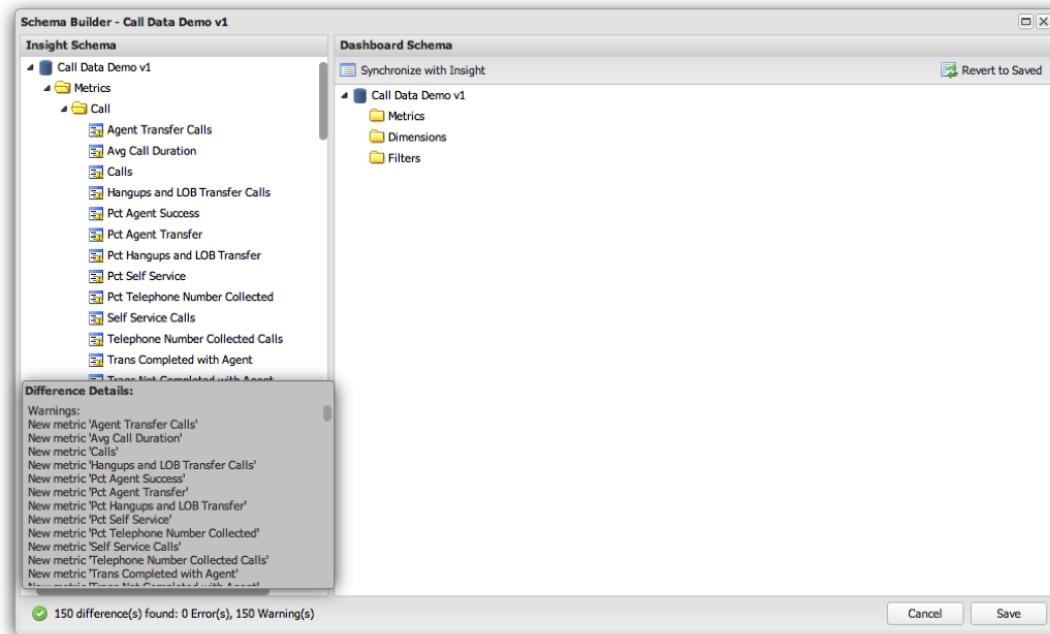
Navigation title:Initializing and Updating a Profile's Schema Definition

1. Open the **Schema Builder** for the profile you would like to set up.
2. A **Loading** message will be displayed while the schema is being retrieved from the Insight profile. The length of time for loading the schema is dependent on the complexity of the profile being loaded.

-
3. When complete, you will see a summary of the differences between the **Insight Schema** in the left pane, and the **Dashboard Schema** in the right pane. This summary will appear in the lower left-hand portion of the **Schema Builder** window.



Note: When setting up the schema for the first time, each metric, dimension, and filter will be listed differently from the dashboard's schema. This is because the dashboard schema objects do not exist at this time.



4. Click the **Synchronize with Schema** button to synchronize all metrics, dimensions, and filters from the Insight Schema view with the Dashboard Schema view.
5. When complete, you should see a message indicating that there are no differences found:



6. If there are any errors with the Dashboard Schema—such as duplicate metrics and dimensions—then you must correct them manually before you can save.



Note: You can selectively remove any metrics, dimensions, or filters from the **Dashboard Schema** that you do not want to appear to end-users of the dashboard. You will receive a warning that items are not present in the Dashboard Schema, but it will not prevent you from saving.

7. When ready, click **Save** to save your changes to the dashboard's schema.
8. The dashboard system will use this schema definition to populate the dimensions, metrics, and filters available to end-users of the dashboard interface.

Managing Folders

Navigation title: Managing Folders

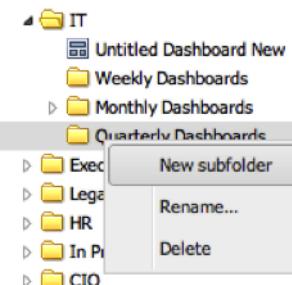
All dashboards are saved into folders. The folder structure is entirely flexible and can be modified at any time. By default, only a single **Main** folder exists at installation. This allows you to create a folder structure or dashboards specific to your organization.

Only administrators have the privilege to create and modify the folder structure. This is performed within the **Folder Manager** and is accessible from the **Folder Manager** tab in the **Administration Console**.

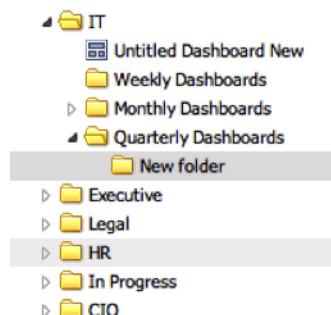
Adding Folders

Navigation title:Adding Folders

1. Right-click an existing folder, and select **New subfolder**.



2. A new subfolder will be created.

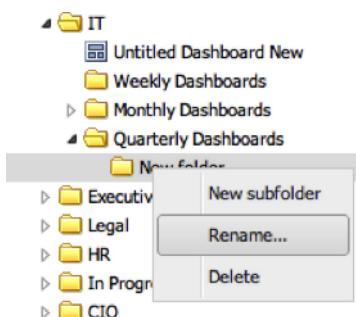


If the operation was successful, you will also see a prompt indicating that the folder has been created.

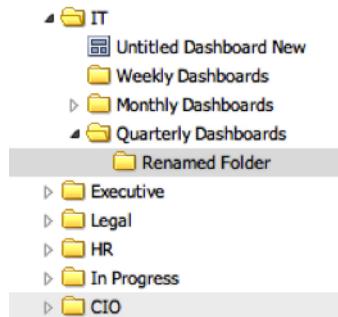
Renaming Folders

Navigation title:Renaming Folders

1. Right-click the folder to rename, and select **Rename....**



-
2. Enter the new name for the folder and press **Enter**.

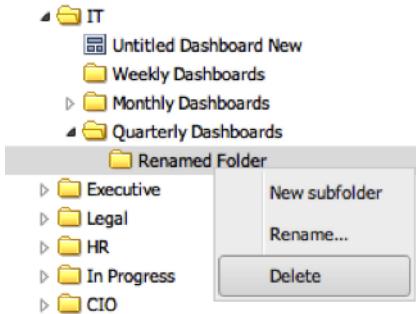


If the operation was successful, you will see a prompt indicating that the folder has been renamed.

Deleting Folders

Navigation title: Deleting Folders

1. Right-click a folder to delete and select **Delete**.



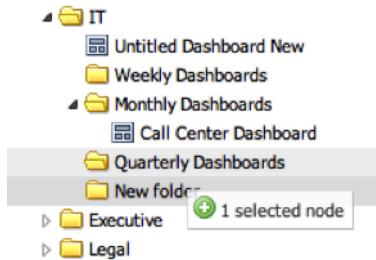
Note: The folder must be empty before you can delete it.

2. You will be prompted to make sure you wish to delete the folder.
3. If the operation was successful, you will also see a prompt indicating that the folder has been deleted.

Moving Dashboards and Folders

Navigation title: Moving Dashboards and Folders

To move folders and dashboards between folders, simply drag-and-drop existing folders and dashboards onto other folders



Changes will be updated live on the server as you make manipulations.

Managing User Accounts

Navigation title:Managing User Accounts

User account management is performed using the User Manager section within the Administration Console.

Users can be created, edited, and deleted using the tools in this area. The number of users you can have at one time depends on how many users are licensed.

All user operations below are performed within the **User Manager** tab in the **Administration Console**.

Adding a User Account

Navigation title:Adding a User Account

1. Click **Add User** to bring up the **New User** prompt.

A screenshot of the 'New User' dialog box. The dialog has a title bar 'New User' with a close button. It contains four input fields: 'Username', 'Password', 'Confirm Password', and a dropdown menu 'Authentication Method' set to 'Forms'. Below these are two tables: 'Available Groups' on the left and 'Assigned Groups' on the right. The 'Available Groups' table lists 'Administrators', 'Finance Department', 'HR Department', and 'IT Team'. The 'Assigned Groups' table currently has no entries. At the bottom are 'Cancel' and 'Add User' buttons.

2. Fill in the necessary fields to complete the form.
 - a) **Username:** Enter the user name.
 - b) **Password:** Enter a password more than 6 characters long.

- c) **Confirm Password:** Reenter the password.
d) **Authentication Method:** select an option from the drop-down list.

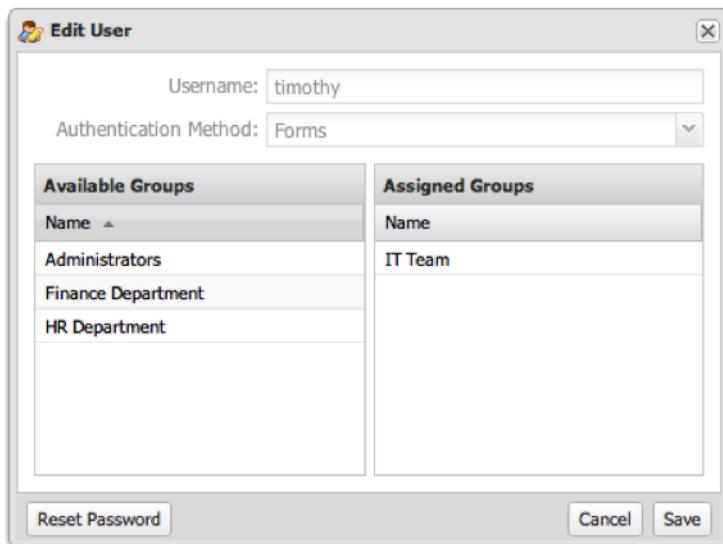
Option	Description
Forms	By default, the dashboard stores the user account and authenticate internally.
LDAP	Select this option if the user is to be authenticated via LDAP. (The user must already exist in the directory).
Windows	Select if the user is to be authenticated using Windows authentication (the user must already exist in the Windows directory).

3. **Assigned Groups:** Choose from the default Administrators group and any other groups that have been created. No groups are required at this time and the user's group membership can be modified at any time.
4. Once the form has been configured appropriately, click **Add User** to add the user to the system. If the operation was successful, you will see a prompt indicating that the user was created.

Editing a User Account

Navigation title: Editing a User Account

- Click the gear icon  at the end of the row for the user you wish to edit



- Edit the user's account settings and click **Save**. If the operation was successful, you will see a message indicating that the user group was successfully saved.

Resetting a User's Password

Navigation title: Resetting a User's Password

- Click the gear icon  at the end of the row to change the user password.
- Click **Reset Password** to bring up the Reset Password window.
- Enter a new password and confirm the password.
- Click **Change Password** to change the user's password. If the operation was successful, you will see a message indicating that the user's password was successfully changed

-
5. If no other changes need to be made to the user account, click **Cancel** in the **Edit User** dialog.

Removing a User Account

Navigation title: Removing a User Account

1. Click the **X** icon at the end of the row for the user whom you wish to remove from the system.
A prompt displays, asking whether or not you want to delete the user.
2. Select **Yes** to delete the user.
If the operation was successful, you will see a message indicating that the user was successfully removed.

Managing User Groups

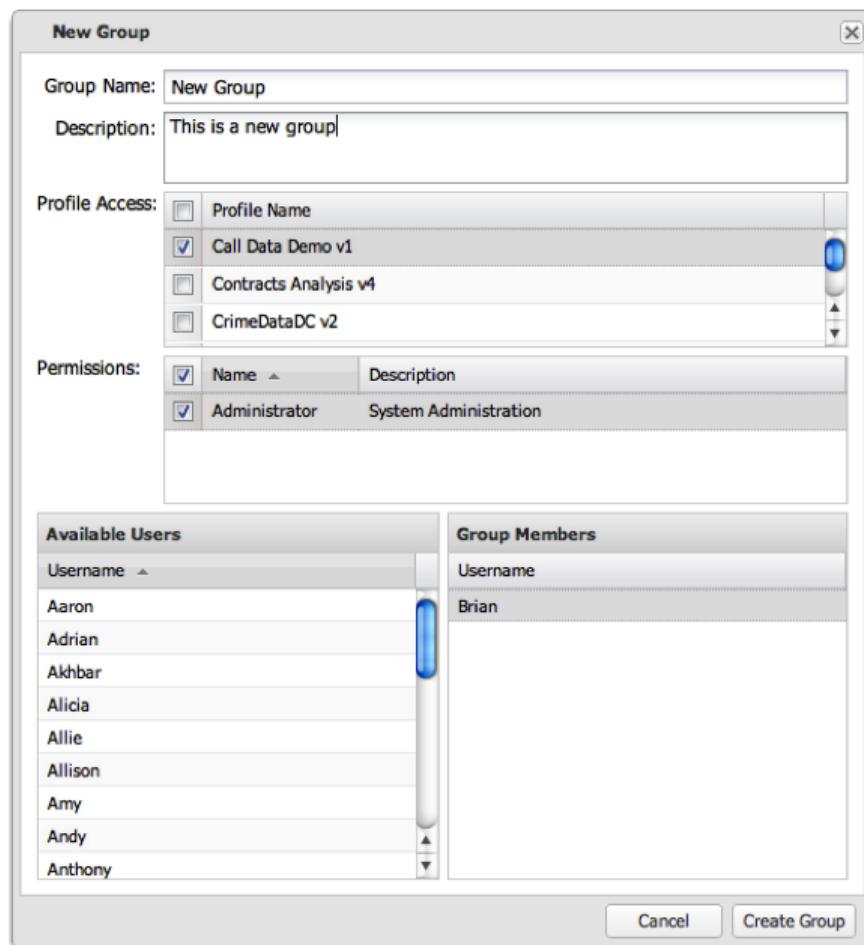
Navigation title: Managing User Groups

User group management is performed within the **Group Manager** section in the **Administration Console**. All operations below are performed within the **Group Manager** tab.

Creating a User Group

Navigation title: Creating a User Group

1. Click **Create Group** to bring up the **New Group** form.



The screenshot shows the 'New Group' dialog box. At the top, there are fields for 'Group Name' (set to 'New Group') and 'Description' (set to 'This is a new group'). Below these are sections for 'Profile Access' and 'Permissions'. In the 'Profile Access' section, 'Call Data Demo v1' is checked. In the 'Permissions' section, 'Administrator' is selected with a checkmark. At the bottom left, the 'Available Users' list includes Aaron, Adrian, Akbar, Alicia, Allie, Allison, Amy, Andy, and Anthony. On the right, the 'Group Members' list shows Brian. At the bottom right are 'Cancel' and 'Create Group' buttons.

-
2. Fill in the necessary fields to complete the form:

Option	Description
Group Name	Enter a unique group name.
Description	Enter a brief description of the group (optional).
Profile Access	Select which profiles this group will have access to.
Permissions	Administrator: Save, view, modify, and delete all dashboards for all profiles. Administrator privileges allow a user to log in to the administration area and perform all administrative functions. Read-Only: Restrict permissions to read-only for dashboards. The user cannot add or modify visualizations or save dashboards. No permissions checked (default): No permissions checked grants regular users permissions to the group. These users can create and modify visualizations, save dashboards, and view the dashboards they have access to.
Available Users / Group Members	Drag-and-drop users from the list of Available Users into the Group Members section. The group members can be modified at any time.

3. Click **Create Group** to create the user group.

If the operation was successful, you will see a message indicating that the user group was successfully created.

Editing a User Group

Navigation title: Editing a User Group

1. Click the gear icon at the end of the row to edit the group.

Edit Group

Group Name:	New Group												
Description:	This is a new group												
Profile Access:	<input type="checkbox"/> Profile Name <input checked="" type="checkbox"/> Call Data Demo v1 <input type="checkbox"/> Contracts Analysis v4 <input type="checkbox"/> CrimeDataDC v2												
Permissions:	<table border="1"> <thead> <tr> <th>Name</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td>Administrator</td> <td>System Administration</td> </tr> </tbody> </table>	Name	Description	Administrator	System Administration								
Name	Description												
Administrator	System Administration												
Available Users	Group Members												
<table border="1"> <thead> <tr> <th>Username</th> </tr> </thead> <tbody> <tr><td>Aaron</td></tr> <tr><td>Adrian</td></tr> <tr><td>Akbar</td></tr> <tr><td>Alicia</td></tr> <tr><td>Allie</td></tr> <tr><td>Allison</td></tr> <tr><td>Amy</td></tr> <tr><td>Andy</td></tr> <tr><td>Anthony</td></tr> </tbody> </table>	Username	Aaron	Adrian	Akbar	Alicia	Allie	Allison	Amy	Andy	Anthony	<table border="1"> <thead> <tr> <th>Username</th> </tr> </thead> <tbody> <tr><td>Brian</td></tr> </tbody> </table>	Username	Brian
Username													
Aaron													
Adrian													
Akbar													
Alicia													
Allie													
Allison													
Amy													
Andy													
Anthony													
Username													
Brian													
		<input type="button" value="Cancel"/> <input type="button" value="Save"/>											

2. Edit the group's settings and click **Save** to save the changes.

If the operation was successful, you will see a message indicating that the user group was successfully saved.

Removing a User Group

Navigation title: Removing a User Group

1. Click the **X** icon at the end of the row for the group you wish to remove from the system.
A prompt displays, asking whether or not you want to delete the group
2. Click **Yes**.
If the operation was successful, you will see a message indicating that the group was successfully removed.

System Logs

Navigation title: System Logs

Error and audit logs can be found in the **Logs** directory within the installation path. These logs are valuable for auditing the dashboard system and debugging any problems you may encounter. The logs are stored in files using the current date of the server:

Error logs: `thinclientweb-yyyy-mm-dd.log`

Audit logs: `thinclientweb-audit-yyyy-mm-dd.log`

Implementing Adobe Data Workbench

The Data Workbench Implementation Guide acts as a reference for onboarding, planning, and technical questions for implementation.

Adobe Data Workbench (DWB) is a component of Adobe Analytics Premium (AAP). It is a powerful, highly-configurable visual analysis and reporting application that provides real-time, multi-dimensional metric analysis, dynamic segmentation, advanced data visualization, and extended analysis features for business professionals, data analysts, and system architects. DWB gauges the effectiveness of a company's business objectives by analyzing how you can interact with all of your data channels, allowing you to define trends, create predictive models for forecasting, and optimize business processes for enhanced performance. By analyzing data from multiple channels, you can generate a complete view of customer behavior and identify actionable responses across your organization.

DWB is highly-configurable analytics tool of AAP customized to fit the needs of your organization. DWB allows you to visualize any type of structured data and lets you to conduct queries against the entire dataset, providing you with the ability to conduct real-time, custom, ad hoc analysis of vast amounts of data. You can then set up and schedule needed reports to be generated and delivered as you need them.

DWB provides:

- Multichannel real-time data correlation
- Flexible data collection from multiple sources
- Powerful advanced data visualizations

The Data Workbench Implementation includes the following phases:

1. Onboarding Tasks for Data Workbench
2. Discovery and Requirements
3. Installation and Provisioning
4. Architecture Setup
5. Administration Setup
6. Configurations
7. Post-implementation

DWB Onboarding

These onboarding instructions are for customers implementing Data Workbench with a single report suite using Adobe managed services without consulting services. If you are new to Adobe Data Workbench and Analytics Premium, the Adobe Onboarding team will be your initial contact. If upgrading from standard Adobe Analytics or an earlier version of DWB, an Adobe Customer Success Manager will contact you to begin the initial onboarding process.

See the [Basic Onboarding Instructions for DWB Managed Services](#) on page 1059 section for more information.

DWB Discovery and Requirements

Gather inputs about the questions and tasks required to devise solution(s) in Data Workbench.

See the [Data Workbench Discovery and Requirements](#) on page 1057 section for more information.

DWB Installation and Provisioning

The Installation and Provisioning phase helps you set up the client workstation and on-premise servers, and to provision and configure these DWB components for your specific needs.

See the [Data Workbench Installation and Provisioning](#) on page 1059 section for more information.

DWB Architecture Setup

The DWB architecture lets you set up data feeds and design the schema.

See the [Data Workbench Architecture](#) on page 1061 section for more information.

DWB Administration Setup

Information about access control, errors and alerts, and server upgrades when administrating DWB.

See the [Data Workbench Administration](#) on page 1065 section for more information.

DWB Configuration and Implementation

Instructions for configuring and implementing DWB.

See the [Data Workbench Configuration and Implementation](#) on page 1065 section for more information.

DWB Post-Implementation

After setting up DWB, you can implement these features.

See the [Data Workbench Implementation of Features](#) on page 1086 section for more information.

Data Workbench Discovery and Requirements

Gather and describe the business questions appropriate to your marketing environment when implementing Data Workbench.

This section lets you to gather inputs about the questions and tasks required to devise solution(s) in Data Workbench (DWB), which can address these questions accurately, unambiguously, and in a technology-independent manner, providing references to the business terminology and the Adobe Analytics Premium solution. This section provides information about these objectives and the associated requirements.

Phase 1: Key Business Goals/Objectives

The following tables prompt you to identify your customer base and analyze construction of your DWB implementation.

- Understanding the your Customer Base
- Understanding specific Business Case (For example, effectiveness of Self-Service and other data channels/offline data sources)

Understanding Your Customer Base

Understand why customers use your site, the challenges you face, and how DWB will help you based on your business model. For example, how to measure, monitor, and analyze your customers to cross-sell other products and services, get the list of active users and account penetration, and other goals.

ID	Business Question/Requirement	Priority	Phase	Dependencies
1a	Specific Business Question 1	High/Medium/Low	1	Common Key, dependent on some other key etc.
1b	Specific Business Question 2	High	1	Any Dependency

Analysis Construction

Workspace Analysis Data Source(s)	Add Workspace name
Workspace Dimensions & Metrics needed	Identify Dimensions: Identify Metrics:
Workspace Filters, Flags, and Tools Needed	Identify Segments: Identify Tools:
What actions can be derived from this analysis?	Understand tasks and content using specific DWB workspaces.

Phase 2: Understanding Specific Business Cases

Understand other data sources and channels, and learn how these will relate to your business cases.

ID	Business questions/requirements	Priority	Phase	Dependencies	
2a	Specific business requirement 1	High/Medium/Low	1	Common Key, dependent on some other key, Account flag/identifier etc.	
2b	Specific Business Requirement 2	High/Medium/Low	1	Any dependency	

Analysis construction

Workspace Analysis Data Source(s)	Sample Workspace name
Workspace Dimensions & Metrics needed	Dimensions: Define required dimensions. Metrics: Define metrics needed.
Workspace Filters, Flags, and Tools Needed	Segments: Identify your customer segments. Tools: Select tools required.
What actions can be derived from this analysis?	What to understand from this workspace

Data Sources

Data sources	Priority	How often is data received?
Site 1 Name report suite (RSID)	1	Hourly
Site 2 Name (If any) (RSID)	1	Hourly
Data Source 1 (If applicable)	2	Daily?
Data Source 2 (If applicable)	3	Daily?

See additional [Data Workbench documentation](#).

Data Workbench Installation and Provisioning

Content to understand, install, set up, and provision Data Workbench (DWB) for On-Premise implementation.

See overview and set up instructions for on-premise DWB servers at [Server Installation and Installation](#).

See overview and set up instructions for the DWB client workstation at [Client Installation Guide](#).

See [Basic Onboarding Instructions for DWB Managed Services](#) on page 1059 to implement DWB managed services.

Basic Onboarding Instructions for DWB Managed Services

Follow these steps to start the onboarding process for Adobe Data Workbench (DWB), a component of Adobe Analytics Premium (AAP).

These onboarding instructions are for customers implementing Data Workbench with a single report suite using Adobe managed services without consulting services. If you are a new AAP customer implementing DWB, the Adobe Onboarding team will be your initial contact. If upgrading from Adobe Analytics standard or from an earlier version of DWB, an Adobe Customer Success Manager will assist you.

Onboarding Info: What We Need from You

Adobe will contact you to:

- Identify a Primary User for DWB to generate a certificate specifically for that user on the network directory. The primary user will also act as the point person to interact with Adobe Customer Care.
- Identify the report suite to be loaded into DWB.

The Adobe Digital Marketing teams will then take your information to create profiles, set accounts, and deliver a configuration file for DWB.

Adobe Onboarding Tasks

- Adobe Customer Care creates a DWB licensed account. Adobe Customer Care generates a DWB certificate for the primary user.
- Adobe Customer Care defines the primary user as a “supported user,” the person identified for supported calls and problem resolution.
- Adobe Customer Care loads software package to the DWB license and software portal (SoftDocs/Software and Docs profile) to be downloaded to your organization.
- Adobe TechOps team prepares the Production and Development environments and their profiles (per contract) for DWB.
- Adobe TechOps team configures data feeds and profile management scripts.

-
- Adobe TechOps team creates and sends the DWB configuration file (Insight.cfg) to the Adobe Onboarding team responsible for on-boarding tasks associated with your organization.

After customizing your data feeds and generating credentials, certificates, and a profile configuration, the Adobe Customer Care will send your configuration file and credentials to take the next step.

Access your Custom Install Files

You will receive these setup files from Adobe Customer Care to install the DWB Workstation on your client computer.

- Your custom DWB configuration file (Insight.cfg)

This configuration file on the client computer will include connections to your managed DWB servers.

- Login credentials for the Licensing Portal to download the DWB Setup Wizard and required certificate (.pem file) with the name of your primary user.

Download additional setup files

1. Browse to: license.visualsciences.com to download you license certificate and the DWB Setup Wizard executable.
2. Enter your Organization (Account Name), the name of the primary User, and the Password that you received from Adobe Customer Care, then click login.



Note: Your browser might prompt you to present a digital certificate at this point. If it does, click Cancel to dismiss the dialog box.

3. Locate the certificate issued for your instance of Adobe Data Workbench (<PrimaryUser>.pem) in the Downloads section and download.
4. Locate Standard Client Installer in the Downloads section to download the DWB Setup Wizard (InsightSetup-x.xx.exe file).
5. After receiving and downloading files from Adobe Customer Care, run the DWB Setup Wizard to install the workstation software to your client computer.



Note: The DWB Setup Wizard will walk you through installation of the DWB client workstation and help locate the Insight.cfg and <PrimaryUser>.pem files to place in the required folders. The Insight.cfg file resides with the Insight.exe file in your installed client workstation. The <PrimaryUser>.pem file resides in the Certificates folder with the trust_ca_cert.pem file. All certificate and configuration files must be present for DWB to function.

For additional information, see the [DWB Setup Wizard](#).

Connection to your DWB Servers

Your managed servers are identified in the Insight.cfg file that you receive from Adobe Customer Care and resides on your client workstation. Adobe TechOps will set up your servers and Adobe Customer Care will add references to these managed servers and profiles to the Insight.cfg file before sending it to you. (The Configure connections to server in step 12 of the DWB Setup Wizard documentation will be completed.)



Note: In the Workstation Configuration workspace on the DWB client workstation, you will be able to see your connected servers and profiles.

Adobe Managed Services

-
- Adobe TechOps manages the infrastructure including network, data center, servers, and storage. Infrastructure monitoring and response to alerts occurs on a 24x7 basis for alerts requiring TechOps action. For other alerts, TechOps will notify Adobe Customer Care to coordinate with you.
 - Adobe TechOps will perform maintenance and firmware updates for your managed servers. For maintenance causing downtime, you will receive maintenance window notifications from Customer Care at least two weeks in advance. Adobe TechOps will address immediate needs as quickly as possible. Notification will depend on urgency and will be resolved once the schedule is known.
 - Adobe TechOps sets up a scheduled task to automatically manage data. Analytic feed data is moved into DWB for processing and transformation every evening beginning at 21:00 report suite time.
 - Adobe TechOps will process other Adobe Managed Services include data backups, FTP accounts, data archiving, and data transfer when necessary.
 - Adobe TechOps will configure the primary production cluster to contain three months of rolling data to be reset and reprocessed monthly. Updates to lookups (Geography, DeviceAtlas, Standard Classifications) will also occur as part of the reprocessing task. By default, the task runs on the first Friday of each month. If necessary, the schedule can be modified by Customer Care.

For additional information contact Adobe Customer Care at datworkbench@adobe.com or call 800.497.0335.

Data Workbench Architecture

Overview of the Data Workbench architecture.

- Overview of the [Schema Design Countable Structures](#) on page 1061.
- Overview of the [Dataset Schema Interface](#)
- Documentation for the Analytics (SiteCatalyst) schema in the [Dimensions and Metrics guide](#).

Schema Design Countable Structures

Explanation of Countables in Dataworkbench (DWB) for designing and implementing the schema.

Understanding Countable in Data Workbench

At the highest level are countable dimensions. Countable dimensions serve two major functions. First, they are dimensions whose elements you want to count. In other words, countables answers the questions like:

- How many visitors visited your homepage?
- How many visits came from Google.com?

Countable dimensions are typically used to create sum metrics, which return the count, or sum, of all the elements of the dimension. You can define countable dimensions to count instances such as reservation bookings or product orders. For example, you could define the countable dimension orders whose elements (log entries corresponding to orders from your online store) could be counted. If you want to show a count of orders within a visualization, you would define the orders sum metric, which can be evaluated over a dimension or have filters applied to it.

Countable dimensions can be parents of other dimensions or children of other countable dimensions.

Although your root countable dimension does not have to be associated with the tracking IDs in the data, Adobe recommends that you configure your dataset's root countable dimension to use the tracking ID field (x-trackingid) as its Key. As a result, each element of the root countable is associated with a unique value of x-trackingid, and all of the data about each element is grouped together.

Countable dimensions are defined by the following parameters:

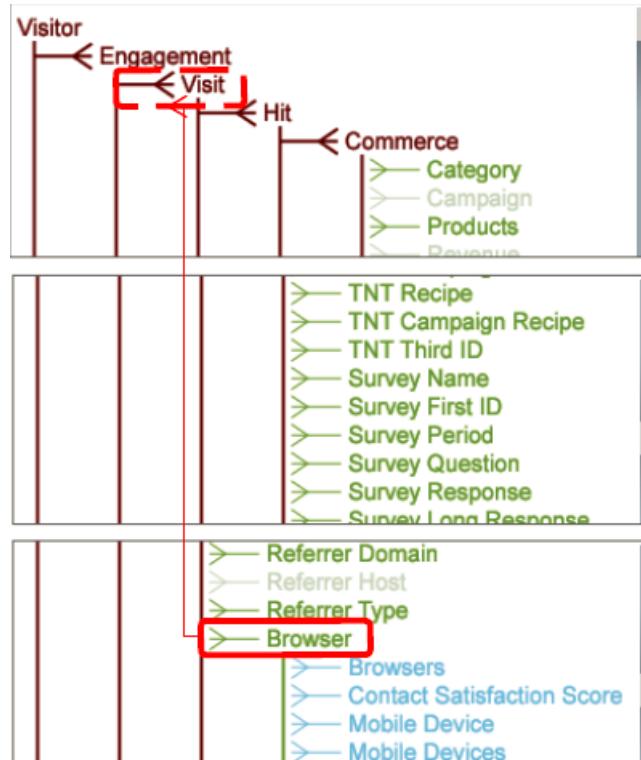
Parameter	Description	Default
Name	Descriptive name of the dimension as appears to the user in data workbench. The dimension name cannot include a hyphen (-).	
Comments	Optional. Notes about the extended dimension.	
Condition	The conditions under which the input field contributes to the creation of the countable dimension. If specified, a condition restricts the set of log entries visible to the dimension and all of its children in the dataset schema.	
Hidden	Determines whether the dimension appears in the data workbench interface. By default, this parameter is set to false. If, for example, the dimension is to be used only as the basis of a metric, you can set this parameter to true to hide the dimension from the data workbench display.	false
Key	Optional. The name of the field to use as the key. If you define this parameter, an element of the countable dimension exists for every combination of an element of the countable dimension's parent and a distinct value of the field specified as the key. Each element of the countable dimension is required to relate to a contiguous set of log entries. Therefore, if the log entries are not ordered by the key, an element of the countable dimension is created each time the key field changes. To prevent this situation, Adobe recommends that you use a unique key which is contiguous in time order.	
Parent	The name of the parent dimension. Any countable dimension can be a parent dimension. To make a dimension the top-level dimension in the dataset's schema, set the parameter to "root." The defined dimension becomes the root countable dimension for the dataset. For example, if you are working with Site, the Visitor dimension is the root countable dimension for your dataset. Note: Although your root countable dimension does not have to be associated with the tracking IDs in the data, Adobe recommends that you configure your dataset's root countable dimension to use the tracking ID field (x-trackingid) as its Key. As a result, each element of the root countable is associated with a unique value of x-trackingid, and all of the data about each element is grouped together. If you would like to configure your dataset differently, contact Adobe.	

This example illustrates the definition of a countable dimension using event data collected from website traffic. The countable dimension counts the web campaign events within a given session. The assumption is that all email campaign resources are requested from the web server with "email=" as part of cs-uri-query. In the example, the number of times that the visitor responds to an email campaign during a given session is of interest, not the actual value of the cs-uri-query(email) field.

-0	Countable
Name	Campaign Event
-Condition	AndCondition
-0	RECondition
-Matches	.+
0	true
Case Sensitive	Comment
Comments	cs-uri-query(email)
Input	
Comments	Comment
Hidden	false
Key	
Parent	Session

The second major function of countables is that they form the backbone of your dataset schema structure. Your data schema and all other dimensions are organized to be grouped under, and belong to a countable. In other words, if we consider dimensions as "categories", then countables are the way we organize these "categories" into groups.

When dimensions are grouped under a countable dimension, they are said to be at the "level" of the countable dimension. For example, in the figure below you can see that 'Email Address' is at the Visitor level and "Browser" is at the Visit level. "Parent" and "child" refer to the relationship between the countable and the dimensions grouped below it. For example, Visitor is a "parent" of Email address. Conversely, Email address is a "child" of Visitor.

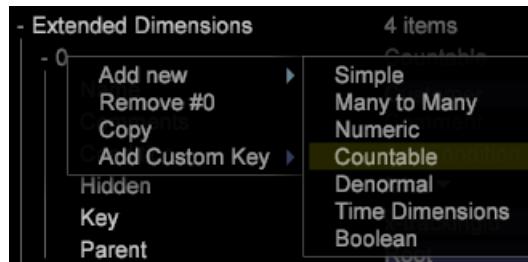




Creating Countable in Data Workbench

Perform the following steps to create the Countable in Dataworkbench:

1. Open Profile Manager
2. Under Transformation folder, create a config file and open it in workstation.
3. Under Extended Dimensions, right click and choose Add new -> Countable as shown below:



4. Enter Name for new Countable. In the example below, Customer Countable is defined. If it is the Highest level Countable, then in the Parent write Root.

- 0	Name	Countable
	Comments	Customer
	Condition	Comment
	Hidden	AndCondition
	Key	false ▾
	Parent	x-trackingid
		Root

If the Countable is not the top level one, then in the parent field give the name of the Parent Countable. In the example below, Engagement Countable is created and the Parent for this countable is Customer.

- 0	Name	Countable
	Comments	Customer
	Condition	Comment
	Hidden	AndCondition
	Key	false ▾
	Parent	x-trackingid
		Root

For additional information on the DWB architecture for schema design, countable structures, and offline data feed configurations, see the [Dataset Schema Interface](#) and [Dimensions and Metrics Reference](#).

Data Workbench Administration

Overview of the administration of Data Workbench.

This section provides basic information about access control, errors and alerts, and server upgrades when administrating DWB.

- For Access Control of User Groups, see [configuring access control](#).
- For Access Control of Server Groups, see [server access groups](#).
- For Errors and Alerts - Decoder Error Logging, see [Log File Properties](#).
- For Errors and Alerts - TnT Integration Audit Log, see [Master Marketing Profile Export](#) and [Exporting to Analytics Core Services](#).
- For Errors and Alerts - Administrative Alerts, see [Administrative Alerts](#).
- For Server Upgrades, see [Upgrading Data Workbench](#).

Data Workbench Configuration and Implementation

Documentation for configuring and implementing DWB.

- For **Data processing - Building the Primary Key**, see [Event Data Records](#) and [Data processing - Building Primary Key](#) on page 1085.
- For **Data processing - Setting up Event Time**, see [Event Data Records](#) and [Setting up Event Time](#) on page 1065.
- For **Data processing - Dimension Setup**, see [Extended Dimensions](#) and [Dimension Setup](#) on page 1068.
- For **Data Processing - Metric Setup**, see [Metrics Descriptions](#) and [Metrics Setup](#) on page 1071.
- For **Cluster Members in Access Control**, see [Configuring Access Control](#).
- For **Processing Servers in Profile Configuration File**, see [Configuring a Profile to Run on a Cluster](#).
- To **Set up Segment Export**, see [Exporting Data using Segment Export](#).
- To **Set up of Normalization Servers**, see the [File Server Configuration Process](#).
- To **Set up of cluster members in Addresses file**, see the [Address Files Installed on Insight Server](#).
- For **Validation of internal and external FTP set up**, see [Validation of Internal and External FTP Servers](#) on page 1072.
- To **Check Consistency of data feeds for historical dates**, see [Validating Historical Data Feeds](#) on page 1074.
- To **Set up and Schedule to Weekly Reprocess Logs**, see [Scripting to Weekly Reprocess](#) on page 1075.
- To **Set up Scripting for the SAINT Scrubber**, see [Scripting for the SAINT Scrubber](#) on page 1076.
- For **Report Server Setup**, see [DWB Report Server Setup](#).
- For **basic DWB Query API Setup instructions**, see [Query API Setup](#) on page 1081.
- For **basic DWB Query API Setup Syntax**, see [Query Language](#).

Setting up Event Time

This section explains how to create timestamps for a Data Workbench dataset.

Understanding the Event Time

Event Time is the date and time at which the request (or event) occurs.

Usually, for online data, `x_hit_time_gmt` is used as timestamp field. The time of the call can be used as the timestamp for offline data (such as call center data). This is a mandatory field and all the data sources should have one field in it that can be used as timestamp. This information should be provided by your organization.

In DWB, the following predefined variables capture the timestamp:

<i>x-timestamp</i>	The date and time (GMT) at which the request was received by the server. The time is expressed as the number of 100 nanoseconds since January 1, 1600. Example: 12771098932000000 would be the <i>x-timestamp</i> value for 11:28:52.0000000 on Tuesday, September 13, 2005.
<i>x-timestring</i>	<i>x-timestamp</i> in the format YYYY-MM-DD HH:MM:SS.mmmm.
<i>x-unixtime</i>	<i>x-unixtime</i> is the epoch time which represent the number of seconds since January 1st, 1970, at 00:00:01.

Based on the format of date field, x-timestamp or x-unixtime or x-timestring is used. For example, if the incoming data is in the format YYYY-MM-DD, then x-timestring is to be used.

The timestamp is defined in one of the formats and DWB internally generates the other two formats. Also, these are pre-defined DWB fields and the same name should not be used for any other field.

Time Zones Defined in DWB

If the date field contains any of the below mentioned time zones, DWB considers the entire row in that particular timezone. For example, one file has the date defined as 2015-01-01 00:00:00 gmt and another file has the value as 2015-01-01 00:00:00 cst, then the first file's date will be considered in GMT timezone whereas the second file's date will be in CST timezone.

Code	Time Zone
gmt	Greenwich Mean
est	Eastern Standard
edt	Eastern Daylight
cst	Central Standard
cdt	Central Daylight
mst	Mountain Standard
mdt	Mountain Daylight
pst	Pacific Standard
pdt	Pacific Daylight



Note: DWB only processes the Time Zones mentioned above.

Setting Custom Time Zones

DWB does not process the offset in the Time Zone. To consider the offset in Time Zone, the data should be formatted in that offset Time Zone.

Example: to consider the date format in CST time zone, the data should come in YYYY-MM-DD HH:MM:SS UTC +/-HHMM format from the client.

2015-10-18 05:00:00 UTC -0200

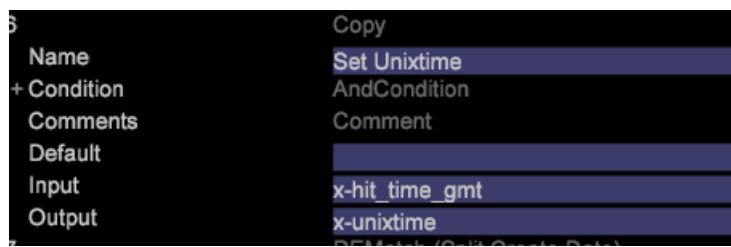
How to set Event Time/Timestamp

Based on the date field format, *x-timestamp*, *x-unixtime* or *x-timestring* variable is used. In the example below, since the *x-hit_time_gmt* comes in unix epoch format, *x-unixtime* is used.

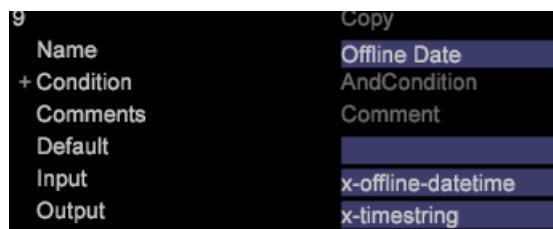
In the DWB foundation.cfg file (or any other configuration file under the Dataset log processing folder), use the Copy transformation to set the Event Time as shown:

Based on the date field format, *x-timestamp*, *x-unixtime* or *x-timestring* variable is used. In the example below, since the *x-hit_time_gmt* comes in unix epoch format, *x-unixtime* is used.

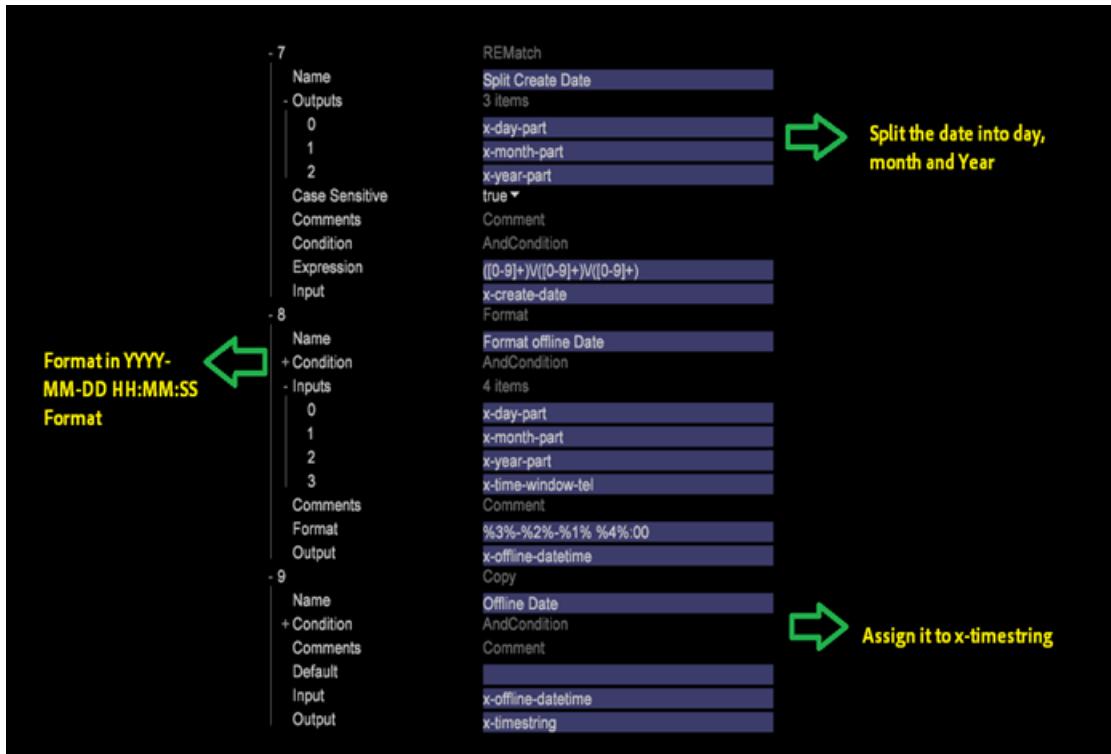
In the insight foundation.cfg (or any other config under Dataset log processing folder), use the Copy transformation to set the Event Time as shown below:



If the date is in YYYY-MM-DD HH:MM:SS.mmm format, *x-timestring* is used.



Example: If the date field is in the format other than defined in DWB say YYYY/MM/DD, then first format it in one of the timestamp format accepted by the DWB and then assign it to the corresponding variable. In the screenshot below, the date is first converted into YYYY-MM-DD format and then assigned to *x-timestring* variable.



Dimension Setup

This section explains the different types of Dimensions and how to set them up in DWB.

What are Dimensions

At the most basic level, dimensions are categories into which the data in the dataset can be broken down.

Best Practice: Dimensions in the data schema can be provided any name. Dimension names used and explained in this course are considered a best practice. Dimensions can be named differently. As you gain exposure to other datasets, you will begin to see differences in datasets. It is important to understand the purpose of the dimensions rather than their name. For example, whether it is called "Visitor", "Customer", "Person", "Consumer", or "User", it is important to understand that these are terms commonly used to refer to the highest level countable dimension that is being used to gather information about a singular person.

For complete information, see the [Dataset Configuration](#) guide.

Types of Dimensions in DWB

There are two types of dimensions in Data Workbench: Extended Dimensions and Derived Dimensions.

Extended Dimensions are created from fields in the "raw" data files. Extended dimensions are used to categorize "raw" data and to specify the relationships that exist among the data. Extended dimensions are created by Data Workbench Architects.

Derived Dimensions are created by a user "on the client- side" after the dataset has been processed using existing Extended dimension definitions. For example, based on the existing URI dimension, a user may choose to create a derived Page Name dimension, which displays a more user-friendly page name in place of a given URI. All dimensions consist of elements or items that have been categorized (grouped) together to form the dimension. Below are three dimensions and their elements.

Many derived dimensions are created automatically to drive different types of visualizations. For example, when a user builds a site or process map, DWB servers create a Prefix dimension. Others, such as the reporting time dimensions, are defined by files in the Dimensions directory of a profile.



Note: The elements appearing in any given dimension will only reflect those values that exist in records that have been chosen to be loaded into the dataset. For instance, if there is no data for "May '12", then that month will not appear in the 'Month' dimension.

Extended Dimensions

Types of Extended Dimensions

1) Countable Dimensions

At the highest level are countable dimensions. Countable dimensions serve two major functions. First, they are dimensions whose elements you want to count. In other words, countables answers the questions such as:

- "How many visitors visited your homepage?"
- "How many visits came from google.com?"

For this reason, countables are often used as the basic fundamental building block to create metrics.

The second major function of countables is that they form the backbone of your dataset schema structure. Your data schema and all other dimensions are organized to be grouped under and belong to a countable. In other words, if we consider dimensions as "categories", then countables are the way we organize these "categories" into groups.

When dimensions are grouped under a countable dimension, they are said to be at the "level" of the countable dimension. For example, the 'Email Address' can be at the Visitor level and "Browser" is at the Visit level. "Parent" and "child" refer to the relationship between the countable and the dimensions grouped below it. For example, Visitor is a "parent" of Email address. Conversely, Email address is a "child" of Visitor.

2) Simple Dimensions

The most common of all dimensions are Simple dimensions. Simple dimensions have a one-to-many relationship with a parent countable dimension and are commonly used in visualizations so you can view their elements. This means that a countable dimension can have one value for a simple dimension but the simple dimension can belong to one or more countables. For example, a customer has a name of 'John' - that customer can only have one first name, however, many other customers can have the name 'John'. As another example only one browser (e.g. Firefox) can be used for any particular visit to a web site but that browser can be used for many different visits.

If countable dimensions answer "How many?", then simple dimensions answer "Which ones?". Using the same example above used in the countable dimension section; Page Name is the simple dimension. Using the table and the simple dimension, Page Name, we can answer questions such as:

- "Which page had the most page views?"
- "Of all the Shopping Cart pages, which one had the most visits?".

3) Many-to-Many Dimensions

Many-to-many dimensions have a many-to-many relationship with a parent countable dimension. For example, if a dimension named External Search Term is at the Visit level; a given External Search Term may be used in one or many Visits, and a given Visit may include one or many External Search Terms. Thus, External Search Term is a many-to-many dimension.

4) Numeric Dimensions

Numeric dimensions are a type of simple dimension that has a numeric value. Numeric dimensions are often created to be used in metrics. Examples of numeric dimensions include 'Revenue', 'Orders', and 'Units'. In the example above, the 'Customer Orders' is a numeric dimension.

5) Denormal Dimensions

Denormal dimensions are dimensions that have a one-to-one relationship with a parent countable dimension. Denormal dimensions are often used with dimensions that have high cardinality (many unique elements) like identification data. For example, a visitor can only have one User ID and a User ID can only belong to one visitor. Thus, this is a one-to-one relationship and can be a denormal dimension.

For example, Geometrixx Web User ID is a denormal dimension at the customer level. Since it is denormal, it has a one-to-one relationship with its parent dimension, meaning that each Web User ID has one customer and each customer only has one Web User ID. Thus, the 'Customers' metric can only be '1' for each element of Geometrixx Web User ID.

6) Time Dimensions

Time dimensions allow you to create a set of periodic or absolute local time dimensions based on the timestamp field that you specify. Examples of time dimensions include 'Day', 'Hour', 'Week', and 'Hour of Day'. In the example above, the 'Hour of Day' table shows how many visits and page views were received during the different hours of the days.



Note: The % escapes used for display formatting is the same as the standard C library *strftime*.

Defining Extended Dimensions

Steps to define Extended Dimension:

1. While working in your dataset profile, open the Profile Manager and click Dataset to show its contents.
2. Open the Transformation.cfg file or the Transformation Dataset Include file in which you want to define the extended dimension.
3. Right -click Transformations and click Add new > <Extended dimension type>.
4. Input the appropriate information for your extended dimension. For descriptions of the transformation types and information about their parameters, see the following sections:
 - [Countable Dimensions](#)
 - [Simple Dimensions](#)
 - [Many-to-Many Dimensions](#)
 - [Numeric Dimensions](#)
 - [Denormal Dimensions](#)
 - [Time Dimensions](#)
5. For any extended dimension that you define, you can add one or more comment lines to the Comments parameter to further describe the dimension or add notes about its use. To add a comment, right-click the *Comments* label and click *Add new > Comment Line*.
6. After you have defined your extended dimension(s) in the configuration file, save the file locally and save it to your dataset profile on the DWB server.

Hiding Extended Dimensions

Extended Dimensions can be hidden so they do not show up on the Dimension Menu in the DWB. To hide the dimension, set the Hidden property to "True" in the dimension definition.

Metrics Setup

This section explains how to create metrics in Data Workbench.

Understanding Metrics

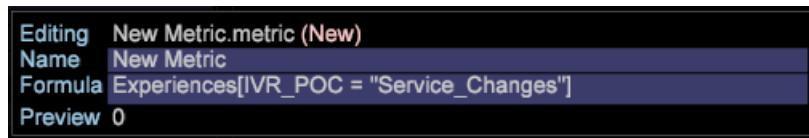
Metrics are quantitative information about customer activity, such as Views, Orders, number of calls made, and Revenue. Metrics are the foundation of reports and help you view and understand data relationships.

Metric Dimension allows you to group metric counts by a specific Level. It also allows you to group metric counts by a specific level.

Creating New Metrics

Follow the steps below to create a new metric:

1. Click **Tool > Metric Editor**.
2. In the metric editor, enter the new Metric name and formula.



3. Save it to Metrics folder.



Creating and Editing Derived Metrics

Use a Metric Editor to define a new metric by name, formula, and format, which is saved to the `User\profile_name\Metrics` folder for later use.

1. Open a new Metric Editor using the **Admin > Profile** menu option or by right-clicking the User column for the folder in which you want to create the metric and clicking **Create > New Metric**. A Metric Editor displays.
2. In the *Name* parameter, type a name for the new metric.



Note: Note that spaces () are allowed while underscores (_) are not. In addition, you cannot use the following symbols: + - * /

Editing	.metric (New)
Name	
Formula	
Preview	

Editing	Ratings.metric
Name	Ratings
Formula	sum(one, rating)
Preview	100,555,569

- In the *Formula* parameter, type an expression for the new metric.



Note: Filters must be defined within brackets [] in the expression. For additional metric expression syntax rules, see [Syntax for Metric Expressions](#).

This table provides sample expressions for extended metrics.

Extended Metric Name	Expression
Percent First Sessions	Sessions [Session_Number="1"]/Sessions
Conversion First Sessions	Conversion [Session_Number="1"]
Average Value Per Visitor	Value/Visitors



Note: When an appropriate expression is entered, the preview line displays the value of the new metric. If there is an error in the expression, the preview line displays an error message.

- Right-click and select **Save**. When you save the metric, a file representing the new metric is created on your computer in the DWB *Installation directory* \User\profile name\Metrics folder.

Editing Existing Derived Metrics

- In the Profile Manager or Metrics Manager, in the profile name column, right-click the check mark for the metric file that you want to edit and click **Make Local**.
- Right-click the check mark for the metric file in the User column and click **Open** from the workbench.



Note: You also can open a Metric Editor by right-clicking any metric-related area within a visualization to display the metric menu.

- In the **Metric Editor**, edit and save the metric definition as necessary using Steps 2-4 in *Creating New Derived Metrics*.

If you would like all users of the profile to use the metric that you edited, you must publish it to the working profile using the Profile Manager.

Please refer to the documentation for more assistance:

[Syntax for Metric Expressions](#)

[Creating and Editing Derived Metrics](#)

Validation of Internal and External FTP Servers

This is a quick guide giving you the minimum steps required to validate Internal and External FTP setup.

An Internal FTP is used when a consultant/architect internal to Adobe has to connect to the FTP site for file upload or download, whereas an External FTP is mainly for you as the user to upload the required data files.

For additional information on setting up FTP servers, see [File Transfer Protocol](#).

Validation Steps - External FTP

1. Open a Command Prompt. (Windows+R and type cmd)
2. Type ftp <ftp server>
3. Provide user name and password.

```
C:\Users\andixit>ftp ftp3.omniture.com
Connected to ftp3.omniture.com.
220 (vsFTPd 2.0.1)
User (ftp3.omniture.com:(none)): saintgobain_insight
331 Please specify the password.
Password:
230 Login successful.
ftp> ls
200 PORT command successful. Consider using PASV.
150 Here comes the directory listing.
dev
226 Directory send OK.
ftp: 5 bytes received in 0.00Seconds 5000.00Kbytes/sec.
```

4. Change local directory from where some file can be moved. Use this command:

```
ftp> lcd C:\Users\andixit\Desktop
local directory now C:\Users\andixit\Desktop.
```

5. Copy file from local to remote location.

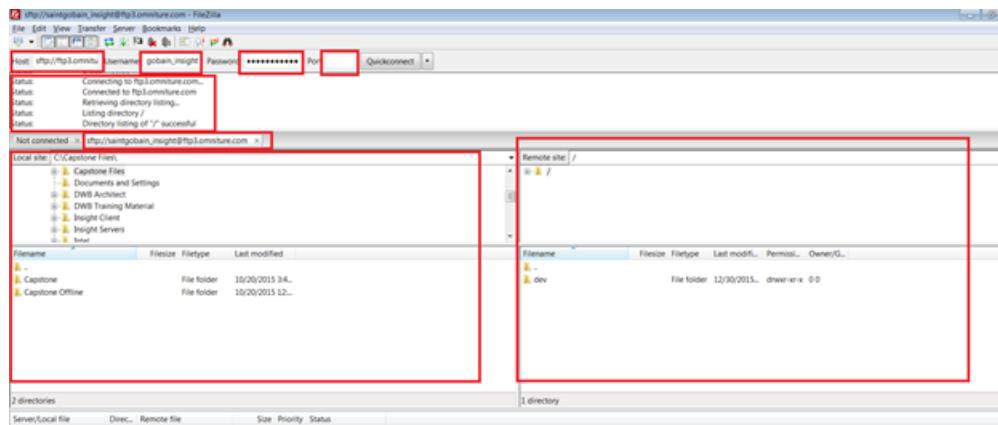
```
ftp> mput Casio.pdf
mput Casio.pdf? yes
200 PORT command successful. Consider using PASV.
150 Ok to send data.
226 File receive OK.
ftp: 440434 bytes sent in 4.30Seconds 102.36Kbytes/sec.
ftp> ls
200 PORT command successful. Consider using PASV.
150 Here comes the directory listing.
Casio.pdf
dev
226 Directory send OK.
ftp: 16 bytes received in 0.00Seconds 5.33Kbytes/sec.
```

6. Logout from remote server. (Use below command)

```
ftp> bye
221 Goodbye
```



Note: Another way to validate FTP is using Filezilla. Provide Host Name, User Name, Password and Port. Right side of panel is remote site and left side is local site. To validate FTP drag and drop files from local to remote site and v.v.



Validation Steps - Internal FTP

The above steps can be followed to validate internal ftp from any Adobe server.

Validating Historical Data Feeds

A quick guide for the minimum steps required to validate and set up historical data feeds.

Validation Steps for Consistency of Data Feeds

1. Login to *drteeth* (<https://oasis.omniture.com/drteeth/>)
2. Go to SiteCatalyst Admin -> Data Feed Definition(new)
3. Jump to server location (Ex. Dallas, London..) Depending on where your organization is located.
4. Provide RSID and select feed type Insight and click on *search*.

Search Data Feed Definitions

Feed ID:	<input type="text"/>
Feed Name:	<input type="text"/>
Billing Customer Name:	<input type="text"/>
RSID:	<input type="text"/> newscorpau-global
Feed State:	Select One
Feed Type:	Insight

Search

We are in the process of migrating old feeds to this new tool. If you do not find your feed here, you might be able to [find it in our old tool](#).

View 25 / page Page 1							
Feed ID	Feed Name	Report Suite	Billing Customer	Feed State	Feed Type	FTP Preset	Notes Actions
1903	newscorpau-global to SIN2	newscorpau-global (100012035)	NC TRANSACTION INC. (100002117)	Canceled	Insight	News Corp AU to SIN2	Edit View Copy History

5. Identify the actual feed name for your client.
6. Click on History in the Actions Section.

Data Feed History
Report Suite: newscontent_global(198012035)
Feed: 1988 /newscontent_global/882/2015NovW1 / Complete / hourly / Insight
Assigned Cluster: insight_cluster
Created by andrea on Thu Dec 24, 2015 11:42am
Last Edited by andrea on Thu Dec 24, 2015 11:42am

Request ID	Audit Log #	Cluster	Requested	Scheduled	Started	Finished	Run No.	Run Status	Actions	
1111216	Audit Log #1	insight_cluster	Sun Nov 1, 2015 12:00am AEDT	Thu Dec 24, 2015 11:43am GMT	Thu Dec 24, 2015 11:43am GMT	Thu Dec 24, 2015 12:00pm GMT	1	Complete	Resend	Reprocess
1111233	Audit Log #1	insight_cluster	Sun Nov 1, 2015 1:00am AEDT	Thu Dec 24, 2015 11:53am GMT	Thu Dec 24, 2015 11:53am GMT		1	Processing	Resend	Reprocess
1111344	Audit Log #1	insight_cluster	Sun Nov 1, 2015 2:00am AEDT	Thu Dec 24, 2015 12:04pm GMT	Thu Dec 24, 2015 12:04pm GMT	Thu Dec 24, 2015 12:24pm GMT	1	Complete	Resend	Reprocess
1111376	Audit Log #1	insight_cluster	Sun Nov 1, 2015 3:00am AEDT	Thu Dec 24, 2015 12:14pm GMT	Thu Dec 24, 2015 12:14pm GMT	Thu Dec 24, 2015 12:33pm GMT	1	Complete	Resend	Reprocess
1111401	Audit Log #1	insight_cluster	Sun Nov 1, 2015 4:00am AEDT	Thu Dec 24, 2015 12:24pm GMT	Thu Dec 24, 2015 12:24pm GMT	Thu Dec 24, 2015 12:43pm GMT	1	Complete	Resend	Reprocess
1111425	Audit Log #1	insight_cluster	Sun Nov 1, 2015 5:00am AEDT	Thu Dec 24, 2015 12:34pm GMT	Thu Dec 24, 2015 12:34pm GMT	Thu Dec 24, 2015 12:59pm GMT	1	Complete	Resend	Reprocess
1111457	Audit Log #1	insight_cluster	Sun Nov 1, 2015 6:00am AEDT	Thu Dec 24, 2015 12:45pm GMT	Thu Dec 24, 2015 12:45pm GMT	Thu Dec 24, 2015 12:55pm GMT	1	Complete	Resend	Reprocess
1111482	Audit Log #1	insight_cluster	Sun Nov 1, 2015 7:00am AEDT	Thu Dec 24, 2015 12:55pm GMT	Thu Dec 24, 2015 12:56pm GMT	Thu Dec 24, 2015 1:42pm GMT	1	Complete	Resend	Reprocess
1111593	Audit Log #1	insight_cluster	Sun Nov 1, 2015 8:00am AEDT	Thu Dec 24, 2015 1:05pm GMT	Thu Dec 24, 2015 1:06pm GMT	Thu Dec 24, 2015 1:51pm GMT	1	Complete	Resend	Reprocess
1111624	Audit Log #1	insight_cluster	Sun Nov 1, 2015 9:00am AEDT	Thu Dec 24, 2015 1:15pm GMT	Thu Dec 24, 2015 1:16pm GMT	Thu Dec 24, 2015 2:13pm GMT	1	Complete	Resend	Reprocess
1111649	Audit Log #1	insight_cluster	Sun Nov 1, 2015 10:00am AEDT	Thu Dec 24, 2015 1:25pm GMT	Thu Dec 24, 2015 1:25pm GMT	Thu Dec 24, 2015 2:44pm GMT	1	Complete	Resend	Reprocess
1111674	Audit Log #1	insight_cluster	Sun Nov 1, 2015 11:00am AEDT	Thu Dec 24, 2015 1:35pm GMT	Thu Dec 24, 2015 1:36pm GMT	Thu Dec 24, 2015 2:55pm GMT	1	Complete	Resend	Reprocess
1111704	Audit Log #1	insight_cluster	Sun Nov 1, 2015 12:00pm AEDT	Thu Dec 24, 2015 1:45pm GMT	Thu Dec 24, 2015 1:45pm GMT	Thu Dec 24, 2015 3:42pm GMT	1	Complete	Resend	Reprocess
1111728	Audit Log #1	insight_cluster	Sun Nov 1, 2015 1:00pm AEDT	Thu Dec 24, 2015 1:56pm GMT	Thu Dec 24, 2015 1:57pm GMT	Thu Dec 24, 2015 3:58pm GMT	1	Complete	Resend	Reprocess
1111839	Audit Log #1	insight_cluster	Sun Nov 1, 2015 2:00pm AEDT	Thu Dec 24, 2015 2:07pm GMT	Thu Dec 24, 2015 2:07pm GMT	Thu Dec 24, 2015 4:17pm GMT	1	Complete	Resend	Reprocess
1111869	Audit Log #1	insight_cluster	Sun Nov 1, 2015 3:00pm AEDT	Thu Dec 24, 2015 2:17pm GMT	Thu Dec 24, 2015 2:17pm GMT	Thu Dec 24, 2015 4:51pm GMT	1	Complete	Resend	Reprocess
1111893	Audit Log #1	insight_cluster	Sun Nov 1, 2015 4:00pm AEDT	Thu Dec 24, 2015 2:27pm GMT	Thu Dec 24, 2015 2:27pm GMT	Thu Dec 24, 2015 5:00pm GMT	1	Complete	Resend	Reprocess
1111917	Audit Log #1	insight_cluster	Sun Nov 1, 2015 5:00pm AEDT	Thu Dec 24, 2015 2:38pm GMT	Thu Dec 24, 2015 2:38pm GMT	Thu Dec 24, 2015 5:49pm GMT	1	Complete	Resend	Reprocess
1111947	Audit Log #1	insight_cluster	Sun Nov 1, 2015 6:00pm AEDT	Thu Dec 24, 2015 2:48pm GMT	Thu Dec 24, 2015 2:48pm GMT	Thu Dec 24, 2015 6:11pm GMT	1	Complete	Resend	Reprocess

Check the status field for any errors and In case some feed is in error state, select the feed and click reprocess. If the error occurred for multiple requests then send an email to datworkbench@adobe.com with Feed ID and Report Suit details to reprocess.

- Post-validation checks the logs in the raw folder of the NAS location.

Scripting to Weekly Reprocess

This section is a quick guide giving you the minimum steps required to set up and schedule scripts to weekly reprocess your log files. This can be used as reference guide to set up or modify your profiles.

What is Reprocess

Loading the data to DWB based on changes in data sources, offline data sources, or time period. The script will reprocess the start date parameter in *Log Processing.cfg* file.

Prerequisites

Report Suite ID, Number of month data should be available in DWB. The Perl64 folder should be available in your C:\ drive.

Reprocess Logs

Provide above details (prerequisites) in windows command script *Reprocess.bat* available at folder \scripts\Log Processing at Main FSU server.

This script will internally call two client specific scripts: One to reprocess the data and other for email alert. These two scripts are also available in the \scripts\Log Processing folder.

The script will change reprocess parameters in the *Log Processing.cfg* file.

Rolling Window for Logs

Provide details (prerequisites) in windows command script *logprocessingdate.bat* available at folder \scripts\Scriptrepository on the maing FSU server. This script will internally call two client specific scripts: One to setting up start date of logs and another for email alerts. These two scripts are also available in \scripts\Scriptrepository

Provide Report suite id and number of months in the *logprocessingdate.bat* file.

The script will change the start date parameter in *Log Processing.cfg*.



Note: If the *Scripository* folder is not available, follow the below process to copy the *Scripository* folder and make changes in above files using details specific for customer. And provide your email address to get an alert in case of any error.

Scheduling Scripts

Follow these steps to schedule the scripts in the Windows task scheduler.

1. Schedule the script in Windows' task scheduler.
 - Open Task Scheduler: Right click on the **Task Scheduler Library** and click **Create Task**.
 - In the **General** tab provide a task name and select **Options**.
 - Under the **Triggers** tab, click **New** and a new window will open.
 - Under the **Actions** tab, click **New** and a new window will open. Then provide script details and other options. (Start in will have a path where script is placed).
2. Validation: Right click and run the job and verify changes in the *Log processing.cfg* file. An email will be sent to the email id provided in the script.

Scripting for the SAINT Scrubber

This section explains the Saint Scrubber script.

Overview of SAINT Classification

Classification is also known by the acronym SAINT for SiteCatalyst Attribute Importing and Naming Tool.

When we "classify" a SiteCatalyst variable, you are establishing a relationship between a variable and meta-data related to that variable. Classifications are most frequently used in the Campaigns area so I will use that as a way to explain them. Most clients send campaign traffic to their site using a tracking code. This tracking code is an identifier that may represent a specific keyword purchased on Google, such as "goog123." This identifier is passed into the *s.campaigns* variable so you can see what site success events take place after visitors come to your site from that campaign code.

But what if, instead of viewing Campaigns just by the tracking code, you want to see campaign results by Search Engine or Keyword or Campaign Channel? Do you have to create a new conversion variable for Search Engine, another for Keyword and yet another for Campaign Channel? If so, you would use up many of your fifty variables on Campaigns alone! Thankfully, you can use Classifications to make your life easier! Since each tracking code could have a Search Engine, Keyword or Campaign Channel, you can simply create three Classifications of the Campaigns variable to represent each. You are essentially telling SiteCatalyst that there is a direct relationship between the Campaigns variable and these three other "meta-data" values. By doing this, SiteCatalyst will allow you to slice and dice site Success Events by all four variables with no additional tagging.

SAINT scrubber Script in DWB

This script is used when you bring in any SAINT Classification data into DWB. The script *SaintScrubber.dat* is normally placed under the *\Scripts\Scripository* folder on the FSU.

The main purpose of this script is to remove the header in the SAINT Classification files. Also, it counts the number of columns mentioned in the column header line and checks all the data rows. If there are rows with less or more number of columns, then it removes these rows from the file.

The *SaintScrubber.dat* internally calls the *saint_scrubber.pl* script. Below are the details for this script file:

Path: *E:\Scripts\Scripository\Library\Perl*

Script Arguments:

1. Input Folder (Mandatory): source_directory
2. Output Folder (Mandatory) : destination_directory
3. Delimiter (Mandatory) : delimiter
4. Reject Folder (Optional)(Parameter can be left blank or omitted from the command line)
5. Log Folder (Optional)(Parameter can be left blank or omitted from the command line)

Steps performed in the Perl script:

1. Replace escaped form feeds, newlines, carriage returns, tabs with spaces.
2. Remove those double bytes that are interpreted as a control character in the UTF-8 BMP (Basic Multilingual Plane) except for:
 - 9 horizontal tab
 - 10 line feed
 - 12 form feed
 - 13 carriage return
 - Use the pipe keyword, as a delimiter for | eg: delimiter pipe
 - Remove other troublesome characters
 - After the above scrubbing drop any lines with a number of columns differing from the first data line (not blank or comment)
 - Support optional rejection files to hold rejected lines instead of just skipping them
 - Support recursive input folder; generate output folders of same structure
 - Move processed input files to processed sub folders so the script does not repeat the effort when run again on the same existing input folder
 - Recognize date in workbench filenames; sort processing first by date then alpha—regardless of folder name. This will ensure that the sequence is correct no matter the workbench file type (ecom, non-ecom) or report suite ID (if you are processing multiple report suites into a single Insight dataset).
 - Support email alerts

File Transfer Governance

A quick guide for different file transfer methods in DWB.

File Tranfer Governance is a standard process to transfer files from an internal directory to any other server or internal file movement.

Different File Tranfer Methods

1. AWS (Amazon Web Services)
 - a. Raise a Ticket to install AWS command line interface on server if not already installed (see <http://docs.aws.amazon.com/cli/latest/userguide/installing.html>).
 - b. How to Check? Try to configure the AWS using command prompt (see <http://docs.aws.amazon.com/cli/latest/userguide/cli-chap-getting-started.html>).
2. Transfer files from FTP server to NAS directory.
 - a. FTP offline feeds from ftp server to NAS directory. The details below are required for FTP.
ftp_username
ftp_password
ftp_port

```
ftp_address
ftp_directory
delete_ftp_files
ftp_file_extension
local_directory
```

(FTP Details will be available in project checklist. Use external ftp user for transferring the files)

- b. Use ftp_winscp_get.pl script attached below and schedule based on requirement.

```
ftp_winscp_get.pl
```

This script should be placed at E:\scripts\Scripository\Library\Perl



Note: If Scripository folder is not available see [Scripting to Weekly Reprocess](#) on page 1075 to download the folder.

- c. Schedule the script based on files availability at ftp_address.
 - d. Naming convention of file should be YYYYMMDD-<offline_feed_name>-00.*
3. Transfer files from NAS directory to FTP server.

- a. Use ftp_winscp_put.pl script and schedule based on requirement.

This script should be placed at E:\scripts\Scripository\Library\Perl

The details below are required to run the script.

```
ftp_username
ftp_password
ftp_port
ftp_address
ftp_directory
delete_ftp_files
ftp_file_extension
local_directory
```

```
#####
# PLUGIN NAME HERE
my $pluginname = "FTP WinSCP";
# 20140421 Script tp put files on the FTP
#####
# INCLUDE SCRIPPOSITORY CORE
use FindBin;                      # locate this script
BEGIN {push @INC, $FindBin::Bin}
require 'core.pl';

# check for the required parameters
GetOptions('local_directory:s'      => \$local_directory,
          'source_file_pattern:s' => \$source_file_pattern,
          'ftp_file_extension:s'  => \$ftp_file_extension,
          'ftp_address=s'        => \$ftp_address,
          'ftp_username=s'       => \$ftp_username,
          'ftp_password:s'       => \$ftp_password,
          'ftp_port:s'           => \$ftp_port,
```

```

        'ftp_directory:s'      => \$ftp_directory,
        'log_directory:s'     => \$log_directory,
        'error_directory:s'   => \$error_directory,
        'mail_from:s'         => \$mail_from,
        'mail_to:s'           => \$mail_to,
        'host:s'              => \$host,
        'trigger_directory:s' => \$trigger_directory,
        'trigger_file_extension:s' => \$trigger_file_extension,
        'delete_ftp_files:s'  => \$delete_ftp_files,);

# FOR LEFT OVER PARAMS, WE CAN CHECK GLOBAL PARAMS
check_parameters(@ARGV);

my $ftp_winscp_script = "winscpscript.txt";
if (index($trigger_file_extension, '.') != -1) {
    my @trigger_file_extension1=split(/./,$trigger_file_extension,2);
    $trigger_file_extension =  $trigger_file_extension1[1];
}
if (index($ftp_file_extension, '.') != -1) {
    my @ftp_file_extension1=split(/./,$ftp_file_extension,2);
    $ftp_file_extension =  $ftp_file_extension1[1];
}
if ($trigger_file_extension ne "_empty_" && $trigger_directory ne
"_empty_") {
    print $trigger_file_extension;
    my $ftp_winscp_trigger_script = "winscpscript_trigger.txt";
    cat winscpscript_trigger.txt > $ftp_winscp_trigger_script;
    sleep(10);
    system("\"E:\\\\Scripts\\\\Scripository\\\\Library\\\\WinSCP\\\\WinSCP.exe\""
    /console /script=$ftp_winscp_trigger_script /log=$logfile");
    $files =
getFiles($trigger_directory,$trigger_file_extension,$days_ago,$months_ago);

    my $ftp_file_pattern="";
    my $numberoffiles = @$files;
    my $i=0;
    foreach my $trigger_file(@$files) {
        $i++;
        my $file_string=substr($trigger_file,length($trigger_directory),
length($trigger_file)-length($trigger_directory));
        my @file_string1=split(/./, $file_string, 2);
        if ($i == $numberoffiles) {
            $ftp_file_pattern.=$file_string1[0]..".$ftp_file_extension;
        }
        else {
            $ftp_file_pattern.=$file_string1[0]..".$ftp_file_extension.", ";
        }
    }

    #unlink($ftp_winscp_trigger_script);
    print $local_directory;
    print $trigger_directory;

    cat winscpscript_trigger.txt > $ftp_winscp_trigger_script;
    runLogger("{$pluginname}: Sleeping for 10 sec to give enough time for

```

```

the temp script to be available");
sleep(10);
runLogger("{$pluginname: FTP started}");
system("\"E:\\Scripts\\Scripository\\Library\\WinSCP\\WinSCP.exe\"/console /script=$ftp_winscp_script /log=$logfile");
runLogger("{$pluginname: FTP ended}");

}

else {
    if ($source_file_pattern eq "_empty_") {
        $source_file_pattern="";
    }
    else {
        if (index($source_file_pattern, '.') != -1) {
            my @source_file=split(/./,$source_file_pattern,2);
            $source_file_pattern = $source_file[0];
        }
    }
    $ftp_file_extension=". ".$ftp_file_extension;
print $local_directory;

create_winscp_script($ftp_script,$ftp_port,$ftp_username,$ftp_password,$ftp_address,$local_directory,$ftp_directory,$delete_ftp_files,$file_pattern,$file_extension);

runLogger("{$pluginname: Sleeping for 10 sec to give enough time for
the temp script to be available}");
sleep(10);
runLogger("{$pluginname: FTP started}");
system("\"E:\\Scripts\\Scripository\\Library\\WinSCP\\WinSCP.exe\"/console /script=$ftp_winscp_script /log=$logfile");
runLogger("{$pluginname: FTP ended}");
}
unlink($ftp_winscp_script);

sub create_winscp_script() {
    my
($ftp_script,$ftp_port,$ftp_username,$ftp_password,$ftp_address,$local_directory,$ftp_directory,$delete_ftp_files,$file_pattern,
$file_extension) = @_;
open (FTP, "> $ftp_script") or die "Can't open log: $!";
print FTP "# Automatically answer all prompts negatively not to
stall\n";
print FTP "option batch on\n\n";
print FTP "# Disable overwrite confirmations that conflict with the
previous\n";
print FTP "option confirm off\n\n";
print FTP "# Connect using a password\n";
if ($ftp_port eq "22") {
    print FTP "open sftp://$ftp_username:$ftp_password@$ftp_address\n\n";
}
else {
    print FTP "open ftp://$ftp_username:$ftp_password@$ftp_address\n\n";
}
print FTP "# Change local directory\n";
print FTP "lcd \\\"$local_directory\\\"\n\n";
print FTP "# Change remote directory\n";
if ($ftp_directory eq "_empty_") {
}
else {
    print FTP "cd \\\"$ftp_directory\\\"\n\n";
}
}

```

```

}
print FTP "\# Force binary mode transfer\n";
print FTP "option transfer binary\n\n";
print FTP "\# Download the file to specified directory\n";
my @get_files=split(//,$file_pattern);
foreach my $file (@get_files){
  if ($delete_ftp_files eq "Y" || $delete_ftp_files eq "Yes") {
    print FTP "put -nopreservetime -nopermissions -delete
$file$file_extension\n";
  }
  else {
    print FTP "put -nopreservetime -nopermissions
$file$file_extension\n";
  }
}

print FTP "\n\n";
print FTP "\# Disconnect\n";
print FTP "close\n\n";
print FTP "\# Exit WinSCP\n";
print FTP "exit\n\n";
close(FTP);
runLogger("$pluginname: creating temporary winscp file");
}

```

- b. Schedule the script based on files availability at ftp_address.
 - c. Naming convention of file should be YYYYMMDD-<offline_feed_name>-00.*
4. Transfer Files from one NAS directory to other NAS directory.
- a. Copy and paste file directly connecting to one NAS directory from other. Follow below process:)

login to server -> go to Run -> \\server_name\E\$ [new folder will open and directly copy or move the files]

 - b. Use "copy_files.pl" script to copy files from one server to other or "move_files.pl" to move files from one server to other. (These files are available in E:\scripts\Scripository)

Query API Setup

A quick guide for setting up a Query API.

Follow the below steps for setting up the Query API:

1. Query API Certificate Acquisition

Send an Email to the Tech Ops Team of Adobe Email – Dataworkbench@adobe.com.

Please provide the CN name you want to use for the Query API(provide a generic name like <Client> Query API).



Note: Tech Ops will generate the certificate and upload it in a URL. Please let the Adobe Consultants know after receiving the notification from Tech Ops on successful generation of the ticket so that the ticket will be sent to you by them back.

2. Downloading and Extracting the API Stunnel. Receive the api-stunnel file from your consultant.

Make sure Perl is installed on your machine.

In the extracted folder (the folder path where you copy the file), copy your Query API certificate inside the *stunnel* folder.

3. Configure the Stunnel.conf

There should be a file called *stunnel.conf* inside the *Stunnel* folder (where you copied your certificate).

Edit the file in Notepad.

```
1 cert=COMMON_NAME.pem
2 CAfile=trust_ca_cert.pem
3 verify = 2
4
5 [insight-PROFILE]
6 client=yes
7 accept = 8002
8 connect = DOMAIN.MYSITE.COM:443
9
```

Change the parameters as follows:

```
1 cert=Aadhithiya Ramani QAPI Client.pem
2 CAfile=trust_ca_cert.pem
3 verify = 2
4
5 [insight-PROFILE]
6 client=yes
7 accept = 8002
8 connect = lon5dpu365.insight.omniture.com:443
```

Two parameters need to be changed in this file.

- *Cert* = The name on your certificate. In this example it is Aadhithiya Ramani QAPI Client.pem.
- *Connect* =The server name for your main DPU.

4. Copying the *Query.pm*.

The *Query.pm* file will be available in the Insight API Folder.

Copy the *Query.pm* file and paste it in your Perl Library folder(usually it will be *C:\Perl64\lib* , but check where the Perl is installed in your machine).

5. Modify the *api-http.pl* file

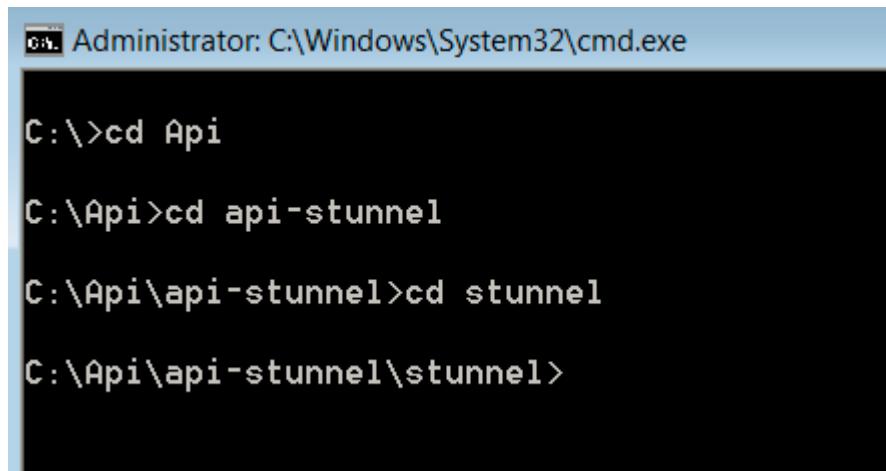
The *api-http.pl* file will be available in the *api-stunnel* folder.

Only one parameter to be modified

My \$profile = The profile name for which you are configuring the Query API.

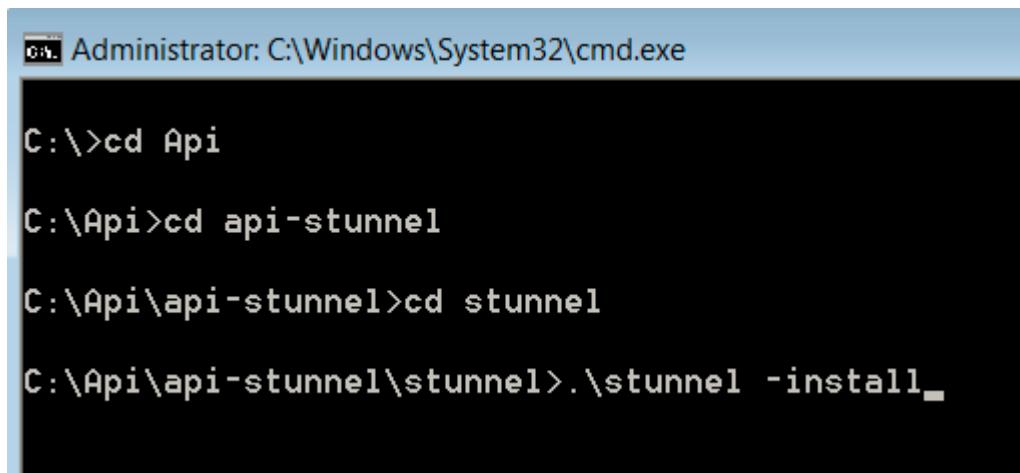
6. Instal the Query API.

Open the command prompt in your system as "Administrator" and navigate to the directory where you extracted the *stunnel* as shown:



```
C:\>cd Api
C:\Api>cd api-stunnel
C:\Api\api-stunnel>cd stunnel
C:\Api\api-stunnel\stunnel>
```

Run the following command `.\stunnel -install`.



```
C:\>cd Api
C:\Api>cd api-stunnel
C:\Api\api-stunnel>cd stunnel
C:\Api\api-stunnel\stunnel>.\stunnel -install_>
```

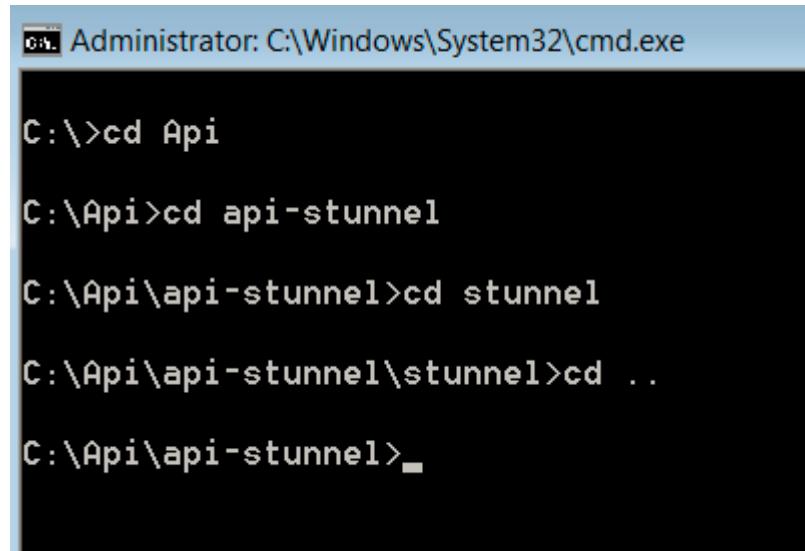
After executing the command a window will pop stating that the *stunnel* is installed.



Note: After executing the command a window will pop stating that the *stunnel* is installed.

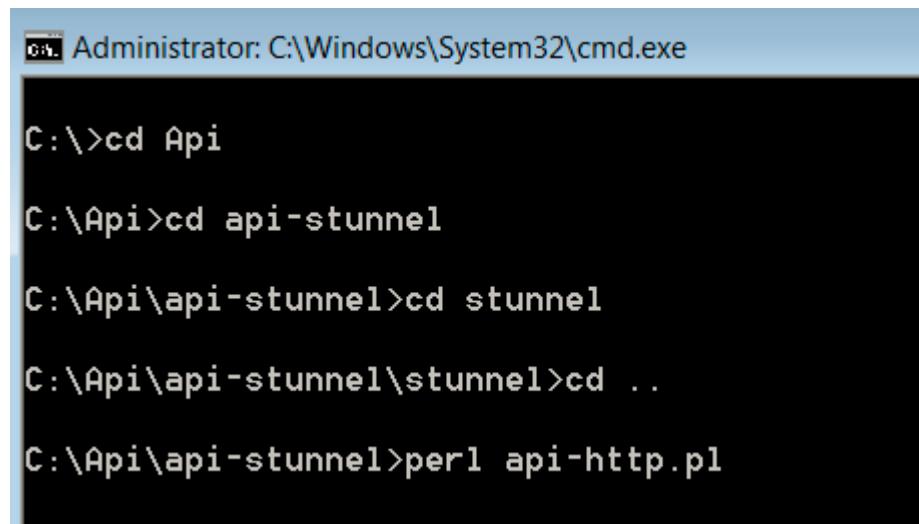
7. Testing the Query API stunnel configuration

The Final step of this process will be to test the Query API configuration. In the command prompt which you used for installing the api-stunnel directory.



```
Administrator: C:\Windows\System32\cmd.exe
C:\>cd Api
C:\Api>cd api-stunnel
C:\Api\api-stunnel>cd stunnel
C:\Api\api-stunnel\stunnel>cd ..
C:\Api\api-stunnel>
```

Run the Perl script available in that folder using the following command *perl api-http.pl*.



```
Administrator: C:\Windows\System32\cmd.exe
C:\>cd Api
C:\Api>cd api-stunnel
C:\Api\api-stunnel>cd stunnel
C:\Api\api-stunnel\stunnel>cd ..
C:\Api\api-stunnel>perl api-http.pl
```

After running the script the results should be like the screenshot below (the date time and values in the result will vary according to the as of time and other parameters in the profile on which you have configured the Query API (in step 6)).



Data processing - Building Primary Key

This section explains how to create Primary keys (Tracking ID) for Data Workbench datasets for schema design and implementation.

Understanding Tracking ID

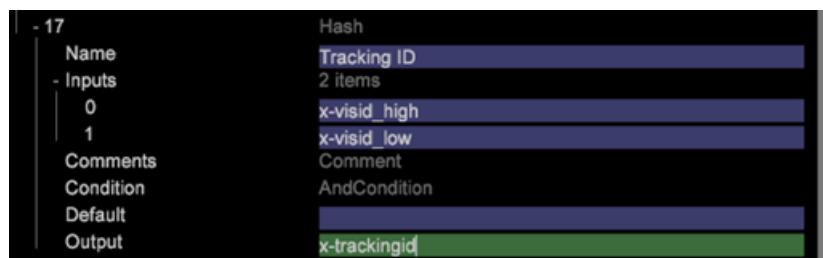
After reading and decoding the data in DWB (using decoders), the first step is to define the Tracking Id and Timestamp. The Tracking ID is an identifier that uniquely identifies a Customer record. It can be any field in the feed like email ID, Social Security Number, Cookie ID, etc. The field to be used as Tracking ID is decided by the client during the discovery session. Tracking ID and Timestamp are mandatory fields and have to be defined for each record.

Usually, for online Data, Cookie ID (combination of *x-visid_high* and *x-visid_low*) is used as the default mechanism for unique Customer Identification, however, this can be changed as per Client's requirement. The date and time at which the request (or event) occurs is the *x-timestamp*. All records in DWB are group by *trackingid* and sorted on timestamp. The Required Field Definitions .cfg file is a Log Processing Dataset Include file that defines the required fields : *x-trackingid* and *x-timestamp*.

Note: *x-trackingid* in DWB is an inbuilt field and this name should not be used for any other field.

Example 1: Creating *x-trackingid* using Cookie ID (when only online data is used)

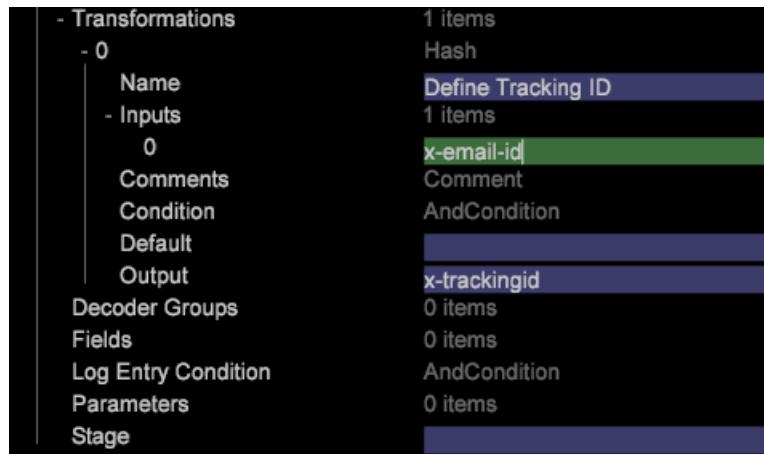
To create the *x-trackingid* in DWB using Cookie ID, use the Hash function to create the *x-trackingid* in the *foundation.cfg* file (it is a best practice to define tracking ID in *foundation.cfg* but it can be defined in any other configuration file under *Dataset > log processing* folder) as shown:



Example 2: Creating *x-trackingid* using Email ID (when both online and offline data is available)

Assuming, both offline and online data is available (for this example), and the Email ID is available in both the data sources. Since, the Email ID uniquely identifies a customer, it will be used to create the *x-trackingid*.

Use the Hash function to create the *trackingId* as shown:



Data Workbench Implementation of Features

After setting up DWB, you can implement these features.

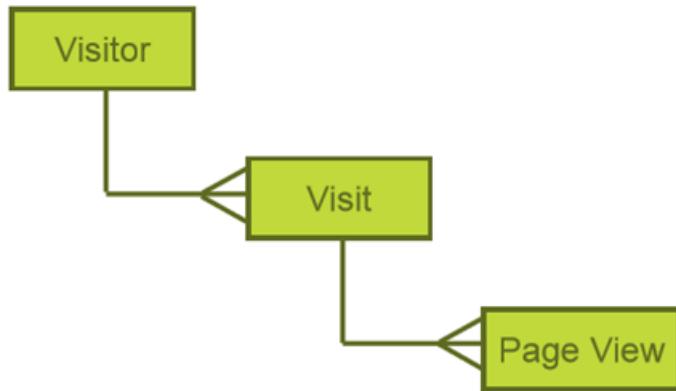
- To Change visualizations for a new schema, see [Analysis Visualizations](#) and [Configuration Changes for New Schema](#) on page 1086.
- For Derived Dimensions in workspaces and visualizations, see the [Derived Dimensions Setup](#) on page 1091.
- For Metric setup, see the [Metrics Setup](#) on page 1071 document.
- To implement the **Globe visualization**, see [Globes](#).
- To implement **Workspaces for Attribution**, see [Rules-based Attribution](#) and [Best Fit Attribution](#).
- To implement **Predictive Analytics** workspaces, see [Working with Visualizations](#) and the visualization listed below.
- To implement **Predictive Intelligence with Clustering visualizations**, see [Clustering](#).
- To implement **Predictive Intelligence with Propensity Scoring**, see [Propensity Scoring](#).
- To implement **Predictive Intelligence with a Lift/Gain Chart**, see [Lift/Gain Chart](#).
- To implement **Predictive Intelligence with Decision Trees**, see [Decision Trees](#).
- To implement **Predictive Intelligence with Binomial Correlation**, see [Binomial Correlation](#).
- To implement **Predictive Intelligence with Binomial Correlation**, see [Correlation Matrix](#).
- To set up and validate the DWB Dashboard, go to the [Adobe Data Workbench Licensing Portal](#). Also, see the [Basic Onboarding Instructions for DWB Managed Services](#) on page 1059.

Configuration Changes for New Schema

This document explains how to change all the default Data Workbench configurations after the new schema is in place.

Understanding the Dataset Schema

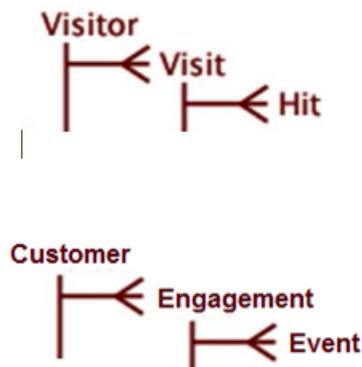
The foundation of the dataset schema consists of a key set of relationships that form the backbone of the Data Workbench web analytics schema. In the following example, typical web analytic schema provides an idea of the relationships between a visitor, a visit, and a page view.



- Any given visitor may have one or more visits.
- Any given visit is generated by only one visitor.
- Any given visit may include one or more page views.
- Any given page view belongs to only one visit.

As the web and the business world has evolved over the years the needs of web and data analysis have changed too. Web sites began as places to look at content. Now, you can look at content; interactively correspond through chats, video, or submissions; purchase products; and much more. In addition, businesses now want to integrate their web data with other channels of data in their business to gain a better view of their business as a whole. For example, a business may want to integrate their web, call center, email, social, and store and customer data together. With this integration of offline and online channels the dataset schemas have evolved over the years where no two dataset schemas are the same.

When you integrate the online and offline data, the term "visitor" doesn't always seem appropriate. As a result, the term "customer" is sometimes used instead of the visitor.



The 'Engagement' level is used to enable a single view of time, when you have data from multiple data sources. For example, suppose you only have a single data source: e-commerce data collected by visitor activity on your web site. In that case, the Visit level indicates visits to your site from those visitors. Note that time dimensions – 'Day', 'Week', 'Month', etc. – are typically captured at the 'Visit' level.

Similarly "Event" level brings all the events (page view, call made to call center etc) happened during an engagement. It combines all the online and offline events for a customer during an engagement.

New Countable Structure in DWB

The new schema structure replaces Visitor by Customer, Visit by Engagement and Hit by Event.

Search:	
-Countables.cfg (modified)	
- Transformation Include	TransformationInclude
- Extended Dimensions	
-0	Countable
Name	Customer
Comments	Comment
Condition	AndCondition
Hidden	false
Key	x-trackingid
Parent	Root
-1	Countable
Name	Engagement
Comments	Comment
Condition	AndCondition
Hidden	false
Key	x-engagementkey
Parent	Customer
-2	Countable
Name	Activity
+Condition	AndCondition
Comments	Comment
Hidden	false
Key	
Parent	Engagement

Configuration Changes as per New Dataset Schema

To change the dataset schema from visitor to customer, you need to change the following configuration files:

1. All configuration files under Dataset folder where countable and extended dimensions are defined.

New Layout		Original Layout
Search:		
- Countables.cfg		
- Transformation Include	TransformationInclude	4 items
- Extended Dimensions		Countable
- 0	Customer	Customer
Name	Comment	Comment
Comments	AndCondition	AndCondition
Condition	false ▾	false ▾
Hidden	x-trackingid	x-trackingid
Key	Root	Root
- 1	Engagement	Countable
Name	Comment	Comment
Comments	AndCondition	AndCondition
Condition	false ▾	false ▾
Hidden	x-visit_num	x-visit_num
Key	Customer	Customer
Parent	Countable (Event)	Countable (Product)
+ 2	1 items	1 items
+ 3	0 items	0 items
+ Transformations		
Parameters		
Stage		

2. Configuration files under Dimension folder, where the "visitor", "visit" or "event" are used as Level.

Example: Campaign.cfg file. In the Adobe SC profile, Campaign is defined at Visit level.

The screenshot shows the 'New Layout' tab selected. The left sidebar lists sections: - Campaign.cfg, - Transformation Include, - Extended Dimensions, - 0, + Condition, Case Sensitive, Comments, Hidden, Input, Load File, Maximum Elements, Operation, Parent, + Transformations, Parameters, and Stage. The right panel shows detailed settings for 'TransformationInclude': TransformationInclude (1 items), Simple, Campaign (highlighted in blue), AndCondition, true (dropdown), Comment, false (dropdown), x-campaign, 0, First Nonblank, Visit (highlighted in green), 2 items, and 0 items.

The following example provides an idea of the parent schema change from Visit to Engagement:

The screenshot shows the 'New Layout' tab selected. The left sidebar lists sections: - Campaign.cfg, - Transformation Include, - Extended Dimensions, - 0, + Condition, Case Sensitive, Comments, Hidden, Input, Load File, Maximum Elements, Operation, Parent, + Transformations, Parameters, and Stage. The right panel shows detailed settings for 'TransformationInclude': TransformationInclude (1 items), Simple, Campaign, AndCondition, true, Comment, false, x-campaign, 0, First Nonblank, Engagement (highlighted in green), 2 items, and 0 items.

3. As some of the metrics are derived or created out of countables, configuration files under Metrics folder needs to be modified or created.

For example: create a new metric `Customers.metric` with `formula = sum(one, customer)` or as Page Views.metric to *definw* it at the hit level. Modify the metric, and then change the level to Event instead of Hit.

Adobe SC Page Views Metirc defined at Hits level:

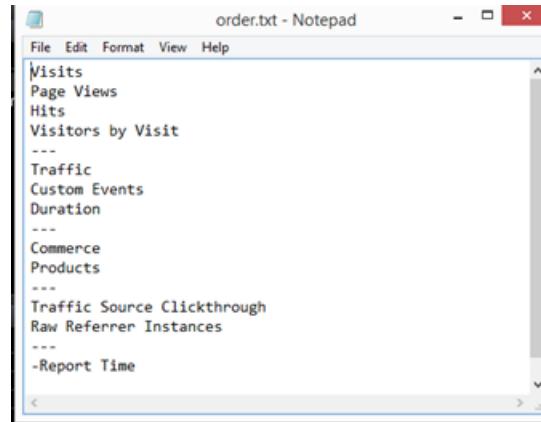
The screenshot shows the 'Editing' tab selected. The configuration details are: Name (Page Views) and Formula (Hits[Page_event_type="0"]).

The following will be the Page Views metric as per the new schema:

Editing	Page Views.metric
Name	Page Views
Formula	Events[Page_event_type="0"]
Preview	0

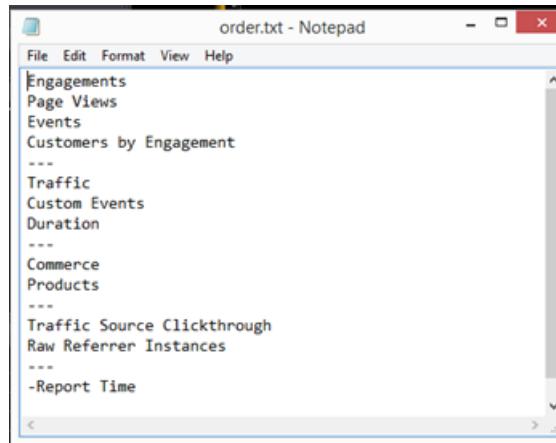
4. Change the *order.txt* in the metrics folder so that it reflects the new or modified metrics related to the Customer, Engagement and Event.

Adobe SC *order.txt* file.



```
order.txt - Notepad
File Edit Format View Help
Visits
Page Views
Hits
Visitors by Visit
---
Traffic
Custom Events
Duration
---
Commerce
Products
---
Traffic Source Clickthrough
Raw Referrer Instances
---
-Report Time
```

Order.txt file with new schema changes:



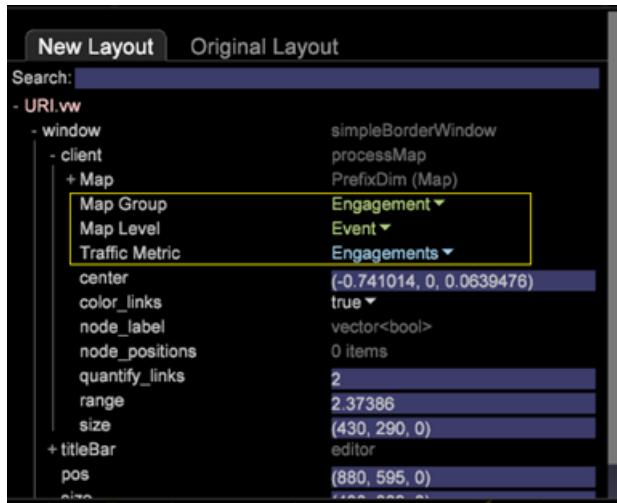
```
order.txt - Notepad
File Edit Format View Help
Engagements
Page Views
Events
Customers by Engagement
---
Traffic
Custom Events
Duration
---
Commerce
Products
---
Traffic Source Clickthrough
Raw Referrer Instances
---
-Report Time
```

5. All the configuration files (.vw) under Visualization folder should be changed to refer to new levels : Customer, Engagement and Event. For example: 2D process Map, 3D process Map etc.

Adobe SC default URI.vw for 2D process Map is defined at Hit level and Visit Group as shown below:



Changes to be made in URI.vw for new schema:



Derived Dimensions Setup

The different types of Derived (Client Side) Dimensions and how to setup those in Data Workbench.

Types of Derived Dimensions

Metric Dimensions

Metric Dimension allows you to group metric counts by a specific Level. It also allows you to group metric counts by a specific level. Once, a Metric Dimension is created, you can segment data based on the metric value.

Example 1: You are a Travel Company and you want to understand difference of behavioral activities on the website between your frequent flyers and customers who have booked flight less than 5 times—how would you do that?

All you have is count of Bookings as a metric, how will you segment customers based on a metric –here, booking –to understand their behavior on the website?

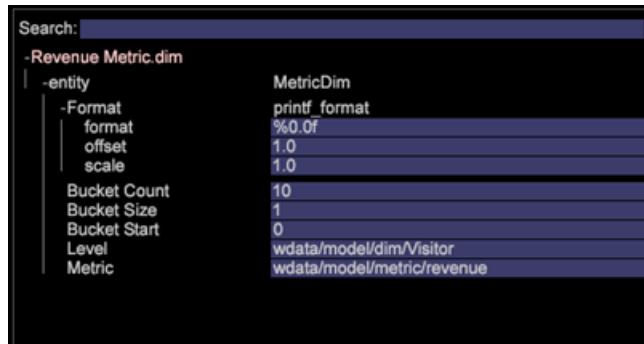
Example 2: You are a Financial Bank and you want to group your customers based on number of CDs they have invested in. You want to segment your customers in 3 Tiers. Tier 1 –Customers with 10+ CDs, Tier 2 –Customers with >5 and <10 CDs and Tier 3 –Customers with >0 and <5 CDs

The information you have is metric which gives you counts of CD investments –how will you create Tired Customer Segments for your analysis?

Creating Metric Dimension – via Workstation

Mark one of the OOB metric dimensions as local and Rename that dimension with a custom name / Make local copy of the RenameDim.example and rename it to the proper dimension name with .dim extension

Open the newly created dimension in the workstation to make changes. Change following parameters of the metric dimension based on the requirements:



Metric - Metric to be grouped

Level - Level at which metrics will be grouped

Bucket Start - Starting element of the Metric Dimension. Enter the same value in offset.

Bucket Size - Grouping size of metric. Enter the same value in scale

Bucket Count - Maximum number of elements to be displayed in the dimension

Save the newly created dimension on the server if you want to share it with others.

Prefix Dimensions

The main purpose of the Prefix dimension is to group elements of the original dimension and to provide user friendly names to the grouped elements.

For example, you own a retail site and your site has various site sections such as Women's Apparel, Men's Apparel, Toys and Games, Home Decor, etc and each of these site sections has several pages associated with it. You want to do path analysis and get insights about the traffic that goes from one site section to other and so on. If you use URI dimension, you will be required to pull each page of each of the site section in the Path Browser or Process Map and carry on the analysis.

The same analysis can be done easily if there is a Prefix dimension which has pages of a site section grouped together as a single element.

Creating Prefix Dimension:

Open a 2D process map from the Visualization menu.

Change following parameters of the prefix dimension based on the requirements.

Change Map Dimension – The Dimension that you want to use for 2D process map (Ex: SMS Typology)

Change Map Level Dimension – Level of the above mentioned dimension

Change Map Clip Dimension – The countable level at which you want to look at the data.

Change Map Metric – The metric that you want to look at.

Once the 2D Process map is set, open the dimension that you mentioned in Change Map Dimension parameter.

Select the elements you want to group. Use CTRL+ALT and drag & drop the elements on to process map.

Right click on the dot that appears and rename the group name. If you have selected 3 elements to group, the default name will be 3 Selected.

Right click on the outline of the visualization and save dimension from the menu that appears.

Rename Dimensions

Rename Dimensions are created off of a pre-existing dimension. The main purpose of the rename dimension is to provide user friendly names to the elements of the dimension. The out of the box Rename dimension is the Page dimension which is created off of the URI dimension. The URI dimension can be confusing for a person who does not know technical names of the pages and that is why the Page dimension allows you to rename elements of the URI dimension.

CREATING CUSTOM RENAME DIMENSIONS:

Elements of the Renamed dimension hold a One-to-One mapping with the original base dimension's elements. You can verify this by opening the .dim file of the Rename Dimension in the Workstation/Note pad. You will notice that each element of the original dimension has only one value (Rename String) against it in the file.

If you have fewer elements for the rename purpose; you can create a .dim file in the workstation and rename each individual elements by the steps explained below.

Steps to create a .dim file for a Rename Dimension- Using the Workstation

Use this option if the numbers of elements to be renamed are less.

1. Open a blank workspace and open the Dimensions Manager. Right Click>Admin>Profile>Profile Manager.
2. Expand the Dimensions Folder in the File Column.
3. Expand the Page Folder in the File Column and Right Click on the Page.dim file in the Second to Last column (This column usually represents the Profile Name) and click on the "Make Local" option.
4. Right Click on the Page.dim in the "User" column and click on the Copy option and Paste the copied .dim file in to the desired folder under the Dimensions directory.
5. Click on OK on the error message.
6. Now, you will notice that there are two Page.dim files under the Dimensions folder. One is the original file under Dimensions\Page directory and the second one is the one that you just copy pasted in step 4.
7. Right Click on the recently pasted Page.dim file under the User column and click on the blue/grey input box that says Page.dim. The input box will turn green with the cursor blinking, indicating that it can be modified. Type the name of the Rename dimension that you want to create.
8. You will notice that the Page.dim file in the File Column got changed to the new file name you gave in the step 7. Right click on the new.dim file in the User column (Last Column) and select Open>In Workstation.
9. Once the .dim file is opened in the workstation; click on the plus sign (+) next to the entity and expand it. Observe the value present against the "Parent" field, it reflects "URI" dimension. It shows "wdata/model/dim/URI" Click on the blue/grey input box to change the URI to the name of the dimension whose elements you want to rename.
10. Ensure the dimension that you want to rename does exist in the dataset. Dimension names are case sensitive so retain the case of the original dimension.
11. Observe the "modified" appearing next to the dimension name. This indicates that the original dimension has been modified. To sustain the changes made in the step 9; Right Click on new.dim (modified) and click on the "Save as" option.
12. Once the dimension is saved per step 10, the newly created rename dimension for the Campaigns is now available to you for the purpose of renaming. This is only available to you locally.

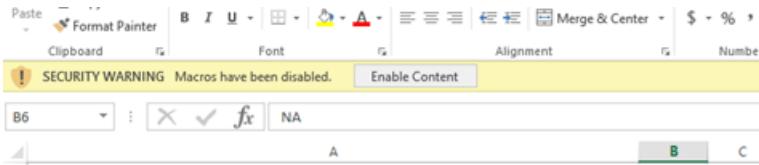
-
13. In order for others to see the dimension created by you, it has to be saved on to the profile. Right click on the .dim file of the new dimension in the "User" column (Last Column) and click on the "Save To>Profile Name" in which you want to save the dimension.
 14. After saving the file to the profile, all Workstation users who have access to this profile will be able to see the rename dimension for the Campaigns.

Prefix and Rename dim creator Tool

Adobe has an Excel tool to generate Prefix and Rename Dimensions.

Below are the steps to generate the Prefix/Rename dimensions using the tool:

1. Save the Excel tool *Adobe_DWB_Dimension_Generator.xlsxm* in a folder. Contact Adobe Customer Care to download the tool.
2. Open the tool and enable macros:



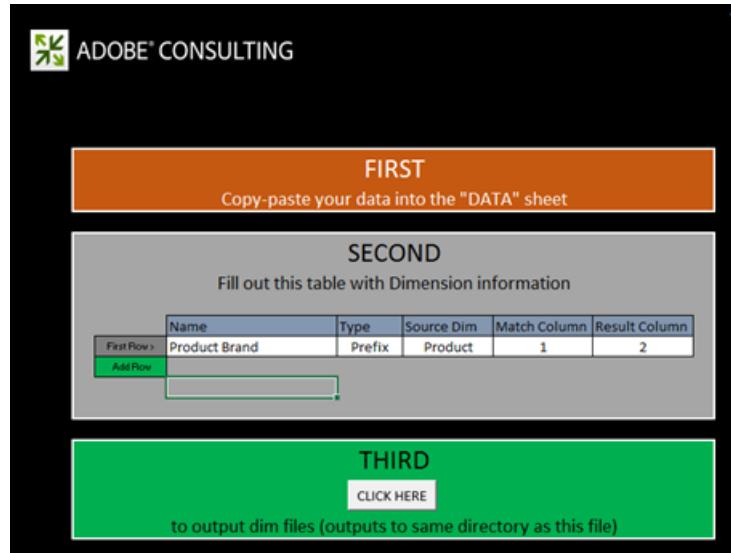
3. Fill the data sheet with the values to be used.

For example, we are creating Product Brand Prefix dimension based on Product Dimension. In the data Sheet, the following information is captured:

1	Air Optix for Astigmatism	Air Optix
2	ACUVUE® ADVANCE® PLUS 6 pack	NA
3	1-DAY ACUVUE® MOIST® for ASTIGMATISM 30 pack	NA
4	ACUVUE® OASYS® with HYDRACLEAR® PLUS 6 pack	NA
5	ACUVUE® OASYS® with HYDRACLEAR® PLUS 12 pack	NA
6	1-DAY ACUVUE® 30 pack	NA
7	Air Optix Aqua	Air Optix
8	FreshLook Colorblends Toric	FreshLook
9	Biofinity Toric	Biofinity
10	Biofinity	Biofinity
11	Vertex Toric (Encore Toric)	Vertex
12	FreshLook Dimensions	FreshLook
13	1-DAY ACUVUE® DEFINE™ 90 pack	NA
14	ACUVUE® OASYS® for ASTIGMATISM	NA
15	FreshLook ONE-DAY	FreshLook
16	ACUVUE® OASYS® with HYDRACLEAR® PLUS 24 pack	NA
17	FreshLook Colorblends	FreshLook
18	FreshLook Colors	FreshLook
19	1-DAY ACUVUE® DEFINE™ 30 pack	NA
20	Expressions	Expressions
21	1-DAY ACUVUE® MOIST® 90 pack	NA
22	Focus DAILIES 90 pack	Focus
23	Air Optix Colors	NA
24	ACUVUE® ADVANCE®	NA
25	SoftLens Toric (SoftLens For Astigmatism)	SoftLens
26	Biofinity Multifocal	Biofinity

Each product is assigned to a brand in the data sheet.

4. In the Configuration tab, fill the information related to the dimension to be created. For the sample data above below information is entered:



Name: Name for the Prefix/Rename dimension

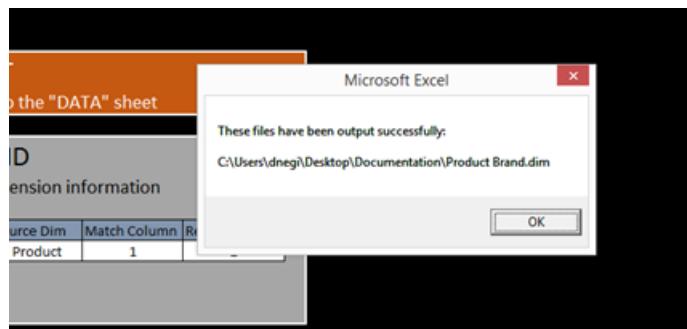
Type: Prefix/Rename

Source Dim: Original Dimension

Match Column: Column to be matched

Result column: Value to be used for new dimension.

5. Click the button titled *Click Here*.



6. The dim file will be generated in the same folder where the tool was saved.

```

1 entity = PrefixDim;
2   Base Dimension = ref: wdata/model/dim/Product
3   Element Names = vector: 35 items
4     0 = string: Acuvue
5     1 = string: Air Optix
6     2 = string: Aqualens
7     3 = string: Aqualite
8     4 = string: Avaira
9     5 = string: Biofinity
10    6 = string: Biomedics
11    7 = string: ClearLab
12    8 = string: ClearSight
13    9 = string: Encore
14    10 = string: Expressions
15    11 = string: Flextique
16    12 = string: Focus
17    13 = string: Frequency
18    14 = string: FreshLook
19    15 = string: Hydrogel
20    16 = string: Mediflex
21    17 = string: NA
22    18 = string: Optima
23    19 = string: Perspecta
24    20 = string: Polysoft
25    21 = string: Preference
26    22 = string: Proclear
27    23 = string: Proflex
28    24 = string: Prosite
29    25 = string: Provision
30    26 = string: PureVision
31    27 = string: Soflens
32    28 = string: Sofmed
33    29 = string: Softcon
34    30 = string: Sterling
35    31 = string: Target
36    32 = string: UltraFlex
37    33 = string: Versaflex
38    34 = string: Vertex
39   Element Prefixes = vector: 35 items
40     0 = vector: 16 items
41       0 = string: Acuvue
42       1 = string: Air Optix
43       2 = string: Aqualens
44       3 = string: Aqualite
45       4 = string: Avaira
46       5 = string: Biofinity
47       6 = string: Biomedics
48       7 = string: ClearLab
49       8 = string: ClearSight
50       9 = string: Encore
51       10 = string: Expressions
52       11 = string: Flextique
53       12 = string: Focus
54       13 = string: Frequency
55       14 = string: FreshLook
56       15 = string: Hydrogel
57       16 = string: Mediflex
58       17 = string: NA
59       18 = string: Optima
60       19 = string: Perspecta
61       20 = string: Polysoft
62       21 = string: Preference
63       22 = string: Proclear
64       23 = string: Proflex
65       24 = string: Prosite
66       25 = string: Provision
67       26 = string: PureVision
68       27 = string: Soflens
69       28 = string: Sofmed
70       29 = string: Softcon
71       30 = string: Sterling
72       31 = string: Target
73       32 = string: UltraFlex
74       33 = string: Versaflex
75       34 = string: Vertex
76   
```

Using the Profile Manager, save the dim file in the Dimension folder.

Shift Dimensions

Shift dimensions allows you to look at the Nth element of any dimension at the within any particular Countable Dimension.

They also give you the ability to look back at -Nth element of any dimension within any particular Countable Dimension

Example 1:

- The Nth page within a session –Next Page Dimension
- The Nth page for a visitor –Next Page for Visitor –across all Sessions
- The Nth call for a user

Why it is important to know Nth element of the countable dimension?

- You want to know the 5thPage Viewed in a Session.
- You want to do pathingon Campaigns to understand which was 2ndcampaign viewed after viewing " Free Checking Account" campaign?
- You want to understand which link visitor clicked before clicking "Chat with an Agent" link?



The Next URI is one of the OOB Shift dimensions which can be used as template. The example above is giving you 2nd(Offset = 1) element of the Campaign (Dim = Campaign) in the Engagement Event (Clip = Engagement Event)

Here offset 1 means look on the right –forward in the Event

Some other OOB Shift Dimensions

Next Page:

The next page viewed in a session after currently selected Page in the Page Dimension

Here offset is 1, Level is Page View, Dim is Page and Clip is Session

Previous Page:

The previous page viewed in a session before currently selected Page in the Page Dimension

Here offset is -1, Level is Page View, Dim is Page and Clip is Session

What will be the previous Campaign viewed before currently selected Campaign by a visitor?

Here offset is -1, Level is Campaign Response, Dim is Campaign Response Attribute Value and Clip is Visitor

Creating Shift Dimension –Via Workstation

- Mark one of the OOB shift dimension as local
- Rename that dimension with a custom name
- Open newly created dimension in the workstation to make changes
- Change following parameters of the metric dimension based on the requirements.
 - Level-Countable dimension
 - Offset-You want to look forward or backwards
 - Dim -Dimension whose elements you want to analyze
 - Clip-Countable in you want to view.

Save the newly created dimension on the server if you want to share it with others.

Last N Dimension

Last N Dimensions operates only on the Time Dimension and on the As of Time of the system. OOB time dimensions are Day, Week, Hour and Month. You can create, Last N dimension for each of these base time dimensions such as Last 10 Days, Last 72 hours, Last 8 Weeks, Last 6 Months, etc. Last N Dimension calculates Last N based on the current "Report Time Metric" or As of Time of the system.

Search:	
-Last10Days.dim	LastNDim
-entity	time_format
-Report Time Format	%x
ftime	true
gmt_to_local	
Count	10
Parent	wdata/model/dim/Day
Range Offset	0
Report Time	wdata/model/metric/Report Time
Start Week Day	Monday
show	false

Count – Total number of elements to be displayed in the dimension

Range offset – Offset value to denote the starting point (Day/Week) to calculate the last N Day/Week.

None.dim

None.dim is a Alias dimension. It is used to create alias from extended dimensions.

Example:

In the None.dim the entity is defined as "wdata/model/dim/Parent/+name" (it can be changed) which means create the dimension as per the name of the dimension file. So if we create a copy of None.dim file under Dimension folder (for this example, copying and renaming the None.dim file under Visitor Profile folder) and rename it to "Log Source ID.dim", a new derived dimension with Log Source ID will appear in the Menu under Visitor Profile as shown below:

Before changes:



After None.dim changes:



The entity can be changed to the extended dimension name, in this case another dimension with other name pointing to the same dimension as shown below:

In this example the "Source Name.dim" has the following content:

-Source Name.dim	wdata/model/dim/Parent/Log Source ID
entity	wdata/model/dim/Parent/Log Source ID
show	true

So another Dimension Source Name pointing to Log Source ID will appear.



Hiding Derived Dimensions

To hide the Derived Dimension, set the *Show* property to "false".

Search:	
-Last10Days.dim	LastNDim
-entity	time_format
-Report Time Format	%x
ftime	true
gmt_to_local	
Count	10
Parent	wdata/model/dim/Day
Range Offset	0
Report Time	wdata/model/metric/Report Time
Start Week Day	Monday
show	false

Data Workbench Glossary

Alert

In data workbench, a message or report that is automatically triggered when a metric reaches a defined threshold within a specified time frame. For example, one might create an alert that issues an e-mail if the number of page views during a 24-hour period falls above or below a specified quantity.

Analyst

An individual who performs analysis, defines reports, or otherwise uses data workbench.

Architect

An individual who determines how data will be captured, processed, and organized for analysis and reporting. This individual generally has significant expertise in configuring the Adobe® Platform for use by analysts.

Bar graph

A form of graph that uses rectangular bars of various sizes to show comparisons among two or more items.

Benchmark

A measurement or standard that serves as a point of reference by which others may be measured or judged. In data workbench, a benchmark illustrates the comparison between the value associated with a selection (a selected subset) and the value that would be shown if the selection were not made.

Cardinality

The number of elements in a set. The number of elements in an Adobe data dimension is referred to as the dimension's cardinality.

Checkpoint

The time at which a copy of an Adobe dataset was written for backup or recovery purposes. The term also refers to the set of data that is written during a checkpoint operation.

Chi-square

A statistical test to determine the probability that an observed deviation from the expected event or outcome occurred by chance. In data workbench, the closer Chi square is to 100%, the smaller the likelihood that the deviation occurred solely by chance.

Clickstream

An informal term that refers to the sequence of pages that a user requests while browsing a Web site. Clickstream information can help site owners understand how visitors use their site and which pages they request most frequently. Site owners can capture the clickstreams of visitors to their sites, but cannot capture clickstreams outside their site (except by using third-party cookies or page tags) because those requests are logged by outside Web servers.

Color Legend

A legend in data workbench that displays the color ramp for a selected metric. Color legends enable the colorcoding of visualizations by various metrics. Color-coding visualizations makes it easier to spot anomalies, exceptions, and trends.

Color Ramp

In a color legend, the range of colors used to represent the range of possible values for a metric. When a color legend is applied to a visualization, graphical elements in the visualization (for example, bars in a bar graph) are color-coded according to the values represented by the color ramp.

Common Key

A common variable that forms a relationship between the rows in two tabular sets of data. For example, a product ID might act as the common key between a table of query-string variables and a table of product data from an inventory-management system.

Confidence Interval

A range of values that has a specified probability of containing the rate or trend. The 80% (p-value = .20), 95% (p-value = .05) and 99% (p-value = .01) confidence intervals are the most commonly used intervals. (Source: <http://www.nci.nih.gov/statistics/glossary>)

Confidence Level

The likelihood that the sampling error in a survey result will fall within a specified range, usually expressed in terms of standard errors (for example, 1 standard error equals 68% likelihood, 2 standard errors equals 95.4% likelihood). (Source: <http://www.magazine.org/research/3410.cfm>)

Conversion Map

In data workbench, a type of visualization in which elements are plotted on the x-axis according to the value of the Conversion metric.

Conversion Rate

In data workbench, the percentage of sessions during which a value event occurred. Conversion rate is calculated by dividing the number of sessions in which a value event occurred by the total number of sessions.

Correlation

A numeric measure of the strength of a linear relationship between two random variables. Variables that tend to move up or down together are positively correlated, while variables that tend to move in opposite directions are negatively correlated. See also Correlation Coefficient.

Correlation Coefficient

The numeric value that signifies the strength of a linear relationship between two random variables. See also Correlation.

Countable Dimension

A dimension in which the number of elements the dimension contains can be discretely counted. Countable dimensions can have child dimensions of the following types: Countable, Numeric, Simple, Many-to-Many, and Denormal.

CrossRows Transformation

In data workbench server, a data transformation that enables data from multiple event records for a Visitor (made at different times) to be incorporated in a calculation.

Crosstab

In data workbench, a tabular visualization that displays the metrics associated with the intersections of two dimensions.

Cube

A multi-dimensional data structure or a group of data cells arranged by the dimensions of the data. For example, a spreadsheet exemplifies a two-dimensional array with data arranged in rows and columns. In a spreadsheet, each row or column is a dimension. A three-dimensional array can be visualized as a cube with each dimension forming a side of the cube, including any slice parallel with that side. Higher dimensional arrays have no physical

metaphor, but they can be used to organize data in ways that users think about their enterprise. Also known as a hypercube, multi-dimensional array, or multi-dimensional database.

Dashboard

A workspace that is created for viewing rather than interaction. Dashboards provide “at-a-glance” status through the display of key performance indicators that are appropriate for a particular manager or operator who is overseeing one or more business objectives. See also Workspace.

Data Mining

The unguided or interactive application of a collection of mathematical procedures to historical business data in an effort to find insights in the form of correlations and other statistical relationships.

Data Subset

A feature that enables users to easily select a subset of a dataset for use in online or offline analysis. (A data subset is a portion of the dataset based on a filter.)

Data Warehouse

A database designed to support decision making in organizations. A data warehouse generally contains large amounts of subject-oriented, time-variant, nonvolatile data that is structured for rapid, online queries and managerial summaries.

Data Processing Unit

A type of data workbench server that processes, stores, and serves data from an Adobe dataset. A DPU can optionally store the VSL log files that contain the source data from which the dataset is constructed or can receive that data from an data workbench server File Server Unit (FSU). A DPU is the type of data workbench server with which data workbench and Report® clients interact directly.

Dataset

The data loaded and processed by data workbench server. The dataset represents the data that can be transmitted to data workbench or Report for analysis, reporting, and alerting purposes. Physically, the dataset resides in the temp.db file. Each data workbench server computer (or data workbench server cluster) maintains one dataset.

Dataset Data

The data created and stored in an data workbench server dataset. It includes the event data and integration data that is admitted to or created in the dataset. It also includes any information derived from such data as determined by the configuration files that define that dataset. Dataset data can be re-created by reprocessing the event and integration data using the same or different configuration files. (Configuration files are system files that are managed as a part of an Adobe profile.)

Dataset Records

Those event data records admitted to an Adobe dataset after all filtering and other processing has been performed. Also called processed log entries.

Dataset Schema

A visualization in data workbench that displays the schema of the dataset that backs the currently selected profile.

Dataset Storage Space

The amount of data (in bytes) entered or created in an Adobe dataset. This data is stored by the dataset in a file called Temp.db on an data workbench server computer. The data in Temp.db is transient and can be recreated by reprocessing the source data (i.e., event data and integration data) with the appropriate configuration files.

Decoder

The component in data workbench server that reads event data from various sources and generates data that is used to produce the dataset. Output from a decoder can be used as input to any of the “log processing” capabilities

in data workbench server. Decoder types include the Sensor decoder (for loading data from various versions of Sensor), the Regular Expression decoder (for loading data from delimited flat files), and the ODBC decoder (for loading data from ODBC data sources).

Dimension

A set of elements, all of which are of a similar type in the user’s perception. The elements define a set of categories into which data can be grouped. For instance, the elements Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, and Sunday make up a “Weekday” dimension.

Dimension Element

An individual category within a dimension. For example, a “Weekday” dimension would contain the individual elements Monday, Tuesday, Wednesday, Thursday, Friday, Saturday, and Sunday.

Dimension Legend

In data workbench, the legend that lists the dimensions that have been defined in (or derived from) the dataset. When a selection is made in a visualization, the Dimension legend identifies those dimensions whose values differ from the benchmark by a statistically significant amount.

Drill Up-Down

A specific analytical technique whereby the user navigates among levels of data ranging from the most summarized (up) to the most detailed (down). For example, when viewing sales data for North America, a drill-down operation in the Region dimension might display Canada, the United States, and Mexico. A further drill-down on Canada might display Toronto, Vancouver, Montreal, and so forth.

Event Data

The data collected by Sensors or other means (for example, a web server log file), which constitutes the primary input to data workbench server. Each event data record represents a transaction record or a single instance of an event.

Extended Dimension

A dimension that is based on extended data. Extended data is any data beyond what is minimally required to form a valid event data record. Extended data can be added to an event data record when the original event is captured or it can be incorporated from other sources and added to the event data record as integration data. Any dimension that is based on this additional data is considered to be “extended.”

File Server Unit (FSU)

A type of data workbench server whose function is solely to receive event data from one or more Sensors or repeater servers and provide data to one or more data workbench server Data Processing Units (DPUs) for their use in constructing Adobe datasets. FSUs optimize the transfer of event data to the DPUs and are significantly faster than ordinary file servers. The use of an FSU reduces hardware costs by enabling log data to be stored on lower cost storage hardware and reduces administrative complexity by allowing multiple Sensors to point to a single data workbench server.

Incremental Query Evaluation

The patented process by which data workbench server provides immediate query results to a user based on a projected random sample of the full population. Under this process, the server incrementally refines the accuracy of the query by considering more data until all data has been considered and an exact count has been obtained.

Integration Data

Integration data is external data from corporate databases or look-up files that you can combine with event data to create the dataset. In general, you use integration data to augment the event data acquired by Sensor. (Conceptually, you can think of using integration data to populate event data records with additional columns of information.)

Legend

A window in data workbench that provides explanatory details about the visualizations displayed in the workspace. Types of legends include Color legends, Dimension legends, and Metric legends. Like any window in data workbench, Legend windows can be generated and distributed by Report.

Line Graph

A type of visualization in data workbench that plots metrics for a specified dimension as successive points on the x-axis of a graph and then connects the points with lines. A line graph is a particularly effective way to visualize a metric over a time-based dimension.

Log Entry Condition

A condition that determines whether an event data record (a log entry) will be included in the dataset. For example, a log entry condition might specify that only event data records associated with a particular Web site are to be admitted to the dataset. Log entry conditions are specified in the log processing configuration file on the data workbench server.

Many-to-Many Dimension

In the Adobe Platform, a dimension that has a many-to-many relationship with a parent Countable dimension. A many-to-many dimension represents a set of values for each element of its parent dimension. For example, in Site, the Search Phrase dimension has a many-to-many relationship to its parent, the Session dimension (that is, a Session can have any number of Search Phrases, and a Search Phrase can have any number of Sessions.)

Masking

A feature in data workbench that enables analysts to temporarily hide elements that they do not want to include in an analysis.

Mean

The arithmetic average of a set of numbers. The sum of the data divided by the sample size.

Median

A number that separates the highest half of a sample, a population, or a probability distribution from the lowest half. Half of the population will have values less than or equal to the median, and half of the population will have values equal to or greater than the median.

Metric

In Adobe, a named formula that describes how to calculate a quantitative value from the data in the dataset. In Site, for example, the “Sessions per Visitor” metric represents a formula that divides the count of Sessions by the count of Visitors.

Metric Legend

A window in data workbench that displays the metrics defined by the active profile. A Metric legend displays the value of each metric as calculated from the dataset or the current selection (if a selection is active in the workspace, the Metric legend displays the values for the selected subset instead of the entire dataset.) Like any window in data workbench, Metric legends can be generated and distributed by Report.

Metric Worksheet

A window in data workbench that enables analysts to define their own metrics for a dataset. A Metric worksheet is similar to a spreadsheet. Using data workbench formula syntax, analysts can enter expressions describing the quantitative values that they want to derive from the dataset. For example, an analyst might define a metric that displays the percentage of Visitors who viewed a page from a particular domain. Like regular metrics, formulas in a Metric worksheet operate on the selected subset when a selection is active in the workspace. Like any window in data workbench, Metric worksheets can be generated and distributed by Report.

New Visitor Condition

The condition that determines whether a new tracking ID is created when data workbench server is presented with an event data record.

Node

A grouping of one or more discrete items into a single logical entity. In a configuration file (.cfg), a node is an item containing related parameters. See also Parameter and Vector. In data workbench, a node on a Process Map represents a single page or defined group of pages.

Numeric Dimension

In the Adobe Platform, a dimension that has ordered, numerical values and has a one-to-many relationship with a parent Countable dimension. Generally, a Numeric dimension represents a numeric property of the elements of the parent dimension. Numeric dimensions are often used to define “sum” metrics.

One-to-Many Relationship

A relationship between two data dimensions in which a single element in one dimension is (or can be) related to one or more elements in the other dimension.

Ordinal

Being or denoting a numerical order in a series. (Source: <http://wordnet.princeton.edu/perl/webwn?s=ordinal>) If a dimension is sorted ordinally in data workbench, the elements of the dimension are displayed in the order in which they are represented internally.

Outliers

In a set of data, a value so far removed from other values in the distribution that its presence cannot be attributed to the random combination of chance causes.

Page Overlay

A type of visualization in data workbench that color-codes the links on an image of a Web page according to a specified metric. You can use a Page Overlay visualization to quickly identify which links on a page attract visitors attention (and take them to other pages on your site) and which do not. You can also use it understand the “value” (as measured by value events) that the various links on a page generate for your site.

Page View Condition

An option in an data workbench server transformation that enables event records to be included or excluded as page views based on their content type or actual content. You might use this option, for example, to exclude event records pertaining to unsuccessful HTTP requests (for example, ones that produced a 404 status code) or requests that return certain content types (for example, image requests). If an HTTP event record is not filtered out by the Page View Condition option, that event record will represent a page view in the dataset.

Path Browser Visualization

A type of visualization in data workbench that enables an analyst to interactively explore a sequences of events (such as Page Views) in a Visitor Session or across Visitor Sessions.

Primary Server

In a data workbench server cluster, the data workbench server that brokers communications between clients (such as data workbench and Report) and the other servers in the cluster. The primary server also functions as the administrative focal point for the cluster. Using the data workbench server’s profile synchronization capabilities, changes that an administrator makes to the primary server are automatically propagated to the other servers in the cluster. A primary server is an data workbench server DPU.

Process Map

A type of visualization in data workbench that enables an analyst to understand the flow of traffic among pages or nodes on a Web site. A process map depicts information about specific pages (such as the number of sessions during which the pages were viewed) and also illustrates traffic volume between pages or nodes.

Processing Server

In a data workbench server cluster, the data workbench servers that are controlled by the configuration information on the primary server. A processing server processes event data into a dataset and responds to queries from clients such as data workbench and Report. When responding to a query, a processing server divides (“partitions”) responsibility for fulfilling the query among the servers in the cluster. When the other servers complete their portions of the query, the processing server combines (“departitions”) their results and returns the combined result to the client.

Profile

A set of configuration files that contains the rules for creating a dataset for a specific analysis purpose. A profile also defines articles such as metrics, derivative dimensions, workspaces, reports, visualizations, and legends that enable analysts to interact with the dataset and obtain information from it. A profile can be structured generally for a Web site (as in a profile for www.mysite.com) or it can be tailored for a particular type of user (such as the “Marketing” user profile or the “Finance” user profile).

Profile Manager

An interactive administrative tool in data workbench that enables an administrator or other user to manage the configuration files associated with a profile.

Regular Expression

A formula that describes or matches a set of strings according to certain syntax rules. Regular expressions (often abbreviated as regexp, regex, or regxp) are used to search and manipulate bodies of text based on certain patterns. Regular expression notation originated in early Unix editors and gained widespread use in other Unix utilities such as vi and Perl. Today, regular expressions are supported by many text editors, scripting languages, and other text-manipulation tools. Data workbench server includes a regular expression engine.

Retention Map

A type of visualization in data workbench that plots elements on the x-axis according to the value of the Retention metric.

Scatter Plot

A type of visualization in data workbench that represents bi-variate data as points on a graph. A scatter plot has two pieces of data for each element being graphed. For example, a scatter plot of the Conversion Rate and Visitors metrics for a set of 10 pages would result in 10 separate points in the plot.

Selection

A feature in data workbench that enables an analyst to restrict the set of data retrieved and displayed for reporting or analysis. A selection can be made interactively within data workbench by clicking elements displayed in one or more dimensions (left-click selects an element, right-click deselects an element). Selections can also be made by defining filters that select certain elements in specified dimensions.

Simple Dimension

In the Adobe Platform, a dimension that has a one-to-many relationship with a parent countable dimension. For example, Visitor Referrer is a simple dimension whose parent is the countable dimension, Visitor. A Visitor has only one Visitor Referrer. However, a Visitor Referrer can have multiple Visitors (that is, one Visitor Referrer can be related to many Visitors).

Smoothing

A mathematical inference of a curve across multiple points in a line graph, which is used to illustrate a more meaningful trend line across relatively sparse data points.

Source

On data workbench server, a resource containing event data that can be used to create a dataset. Adobe decoders look to sources for event data to decode for use by data workbench server.

Source Data

On data workbench server, data that is the input to one of its decoders. Source data can be input to the Sensor decoder (which loads source data from various versions of Sensor), the Regular Expression decoder (which loads source data from delimited flat files), and the ODBC Decoder (which loads source data from ODBC data sources).

Table Graph

A type of visualization in data workbench that displays data in tabular format. The metric values in a table graph can be expressed numerically or represented in bar format.

Time Series

A graph that shows how a given property or value changes over time.

Tracking ID

An identifier that uniquely distinguishes the primary entities being analyzed in an Adobe dataset. A tracking ID can be constructed from various sources including a unique ID from a Web client cookie, an IP number and user-agent hash, or an x.509 name. Although these are some common sources for a tracking ID, any value that is capable of uniquely identifying the entities populating the Visitor dimension in the dataset can be used.

Transformation

A method of changing the value of a variable using some mathematical operation. On the data workbench server, for example, analysts can use the split transformation to break the name-value pairs in a query string into individual variables.

Two-Dimensional Bar Graph

A type of visualization in data workbench that simultaneously displays two dimensions and up to two metrics in a three-dimensional graphical view.

Value Legend

A window in data workbench that enables an analyst to associate a monetary value with a selected event and view the results when that monetary value is summed for all instances of the selected event in the dataset.

Visitor

The dimension in the dataset that identifies the entity that generated the event. Each member of the Visitor dimension is identified by a unique tracking ID. In Site, for example, the tracking ID is typically derived from a unique ID within the client's cookie. In Call, the tracking ID might be the caller's a telephone number.

Visitor Referrer

The first HTTP referrer for a Visitor within the time span of an Adobe dataset.

Visitor Splitting

A feature in data workbench server that allows visitors with large amounts of event data to be split between two tracking IDs. Visitor splitting is used to prevent event data from being filtered from the dataset when a visitor exceeds the maximum configured number of events per visitor (a parameter that is set to ensure correct system performance). Although visitor splitting artificially increases the number of visitors in the dataset, it does not inflate the total number of event records, which ensures that the total number of countable events (for example, page views, bookings) remains accurate.

VSL

Log file. The type of file in which data workbench server store event data that they receive from Sensor. VSL files are compressed, but can be output in uncompressed form using data transformation functionality. VSL is the file extension for a log file.

What-If Analysis

A type of analysis used to understand the consequences of changing certain variables in a data model by observing the effect that changes to those variables have on other data in the model.

Workspace

In data workbench, a workspace is a container for a particular analysis and visualization task. A workspace can contain multiple visualizations), all of which operate against a common set of data (that is, all visualizations in the workspace render the same set of query results). When an analyst performs a subsetting or filtering operation in one visualization, the selected subset is reflected throughout the entire workspace.

Worktop

The “home” area within the data workbench user interface that enables you to organize and access all of your workspaces and reports. It also enables you to create and save new and updated workspaces and reports to the data workbench server so that others using the same profile can access them as well.

Third-Party Licenses

Apache Xerces: This product includes software developed by the Apache Software Foundation (<http://www.apache.org/>). Copyright (c) 2002 The Apache Software Foundation, all rights reserved. Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met: 1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer. 2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution. 3. The end-user documentation included with the redistribution, if any, must include the following acknowledgment: "This product includes software developed by the Apache Software Foundation (<http://www.apache.org/>).” Alternately, this acknowledgment may appear in the software itself, if and wherever such third-party acknowledgments normally appear. 4. The names “Xerces” and “Apache Software Foundation” must not be used to endorse or promote products derived from this software without prior written permission. For written permission, please contact apache @apache.org. 5. Products derived from this software may not be called “Apache”, nor may “Apache” appear in their name, without prior written permission of the Apache Software Foundation. THIS SOFTWARE IS PROVIDED ``AS IS'' AND ANY EXPRESSED OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE APACHE SOFTWARE FOUNDATION OR ITS CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

OpenSSL Toolkit: This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit (<http://www.openssl.org/>). Copyright (c) 1998-2002 The OpenSSL Project. All rights reserved. Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met: 1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer. 2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution. 3. All advertising materials mentioning features or use of this software must display the following acknowledgment: "This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit (<http://www.openssl.org/>).” 4. The names “OpenSSL Toolkit” and “OpenSSL Project” must not be used to endorse or promote products derived from this software without prior written permission. For written permission, please contact openssl-core@openssl.org 5. Products derived from this software may not be called “OpenSSL” nor may “OpenSSL” appear in their names without prior written permission of the OpenSSL Project. 6. Redistributions of any form whatsoever must retain the following acknowledgment: "This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit (<http://www.openssl.org/>).” THIS SOFTWARE IS PROVIDED BY THE OpenSSL PROJECT ``AS IS'' AND ANY EXPRESSED OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE OpenSSL PROJECT OR ITS CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

This product includes cryptographic software written by Eric Young (eay@cryptsoft.com). Copyright (C) 1995-1998 Eric Young (eay@cryptsoft.com) All rights reserved. This package is an SSL implementation written by Eric Young (eay@cryptsoft.com). The implementation was written so as to conform with Netscape's SSL. This library is free for commercial and non-commercial use as long as the following conditions are adhered to. The following conditions apply to all code found in this distribution, be it the RC4, RSA, Lhash, DES, etc., code; not just the SSL code. The SSL documentation included with this distribution is covered by the same copyright terms except that the holder is Tim Hudson (tjh@cryptsoft.com). Copyright remains Eric Young's, and as such any Copyright notices in the code are not to be removed. If this package is used in a product, Eric Young should be given attribution as the author of the parts of the library used. This can be in the form of a textual message at program startup or in documentation (online or textual) provided with the package. Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met: 1. Redistributions of source code must retain the copyright notice, this list of conditions and the following disclaimer. 2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution. 3. All advertising materials mentioning features or use of this software must display the following acknowledgement: "This product includes cryptographic software written by Eric Young (eay@cryptsoft.com)" The word 'cryptographic' can be left out if the routines from the library being used are not cryptographic related. 4. This product includes software written by Tim Hudson (tjh@cryptsoft.com). THIS SOFTWARE IS PROVIDED BY ERIC YOUNG ``AS IS'' AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE AUTHOR OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Liblzf: This software uses, in a modified form, the LZF compression library. Copyright (c) 2000 Marc Alexander Lehmann pcg@gof.com. Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met: 1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer. 2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution. 3. The name of the author may not be used to endorse or promote products derived from this software without specific prior written permission. THIS SOFTWARE IS PROVIDED BY THE AUTHOR ``AS IS'' AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE AUTHOR BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

PCRE: Copyright (c) 1997-2001 University of Cambridge. Permission is granted to anyone to use this software for any purpose on any computer system, and to redistribute it freely, subject to the following restrictions: 1. This software is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. 2. The origin of this software must not be misrepresented, either by explicit claim or by omission. In practice, this means that if you use PCRE in software which you distribute to others, commercially or otherwise, you must put a sentence like this "Regular expression support is provided by the PCRE library package, which is open source software, written by Philip Hazel, and copyright by the University of Cambridge, England." Somewhere reasonably visible in your documentation and in any relevant files or online help data or similar. A reference to the ftp site for the source, that is, to <ftp://ftp.csx.cam.ac.uk/pub/software/programming/pcre/> should also be given in the

documentation. 3. Altered versions must be plainly marked as such, and must not be misrepresented as being the original software.

Truerand: Truerand is based on code by D. P. Mitchell. Truerand was modified by Matt Blaze 7/95, 11/96. Truerand is Version 2.1. This is completely unsupported software. The authors of this software are Don Mitchell and Matt Blaze. Copyright (c) 1995, 1996 by AT&T. Permission to use, copy, and modify this software without fee is hereby granted, provided that this entire notice is included in all copies of any software which is or includes a copy or modification of this software and in all copies of the supporting documentation for such software. This software may be subject to United States export controls. THIS SOFTWARE IS BEING PROVIDED "AS IS", WITHOUT ANY EXPRESS OR IMPLIED WARRANTY. IN PARTICULAR, NEITHER THE AUTHORS NOR AT&T MAKE ANY REPRESENTATION OR WARRANTY OF ANY KIND CONCERNING THE MERCHANTABILITY OF THIS SOFTWARE OR ITS FITNESS FOR ANY PARTICULAR PURPOSE.

libJPEG: The authors make NO WARRANTY or representation, either express or implied, with respect to this software, its quality, accuracy, merchantability, or fitness for a particular purpose. This software is provided "AS IS", and you, its user, assume the entire risk as to its quality and accuracy. This software is copyright (C) 1991-1998, Thomas G. Lane. All Rights Reserved except as specified below. Permission is hereby granted to use, copy, modify, and distribute this software (or portions thereof) for any purpose, without fee, subject to these conditions: (1) If any part of the source code for this software is distributed, then this README file must be included, with this copyright and no-warranty notice unaltered; and any additions, deletions, or changes to the original files must be clearly indicated in accompanying documentation. (2) If only executable code is distributed, then the accompanying documentation must state that "this software is based in part on the work of the Independent JPEG Group". (3) Permission for use of this software is granted only if the user accepts full responsibility for any undesirable consequences; the authors accept NO LIABILITY for damages of any kind. These conditions apply to any software derived from or based on the IJG code, not just to the unmodified library. If you use our work, you ought to acknowledge us. Permission is NOT granted for the use of any IJG author's name or company name in advertising or publicity relating to this software or products derived from it. This software may be referred to only as "the Independent JPEG Group's software". We specifically permit and encourage the use of this software as the basis of commercial products, provided that all warranty or liability claims are assumed by the product vendor.

libtiff 3.5.7: Copyright (c) 1988-1997 Sam Leffler. Copyright (c) 1991-1997 Silicon Graphics, Inc. Permission to use, copy, modify, distribute, and sell this software and its documentation for any purpose is hereby granted without fee, provided that: (i) the above copyright notices and this permission notice appear in all copies of the software and related documentation, and (ii) the names of Sam Leffler and Silicon Graphics may not be used in any advertising or publicity relating to the software without the specific, prior written permission of Sam Leffler and Silicon Graphics. THE SOFTWARE IS PROVIDED "AS-IS" AND WITHOUT WARRANTY OF ANY KIND, EXPRESS, IMPLIED OR OTHERWISE, INCLUDING WITHOUT LIMITATION, ANY WARRANTY OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. IN NO EVENT SHALL SAM LEFFLER OR SILICON GRAPHICS BE LIABLE FOR ANY SPECIAL, INCIDENTAL, INDIRECT OR CONSEQUENTIAL DAMAGES OF ANY KIND, OR ANY DAMAGES WHATSOEVER RESULTING FROM LOSS OF USE, DATA OR PROFITS, WHETHER OR NOT ADVISED OF THE POSSIBILITY OF DAMAGE, AND ON ANY THEORY OF LIABILITY, ARISING OUT OF OR IN CONNECTION WITH THE USE OR PERFORMANCE OF THIS SOFTWARE.

Bzip: This program, "bzip2" and associated library "libbzip2", are copyright (C) 1996-2002 Julian R Seward. All rights reserved. Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met: 1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer. 2. The origin of this software must not be misrepresented; you must not claim that you wrote the original software. If you use this software in a product, an acknowledgment in the product documentation would be appreciated but is not required. 3. Altered source versions must be plainly marked as such, and must not be misrepresented as being the original software. 4. The name of the author may not be used to endorse or promote products derived from this software without specific prior written permission. THIS SOFTWARE IS PROVIDED BY THE AUTHOR ``AS IS'' AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO

EVENT SHALL THE AUTHOR BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE. Julian Seward, Cambridge, UK. jseward@acm.org. bzip2/libbzip2 version 1.0.2 of 30 December 2001

Libpng: COPYRIGHT NOTICE, DISCLAIMER, and LICENSE: Libpng versions 1.0.7, July 1, 2000, through 1.2.3, May 21, 2002, are Copyright (c) 2000-2002 Glenn Randers-Pehrson, and are distributed according to the same disclaimer and license as libpng-1.0.6 with the following individuals added to the list of Contributing Authors, Simon-Pierre Cadieux, Eric S. Raymond and Gilles Vollant. There is no warranty against interference with your enjoyment of the library or against infringement. There is no warranty that our efforts or the library will fulfill any of your particular purposes or needs. This library is provided with all faults, and the entire risk of satisfactory quality, performance, accuracy, and effort is with the user. Libpng versions 0.97, January 1998, through 1.0.6, March 20, 2000, are Copyright (c) 1998, 1999, 2000 Glenn Randers-Pehrson. Libpng is distributed according to the same disclaimer and license as libpng-0.96, with the following individuals added to the list of Contributing Authors: Tom Lane, Glenn Randers-Pehrson, and Willem van Schaik. Libpng versions 0.89, June 1996, through 0.96, May 1997, are Copyright (c) 1996, 1997 Andreas Dilger Distributed according to the same disclaimer and license as libpng-0.88, with the following individuals added to the list of Contributing Authors: John Bowler, Kevin Bracey, Sam Bushell, Magnus Holmgren, Greg Roelofs, and Tom Tanner. Libpng versions 0.5, May 1995, through 0.88, January 1996, are Copyright (c) 1995, 1996 Guy Eric Schalnat, Group 42, Inc. For the purposes of this copyright and license, "Contributing Authors" is defined as the following set of individuals: Andreas Dilger, Dave Martindale, Guy Eric Schalnat, Paul Schmidt and Tim Wegner. The PNG Reference Library is supplied "AS IS". The Contributing Authors and Group 42, Inc. disclaim all warranties, expressed or implied, including, without limitation, the warranties of merchantability and of fitness for any purpose. The Contributing Authors and Group 42, Inc. assume no liability for direct, indirect, incidental, special, exemplary, or consequential damages, which may result from the use of the PNG Reference Library, even if advised of the possibility of such damage. Permission is hereby granted to use, copy, modify, and distribute this source code or portions hereof, for any purpose, without fee, subject to the following restrictions: 1. The origin of this source code must not be misrepresented. 2. Altered versions must be plainly marked as such and must not be misrepresented as being the original source. 3. This Copyright notice may not be removed or altered from any source or altered source distribution. The Contributing Authors and Group 42, Inc. specifically permit, without fee, and encourage the use of this source code as a component to supporting the PNG file format in commercial products. If you use this source code in a product, acknowledgment is not required but would be appreciated.

STLPort: Boris Fomitchev grants End User a non-exclusive, non-transferable, royalty-free license to use STLport and its documentation without fee. By downloading, using, or copying STLport or any portion thereof, End User agrees to abide by the intellectual property laws and all other applicable laws of the United States of America, and to all of the terms and conditions of this Agreement. End User shall maintain the following copyright and permission notices on STLport sources and its documentation unchanged: Copyright 1999, 2000 Boris Fomitchev. This material is provided "as is", with absolutely no warranty expressed or implied. Any use is at your own risk. Permission to use or copy this software for any purpose is hereby granted without fee provided the above notices are retained on all copies. Permission to modify the code and to distribute modified code is granted, provided the above notices are retained, and a notice that the code was modified is included with the above copyright notice. The End User may distribute binaries compiled with STLport (whether original or modified) without any royalties or restrictions. The End User may distribute original or modified STLport sources, provided that: The conditions indicated in the above permission notice are met; The following copyright notices are retained when present, and conditions provided in accompanying permission notices are met: Copyright 1994 Hewlett-Packard Company; Copyright 1996,97 Silicon Graphics Computer Systems, Inc.; Copyright 1997 Moscow Center for SPARC Technology; Permission to use, copy, modify, distribute and sell this software and its documentation for any purpose is hereby granted without fee, provided that the above copyright notice appear in all copies and that both that copyright notice and this permission notice appear in supporting documentation. Hewlett-Packard Company makes no representations about the suitability of this software for any purpose. It is provided "as is"

without express or implied warranty. Permission to use, copy, modify, distribute and sell this software and its documentation for any purpose is hereby granted without fee, provided that the above copyright notice appear in all copies and that both that copyright notice and this permission notice appear in supporting documentation. Silicon Graphics makes no representations about the suitability of this software for any purpose. It is provided "as is" without express or implied warranty. Permission to use, copy, modify, distribute and sell this software and its documentation for any purpose is hereby granted without fee, provided that the above copyright notice appear in all copies and that both that copyright notice and this permission notice appear in supporting documentation. Moscow Center for SPARC Technology makes no representations about the suitability of this software for any purpose. It is provided "as is" without express or implied warranty.

Zlib: (c) 1995-2002 Jean-loup Gailly and Mark Adler. This software is provided 'as-is', without any express or implied warranty. In no event will the authors be held liable for any damages arising from the use of this software. Permission is granted to anyone to use this software for any purpose, including commercial applications, and to alter it and redistribute it freely, subject to the following restrictions: 1. The origin of this software must not be misrepresented; you must not claim that you wrote the original software. If you use this software in a product, an acknowledgment in the product documentation would be appreciated but is not required. 2. Altered source versions must be plainly marked as such, and must not be misrepresented as being the original software. 3. This notice may not be removed or altered from any source distribution.

GDAL Library: (c) 2000 Frank Warmerdam. Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions: The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software. THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

NetRatings, Inc.: Adobe Systems, Inc., grants End User a limited, non-exclusive sublicense under the "NetRatings Patents" solely for the use by End User of the licensed Software Applications. End User expressly acknowledges and agrees that the sublicense granted under this Agreement does not permit the End User to aggregate "Video Stream Data" (as defined below) from multiple sources for resale or distribution to third parties. "Video Stream Data" is defined as audience measurement data relating to the exposure to, or viewership of, video streams (including, without limitation, television and/or IPTV content, sports content, news content, movies and commercials). For purposes of this license, "NetRatings Patents" shall mean the following Patent/Publication Numbers: US Patents - 5,675,510; 6,115,680; 6,108,637; US20040078292; 5,796,952; 6,138,155; 6,643,696; 6,763,386; US 20050114511; and US 20040221033; Australia: AU701813; AU735285; and AU727170; Brazil - BR9609217 and BR9808033; Canada - CA2223919; CA2246746; and CA2284530; EP - EP0843946; EP0870234; EP1130526; EP1168196; and EP1010116; Japan - JP03317705 and JP2000514942; Mexico - MX193614; Norway - NO09705728; WO - WO9641495; WO9810349; and WO9843380; China - CN1174316; CN1547123; CN1547124; CN1209891; and CN1251669; Denmark - DK870234; Germany - DE69720186; Spain ES2195170; Israel - IL131871; Korea - KR341110; and New Zealand - NZ337756.

HTML Snapshot: Effective this 14 day of January, 2005, Guangming Software, an entity formed and located in the People's Republic of China ("Licensor") grants the following rights and licenses to Adobe Systems, Inc. f/k/a Adobe, Inc. ("Adobe"): Licensor, as the sole owner of the rights and title to the source code of "HTML Snapshot," hereby grants Adobe a perpetual royalty-free, non-exclusive license to use the source code and object code of Licensor's "HTML Snapshot" software product. Licensor grants Adobe the right to use the licensed source code of "HTML Snapshot": (1) to create derivative works thereof and therefrom and to modify it as needed; (2) in combination with Adobe' other proprietary source code; and (3) to redistribute, sub-license or otherwise make available to, or permit the use of, such licensed source code and object code to those third parties which license Adobe proprietary software applications. Licensor grants no other rights in the source code of

"HTML Snapshot" except the rights expressly granted hereunder. Licensor retains all rights, title and ownership of the source code of "HTML Snapshot." Licensor authorizes Adobe to deposit the source code of "HTML Snapshot" with an independent third party software escrow agent along with the source code of its own proprietary software applications upon the request of Adobe' customers who have licensed its proprietary software applications. Notwithstanding the foregoing, in the event the source code of "HTML Snapshot" by such independent third party escrow agent is released to a licensee of Adobe, such third party customer shall be contractually restricted from using the source code of "HTML Snapshot" for the purposes of creating any derivative works therefrom or from distributing such source code to any third parties. Licensor prohibits Adobe from using the licensed source code of "HTML Snapshot" to release a product that is competitive with "HTML Snapshot." Licensor is providing the source code license of "HTML Snapshot" "AS IS." LICENSOR EXPRESSLY DISCLAIMS ANY WARRANTY, WHETHER EXPRESS OR IMPLIED. ADOBE USE OF THE SOURCE CODE LICENSE OF "HTML SNAPSHOT" SHALL BE AT ITS OWN RISK. LICENSOR AND ITS EMPLOYEES, AGENTS AND REPRESENTATIVES EXPRESSLY DISCLAIM ANY LIABILITY FOR LOSS OF DATA, DAMAGES, LOSS OF PROFITS OR ANY OTHER KIND OF LOSS ARISING FROM OR RELATING TO ADOBE USE OF THE SOURCE CODE LICENSE OF "HTML SNAPSHOT."

MD5: Copyright (C) 1999, 2000, 2002 Aladdin Enterprises. All rights reserved. This software is provided 'as-is', without any express or implied warranty. In no event will the authors be held liable for any damages arising from the use of this software. Permission is granted to anyone to use this software for any purpose, including commercial applications, and to alter it and redistribute it freely, subject to the following restrictions: 1. The origin of this software must not be misrepresented; you must not claim that you wrote the original software. If you use this software in a product, an acknowledgment in the product documentation would be appreciated but is not required. 2. Altered source versions must be plainly marked as such, and must not be misrepresented as being the original software. 3. This notice may not be removed or altered from any source distribution.

StdString: COPYRIGHT: 1999 Joseph M. O'Leary. This code is free. Use it anywhere you want. Rewrite it, restructure it, whatever. Please don't blame me if it makes your \$30 billion dollar satellite explode in orbit. If you redistribute it in any form, I'd appreciate it if you would leave this notice here.