Introducing the Pentaho BI Suite Community Edition

About This Document

If you have questions that are not covered in this guide, or if you find errors in the instructions or language, please contact the Pentaho Technical Publications team at documentation@pentaho.com. The Publications team cannot help you resolve technical issues with products.

Support-related questions should be submitted through the Pentaho Community Forum at http://forums.pentaho.org/. There is also a community documentation effort on the Pentaho Wiki at: http://wiki.pentaho.org/ that you may find helpful.

For information about how to purchase enterprise-level support, please contact your sales representative, or send an email to sales@pentaho.com.

For information about instructor-led training on the topics covered in this guide, visit http://www.pentaho.com/training.

Limits of Liability and Disclaimer of Warranty

Trademarks

Company Information

Pentaho Corporation
Citadel International, Suite 340
5950 Hazeltine National Drive
Orlando, FL 32822
Phone: +1 407 812-OPEN (6736)
Fax: +1 407 517-4575
http://www.pentaho.com

E-mail: communityconnection@pentaho.com
Sales Inquiries: sales@pentaho.com
Documentation Suggestions: documentation@pentaho.com

Sign up for our newsletter:
http://community.pentaho.com/newsletter/

Contents

Introduction

The Pentaho BI Suite Community Edition is an open source business intelligence package that includes ETL, analysis, metadata, and reporting capabilities. It is entirely open source software, licensed mostly under the GNU General Public License version 2, with parts under the LGPLv2, the Common Public License, and the Mozilla Public License. Pentaho optimizes, platform-tests, and guarantees certified builds of the BI Suite; this enhanced version of the software is packaged with a powerful service management tool called Enterprise Console, user and developer support, IP indemnification, and professional documentation, and sold by Pentaho as Enterprise Edition.

The purpose of this guide is to introduce new users to the Pentaho BI Suite, explain how and where to interact with the Pentaho community, and provide some basic instructions to help you get started using the software.

Community Edition or Enterprise Edition?

The BI Suite Community Edition is ideal for:

- Business intelligence aficionados
- Open source software programmers
- Early adopters
- College students

Pentaho no longer suggests using Community Edition for enterprise evaluations. If you are a business user interested in trying out the BI Suite Enterprise Edition, follow the Enterprise Edition evaluation link on the pentaho.com front page, or contact a Pentaho sales representative.

The enhancements, service, and support packaged with the BI Suite Enterprise Edition are designed to accommodate production environments, especially where downtime and time spent figuring out how to install, configure, and maintain a business intelligence solution are prohibitively expensive. If your business will save money or make more money as a result of a successful business intelligence implementation, then Enterprise Edition is the most appropriate choice.

Community Edition Support Options

As a Pentaho BI Suite Community Edition user, you will have to install, configure, and maintain the software on your own. Your only support options are the community forum (http://forums.pentaho.org) and the community Wiki (http://wiki.pentaho.com). If you do not find an answer right away, please be a good community participant and contribute a Wiki article that explains the solution after you've figured it out.

At any time, you can contact Pentaho sales and upgrade to Enterprise Edition. Enterprise Edition customers get phone support, access to Pentaho software engineers, and a Web-based knowledge base that is updated weekly with new support articles, tips, and comprehensive user guides.

The Pentaho Client Tools

The Pentaho client tools are:

- **Report Designer:** An advanced report creation tool. If you want to build a complex data-driven report, this is the right tool to use. Report Designer offers far more flexibility and functionality than the ad hoc reporting capabilities of the Pentaho User Console.
- **Design Studio:** An Eclipse-based tool that enables you to hand-edit a report or analysis view xaction file. Generally, people use Design Studio to add modifications to an existing report that cannot be added with Report Designer.
- **Aggregation Designer:** A graphical tool that helps improve Mondrian cube efficiency.
- **Metadata Editor:** Enables you to add a custom metadata layer to an existing data source. Usually you would do this for a data source that you intend to use for analysis or reporting; it's not required, but it makes it easier for business users to parse the database when building a query.
- **Pentaho Data Integration:** The Kettle extract, transform, and load (ETL) tool, which enables you to access and prepare data sources for analysis, data mining, or reporting. This is generally where you will start if you want to prepare data for analysis.
- **Schema Workbench:** A graphical tool that helps you create ROLAP schemas for analysis. This is a required step in preparing data for analysis.

After they're installed, you can find all of these tools in their own directories in the `/pentaho/design-tools/` directory. The scripts that run them should be fairly self-explanatory. If you are using Windows, there should be a Pentaho program group with icons that will initialize the BI Server and run the client tools.

Installation

Follow the instructions below to download and install the Pentaho BI Suite Community Edition.

Hardware Requirements

The Pentaho BI Suite software does not have strict limits on computer or network hardware. As long as you meet the minimum software requirements (note that your operating system will have its own minimum hardware requirements), Pentaho is hardware agnostic. There is, however, a recommended set of system specifications:

RAM	At least 2GB
Hard drive space	At least 1GB
Processor	Dual-core AMD64 or EM64T

It's possible to use a less capable system, but in most realistic scenarios, the too-limited system resources will result in an undesirable level of performance.

Your environment does not have to be 64-bit, even if your processor architecture supports it.

Software Requirements

Note: The system requirements listed below apply to the BI Suite. The BI Suite graphical installation utility, however, will only work on Windows or Linux.

In terms of operating systems, Windows XP with Service Pack 2, modern Linux distributions (SUSE Linux Enterprise Desktop and Server 10 and Red Hat Enterprise Linux 5 are officially supported, but most others should work), Solaris 10, and Mac OS X 10.4 are all officially supported.

No matter which operating system you use, you must have the Sun Java Runtime Environment (JRE) version 1.5 (sometimes referenced as version 5.0) installed. 1.4.2 will not work, and while 1.6 (6.0) will probably work in most cases, Pentaho does not yet officially support it.

Note: The GNU Compiler for Java, or GCJ for short, interferes with the way many native Java programs work on Linux, including some of the components of the Pentaho BI Suite. If you are using a Linux distribution that installs GCJ by default (which includes all of the most popular distros), then before you begin installation you must remove, disable, or circumvent GCJ. If you cannot remove it, you can simply ensure that your JAVA_HOME variable is properly set, and add the Java Runtime Environment's /bin/ directory to the beginning of your PATH variable in ~/.bashrc or /etc/environment, then relog before continuing.

Workstations will need to have reasonably modern Web browsers to access Pentaho's Web interface. Internet Explorer 6 or higher; Firefox 3.0 or higher (or the Mozilla or Netscape equivalent); and Safari 2.0.3 or higher will all work.

Your environment can be either 32-bit or 64-bit as long as it meets the above requirements.

The aforementioned configurations are officially supported by Pentaho. Other operating systems such as Windows Vista, FreeBSD, and OpenBSD; other Java virtual machines like Blackdown;

other application servers such as Liferay and Websphere; and other Web browsers like Opera may work without any problems. However, the Pentaho support team may not be able to help you if you have trouble installing or using the BI Suite under these conditions.

Note: Some Pentaho client tools, such as Metadata Editor, Aggregation Designer, Pentaho Data Integration, and Design Studio, require that the Eclipse SWT JAR be in your Java classpath. This can be an issue in scenarios where standalone client tools are installed onto a machine that does not also have the BI Platform installed, particularly on platforms other than Windows and Linux.
Note: The Pentaho Reporting engine requires a graphical environment in order to create charts. If you are installing the BI Platform onto a headless Linux, BSD, or Solaris server and do not have X11R6 on it, you should install the **Xvfb** package on your server to satisfy the charting dependency.

Downloading and Installing the BI Suite

Follow the below process to download and install the Pentaho BI Suite Community Edition.

1. Open a Web browser and navigate to the Pentaho page on SourceForge.net.

   http://sourceforge.net/projects/pentaho/ . If you cannot click on links in this document, you can simply navigate to http://sourceforge.net/projects/pentaho/
2. Click **Download**.
3. At the SourceForge download screen, click **Business Intelligence Server**.
4. In the **Latest** category at the top of the list, click either the .zip or .tar.gz file for the **biserver-ce** project.

   This is an archive package of the Pentaho BI Platform, along with a Tomcat Java application server configured to run it. There is no functional difference between the zip and tar archives; they are merely in compression formats that are generally preferred by Windows and Linux users, respectively.
5. Once the file is downloaded, create a `/pentaho/server/` directory in an accessible place in your filesystem, and unpack the files using your preferred archive utility.

   Ideally you would be unpacking this on what you expect to be your server, though there is no reason why you can't install the Pentaho client tools on the same machine.
6. Repeat this process (creating a `/pentaho/design-tools/` directory to contain them) for any or all of the following Pentaho client tool projects:

   - Report Designer
   - Pentaho Metadata
   - Design Studio
   - Data Integration
   - Schema Workbench

   You may not need all of these tools, but it can't hurt to download all of them.

You have retrieved all of the relevant Pentaho software, and are ready to configure the BI Platform.

Starting the BI Platform

In order to use and configure the Pentaho BI Platform, you must start the BI Server, then the Pentaho Administration Console.

Starting the BI Server

To start the BI Server, run the **start-pentaho** script in the `/pentaho/server/biserver-ce/` directory.

Starting the Pentaho Administration Console

To start the Pentaho Administration Console, run the **start** script (on Windows) or **startup** script (on Linux) in the `/pentaho/server/administration-console/` directory.

Configuring the BI Platform With the Administration Console

Follow the below process to log into the Pentaho Administration Console.

1. Open a Web browser and type in the Web or IP address of the Pentaho Administration Console server, which is `http://localhost:8099/admin` by default.
2. Type in your administrator credentials, then click **Login**.

   The default credentials are **admin** for the user, and **password** for the password.
3. Click the **Administration** tab on the left side of the window.
4. Remove the default sample users and roles and create your own.
5. Click the **Data Sources** tab at the top of the window.
6. Enter the connection details for the data source you want to use for reporting and analysis.

   By default, there is a **sampledata** source listed. If you intend to follow the examples later in this guide, you must leave this data source intact.

You are now logged into the Pentaho Administrator Console and ready to finish configuring the BI Platform.

Getting Started

Your workflow will vary depending on your BI goals. Typically, Pentaho BI Suite users will start with Pentaho Data Integration to prepare a data source, then use Metadata Editor to create a metadata layer for that data source, then potentially Schema Workbench to create a ROLAP schema. At that point, you're ready to create reports and analysis views.

If you just want to create a quick report, the ad hoc reporting component of the Pentaho User Console is the best tool for the job. If you want to create a detailed report, go directly to Report Designer instead. If you have created a ROLAP schema, then you can do some data exploration first by using an analysis view, which allows you to drill down into the smallest of details in a data source.

Ideally, everything will end up being published to the BI Platform, which enables you to display, run, and share your reports with others, or to schedule them to run at given intervals.

Once you've got some reports and/or analysis views that you like, you might create some dashboards that display them in creative and useful ways for your business users.

Follow the instructions below to log into the Pentaho User Console and familiarize yourself with its graphical interface.

How to Log Into the Pentaho User Console

Follow the below process to log into the Pentaho User Console.

1. Open a Web browser and type in the Web or IP address of the Pentaho server, which is `http://localhost:8080/pentaho/` by default.

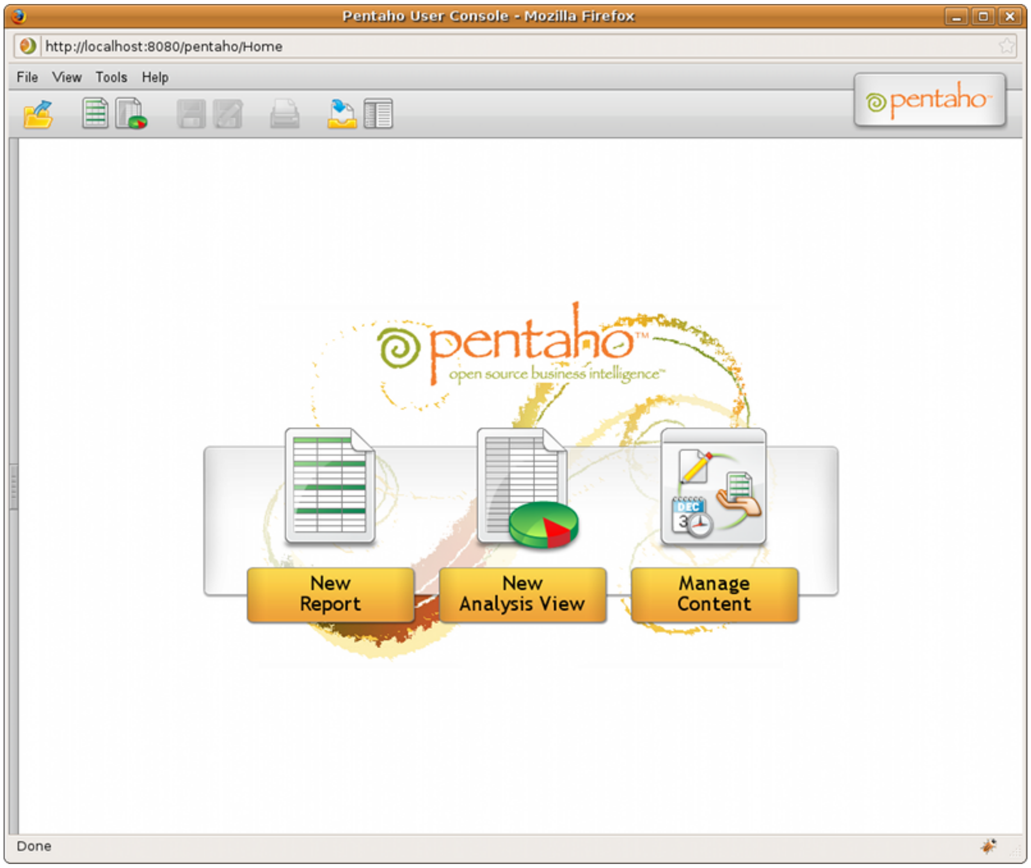
   You'll see an introductory screen with some Pentaho-related information and a **Login** button in the center of the screen.
2. Click **Login**.

   The login dialog will appear.
3. For the locally installed version of the BI Suite, select **Joe** from the user drop-down box, and type in **password** into the password field, then click **Login**. For hosted demo users, select **Guest** and type in **guest** as the password instead.

You are now logged into the Pentaho User Console and ready to start creating and running reports.

Navigating the Pentaho User Console

The first thing you will see when logging in is the quick launch screen, shown here:



If you'd like to experiment on your own before continuing on to the tutorials, click one of the three icons in the center of the screen to create a new ad hoc report, start a new analysis view, or edit existing solutions.

The button bar near the top of the page also contains icons for creating new ad hoc reports and analysis views, along with a button to print the current report or analysis view, and to open a previously saved solution.

Different user roles have different levels of access in the Pentaho User Console. The menu above the button bar performs these same functions as the buttons, plus administrative actions if you are logged in as an administrator, and also offers access to My Workspace and external links to help and support resources.

The three buttons in the quick launch screen will appear when you log into the Pentaho User Console for the first time, and when you close all tabs in the solution browser.

You can change views between My Workspace and the solution browser at any time by clicking the rightmost icons in the top button bar, or through the **View** menu.

Tutorials

The below sections offer, in no specific order, basic tutorials for the three major pillars of the BI Suite: Reporting, analysis, and data integration. These tutorials assume that you are working with the included sample data source, and that you have Report Designer and Pentaho Data Integration installed, and that you are logged into the Pentaho User Console.

Report Designer Tutorials

The tutorials below are for Pentaho Report Designer.

Starting Report Designer

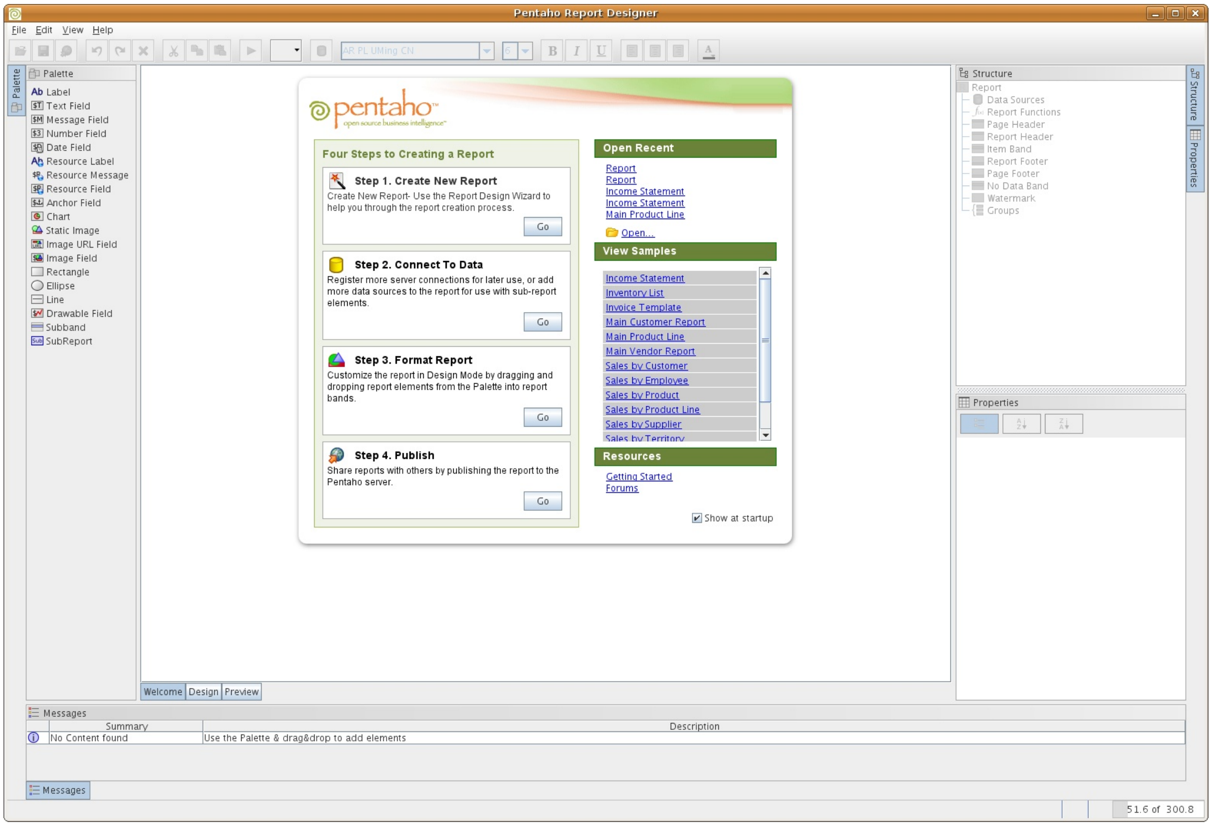
Follow the below process to start Report Designer.

1. Click the **Report Designer** entry in the **Pentaho** folder in the **Programs** section of your Start menu, or navigate to the **/pentaho/design-tools/report-designer/** directory and run **ReportDesigner.exe** on Windows, or **startdesigner_linux.sh** on Linux.
2. Before the program starts, it will run a version checking utility. You can click either option in the version checker screen to start Report Designer – if you just downloaded this file, it is assuredly the most current version, so you may not need to enable this feature right now.

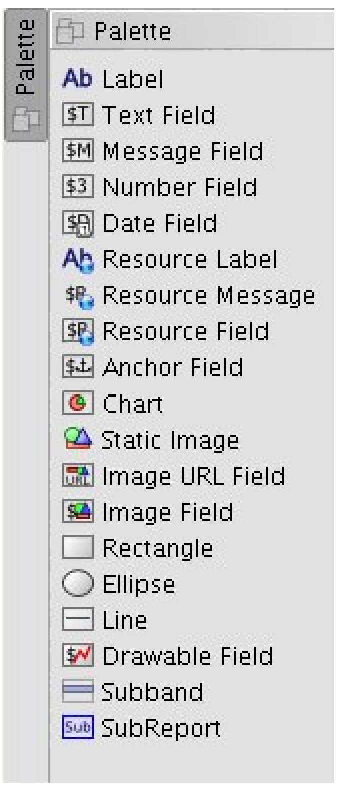
After the version check is complete (or skipped), Report Designer will start. Report Designer displays a Welcome screen and a default workspace at startup. The Welcome screen provides you with a brief introduction to the program, some instructions for getting started, and access to sample content and recently opened reports. For this walkthrough, you won't be following the instructions on the Welcome screen, though if you would like to experiment with Report Designer a little before continuing, feel free.

Exploring the Report Designer Interface

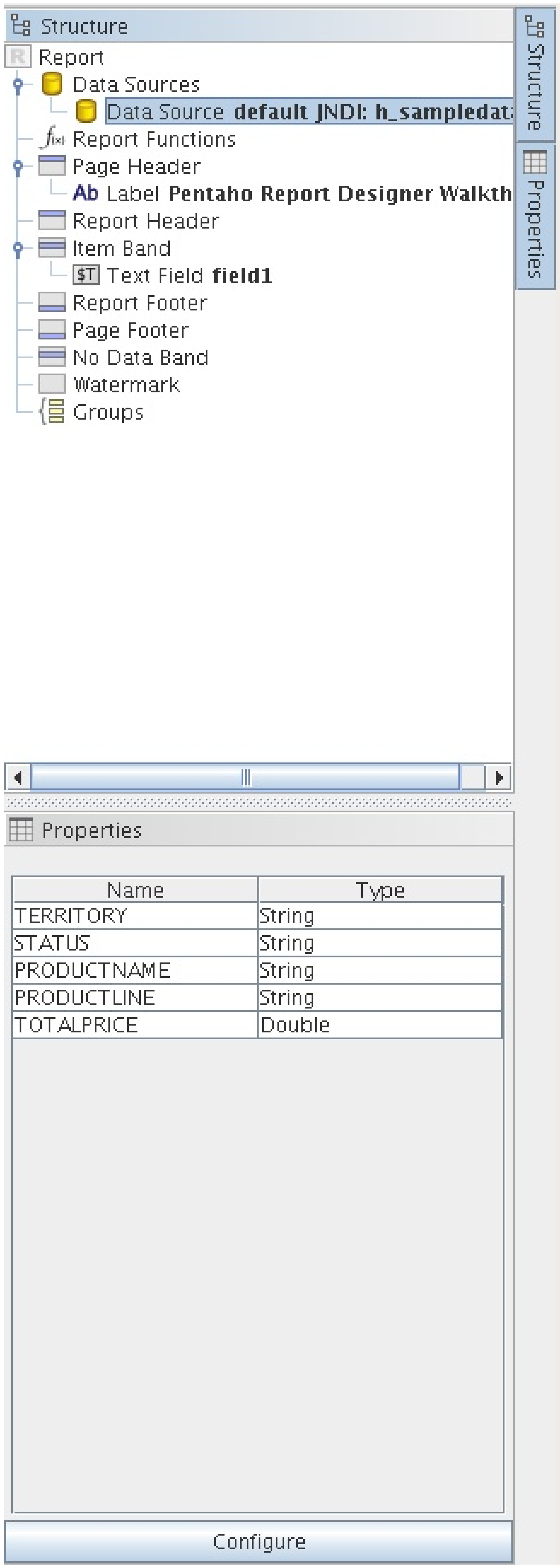
Report Designer's interface is similar to that of other graphic design and layout tools. A typical menu and button bar are at the top of the screen; a tool palette for adding design elements is on the left; and on the right are two panes that contain data and structure elements, and show the properties of a selected report element.



The Palette is where most of the design tools are:



The Properties and Structure panes show your data source and report elements, and their configurable options:


Structure
Report
Data Sources
Data Source default JNDI: h_sampledat
Report Functions
Page Header
Ab Label Pentaho Report Designer Walkth
Report Header
Item Band
$T Text Field field1
Report Footer
Page Footer
No Data Band
Watermark
Groups
Structure
Properties
Properties
Name
Type
TERRITORY
String
STATUS
String
PRODUCTNAME
String
PRODUCTLINE
String
TOTALPRICE
Double
Configure

Creating a Basic Report

As an introduction to Report Designer, follow this procedure to create a simple report that shows which products are associated with cancelled orders, sorted by product line and territory.

1. From the Welcome screen, click the **Design** tab in the lower left corner of the workspace.

   The Design view enables you to create and adjust your report's layout by hand.
2. Go to the **File** menu, then click **Open**. In the ensuing file dialog, navigate to `/bi-suite-2.0.0/client/report-designer/samples/`, select the **evaluation_blank.report** file, then click **Open**.

   The sample report file you just opened was designed specifically for this evaluation guide. It comes with a predefined data source linked to the sample data included in the evaluation package. The only element in this report is a header at the top that says "Pentaho Report Designer Sample Report." You can remove this header if you like.
3. Drag and drop a **Label** element from the Palette onto the **Page Header** band.

   The page header appears at the absolute top of every page in the report.
4. Double-click the label to edit it, then type in **Cancelled orders for all regions**.
5. Click the font size drop-down box in the top toolbar, and select **24** in the list.

   This will increase the text size to 24 point, which is more appropriate for a page title than 12 point.
6. Click and drag the horizontal resize handles on the Label element you just created until the field is as large as the page; drag the vertical resize handles until the text fits appropriately in the space.

   Expanding the field out to the size of the page will allow you to center the text more easily.
7. With the label selected, click the **AlignCenter** icon on the left side of the top toolbar.

   The text is now perfectly centered in the page header.
8. In the Structure pane, click **Data Source default JNDI:SampleData** to select it.
9. In the Properties pane, click **TERRITORY** to select it, then click and drag it into the left side of the **Item Band** section of the grid.

   The Item Band will repeat itself for as many iterations of the query elements as there are. So if TERRITORY contains 5 rows, there will be 5 item elements in the report, spread across as many pages as is necessary.
10. Repeat the previous step for the **PRODUCTLINE** column, but drop it into the center of the Item Band instead.
11. Repeat the previous step for the **PRODUCTNAME** column, but drop it into the right side of the Item Band instead.
12. Drag the TERRITORY field to the absolute top left of the Item Band.
13. Drag the PRODUCTLINE to the absolute top center of the Item Band, then click the **AlignCenter** icon in the top toolbar.
14. Drag the PRODUCTNAME to the absolute top right of the Item Band, then click the **AlignCenter** icon in the top toolbar.
15. Drag the TOTALPRICE to the absolute top right of the Item Band, then click the **AlignRight** icon in the top toolbar.

16. Drag the resize handles on the **TERRITORY**, **PRODUCTLINE**, **PRODUCTNAME**, and **TOTALPRICE** fields so that they fill all of the horizontal space in the grid. Don't change their vertical size -- just adjust their horizontal borders and orientations so that they are at the top of the Item Band with their borders abutting to one another and the sides of the report, ensuring that none of them overlap.

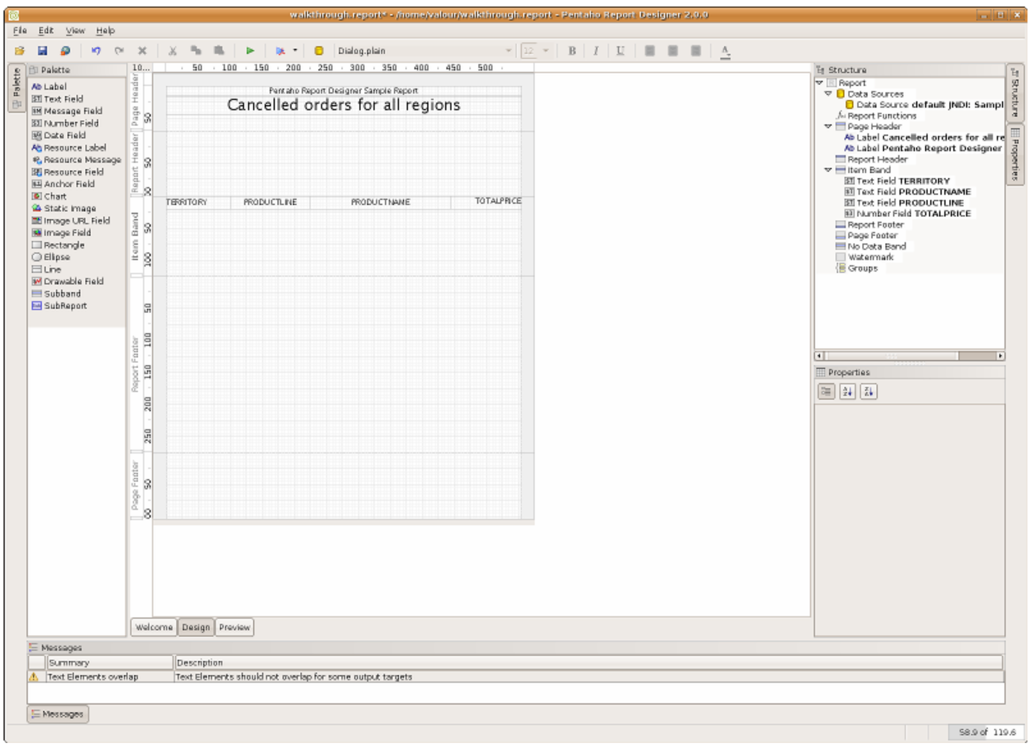
    The three fields should appear to be one solid line when you're done resizing and repositioning them.

    CAUTION: Make sure elements do not overlap; if they do, they may not appear in the output.
17. Click **Preview** near the bottom left of the grid.

    The Preview tab is a good way to verify the effects of any changes you make to a report.
18. Go to the **View** menu, then select **Preview as PDF**.

    A PDF version of the report will appear in your default PDF reader.

You now have a basic report that shows information that would be quite useful to the fictional company it belongs to. If you'd like to see some of Report Designer's more advanced functionality, continue on to the next section. Otherwise, skip ahead to the next portion of this guide that appeals to you.



Refining Your Report

To complete this process, you must have created a report according to the directions in the previous section. You should have the modified **evaluation_blank.report** file open in Report Designer at this time.

Being able to view properly refined and constrained data is useful, but Report Designer is capable of much more than just collecting and printing data; it also has many formatting options to increase the readability and visual appeal of your reports. Follow the below process to enhance your report with conditional formatting, borders, a company logo, a dynamic report date in the footer, and a pie chart. Feel free to go to the Preview tab at any time to see your progress.

1. Click in the grid and drag a selection box around all four fields in the Item Band.
2. In the Properties pane, click the right half of the **Border** property, then click the **...** button.
3. In the border dialog, click the checkbox next to **Same border for all sides**.

   This applies the settings from one border element to all eight.
4. Click any of the border sides, then select black for the color, 1 for the width, and SOLID in the **Dotted** drop-down box, then click **OK**.

   This will add a solid black line around each field.
5. Click the **Top** drop-down box in the **Vertical Alignment** property, and select **Middle**.

   This will center the text in the middle of the field, vertically.
6. Change the **Padding** property to **0;0;5;5**.

   This will put a little more space between the borders and the field text.
7. Click the **PRODUCTLINE** field to select it.
8. In the Properties pane, click the right half of the **Style Expressions** property, then click the **...** button.
9. Click **Add Expression**, then select **background-color** from the list, then click **OK**.
10. Click **background-color** to select it, then click **Edit Formula**.
11. Select **Logical** from the Category drop-down box.
12. Double-click the **IF** statement in the list on the left.
13. Erase the default value from the **A Condition** field, then click the field selection button to the right.
14. Select the **PRODUCTLINE** field, then click **OK**.
15. Add **="Classic Cars"** to the A Condition line.

    Your condition should look like this: **[PRODUCTLINE]="Classic Cars"**
16. Type in **"red"** and **"white"** into the true and false fields, respectively.
17. Click **OK** to leave the Formula Editor, then click **OK** to leave the Style Expressions dialog.

    All cancelled orders in the Classic Cars product line will now be highlighted in red.
18. Drag and drop a **Static Image** element from the Palette to the **Report Header** band.
19. In the Properties pane, click the right half of the **URL** property, click the **...** button, then navigate to `/pentaho/design-tools/report-designer/samples/logo.png`, then click **Open**.

    This graphic is a Steel Wheels Inc. logo developed by Pentaho for demonstration purposes.
20. Reposition and resize the logo until it is the width of the report.

    Report Designer will downsize images proportionately according to the size of the element.
21. Drag and drop a **Date field** element to the right corner of the Page Footer.
22. Double-click the new date field, and type in **Today's date** for a field name.
23. In the Properties pane, click the right half of the **Formula** field, then click the **...** button.

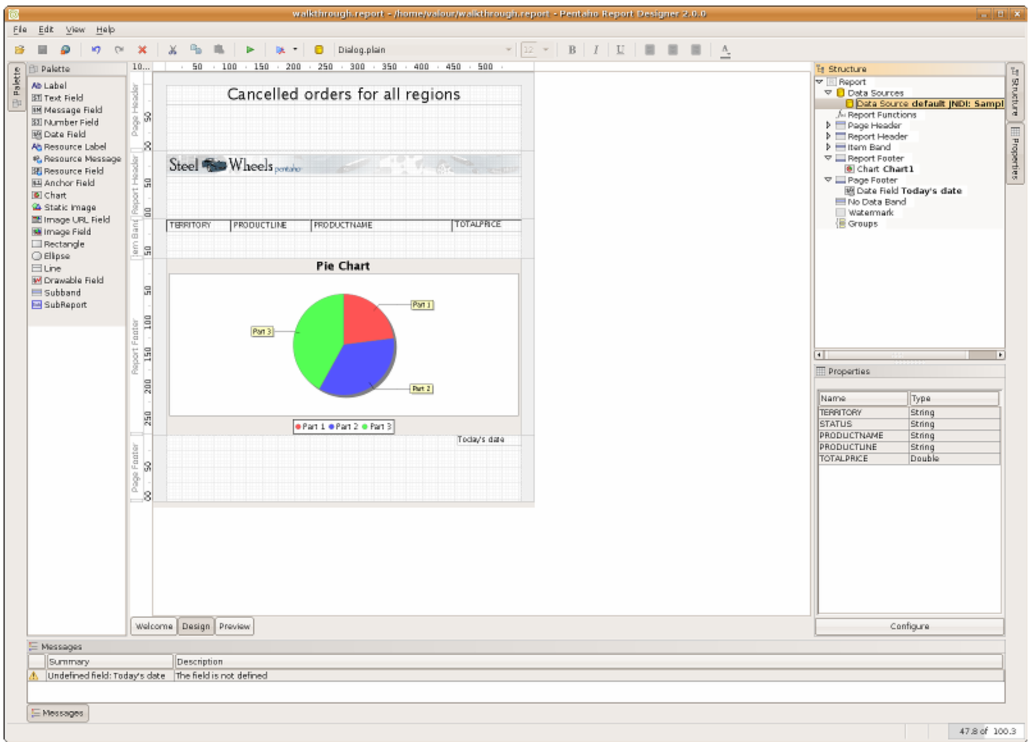
    The Formula Editor will appear.
24. Double-click **NOW** in the list on the left, then click **OK**.
25. Drag and drop a **Chart** element from the Palette to the Report Footer band, below the Item band.
26. In the Properties pane, click **Edit Chart**.
27. Click **Pie** in the list on the left.
28. Double-click the right half of the **Title** field, then type in **Highest Cancelled Order Costs**.

29. Click the right half of the **Value Column** field, then click **...**. In the ensuing popup dialog, select **TOTALPRICE**.

    This column determines the numerical data that the pie chart will represent.
30. Click the right half of the **Series Column** field, then click **...**. In the ensuing popup dialog, select **PRODUCTLINE**.

    This determines how the TOTALPRICE data will be constrained in the pie chart.
31. Un-check the **Summary Only** box.
32. Click **Apply** to save the chart settings.

Your report is now branded for the fictional Steel Wheels company, easier to read, and specifically highlights cancelled orders relating to the Vintage Cars product line. This is still only the beginning, but you should now have a good impression of the power and flexibility of Pentaho Report Designer.



Adding a Parameter and Publishing Your Report to the Pentaho BI Platform

In addition to locally executing a report directly from Report Designer, you can also publish it to the Pentaho BI Platform where it can then be viewed using the Pentaho User Console. If you've created a parameterized report (a report that requires user input through the BI Platform Web interface in order to run), you must publish to the BI Platform in order to take advantage of its interactive elements.

Follow the process below to add parameterization to your report, publish it to the BI Platform, and run it from the solution repository.

1. Edit the **default** data source by double-clicking its entry in the Structure pane.
2. In the **WHERE** portion of the SQL statement shown in the Query Details field, add the following line: `PRODUCTLINE = '${ENTER_PRODUCTLINE}' AND`

   This is the value that BI Platform users will be asked to supply when they run this report.

   Your WHERE statement should now look like this:

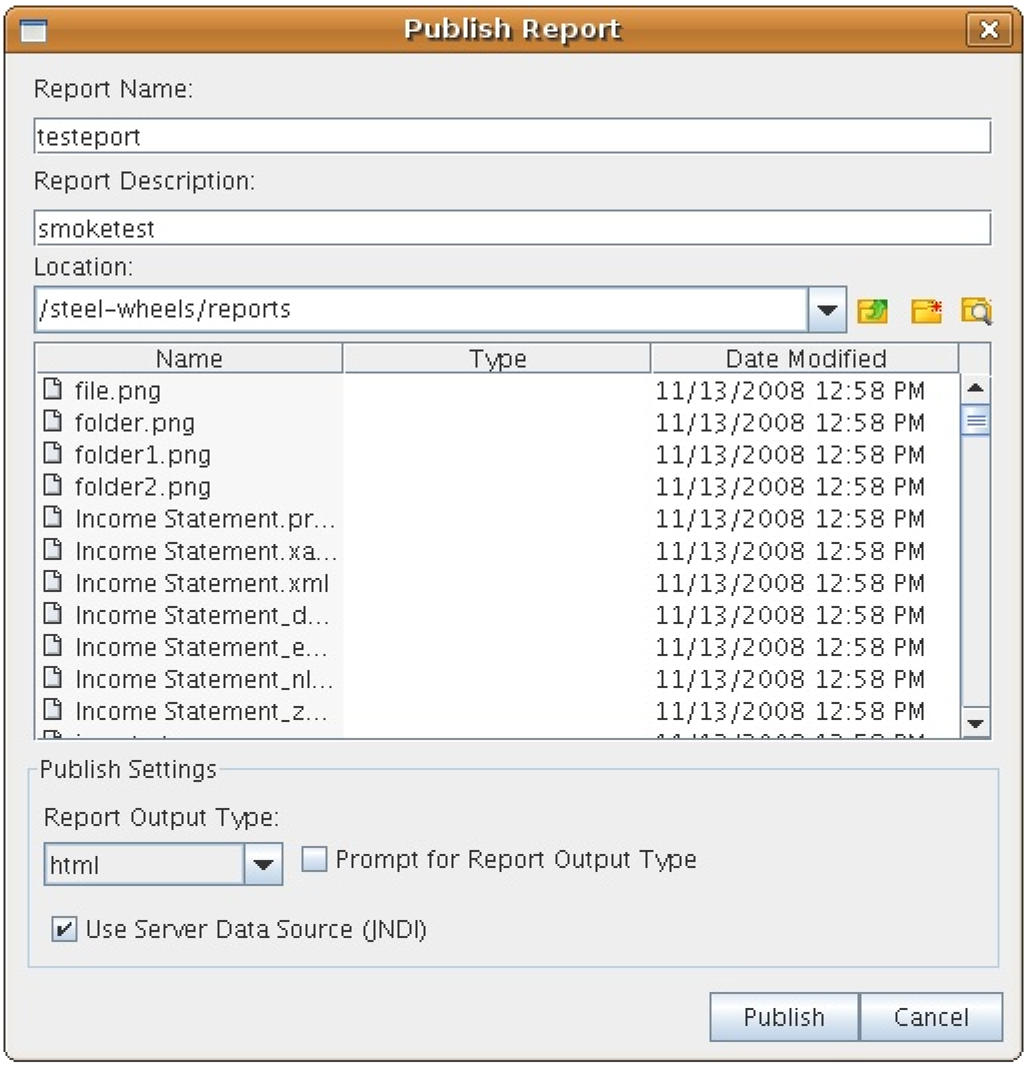
```
WHERE
    `PRODUCTLINE` = '${ENTER_PRODUCTLINE}' AND
    `ORDERFACT`.`STATUS` = 'Cancelled'
```

3. Click **OK** to save the query.

   Note: Once you add parameterization, you will be unable to preview your report in Report Designer. This is because the report has become dependent on user input via the BI Platform's Web interface. To regain preview functionality, simply remove the query line you added above.
4. From the **File** menu, select **Publish to Server...**

   The Repository Login dialog will appear, pre-populated with credentials valid for the evaluation sample data and users.
5. Type your publish password (you set this during installation) into the appropriate field.
6. Ensure that the port number is correct (usually either 8080 or 18080) then click **OK**.

   A Publish Report dialog will appear, as shown below.



7. Double-click the **steel-wheels** directory, then the **reports** subdirectory.
8. Type in a report name and description into the appropriate fields.
9. Select **html** from the Report Output Type drop-down box.
10. Click the checkbox next to **User Server Data Source (JNDI)**, then click **Publish**.

You should see a message that says the publish operation was successful.

11. Click **Yes** when prompted if you would like to go directly to the Pentaho User Console to log in and view the report you just published to the BI Platform.

    If you want to access the report later, just log into the BI Platform by going to **http://localhost:18080** (change the port number and hostname appropriately, if necessary) in your Web browser, then navigate to the **/steel-wheels/reports/** directory in the Solution Browser. You should see your published report in the list.
12. Log in as **Joe**. Joe's password is **password**.

13. In the **ENTER_PRODUCTLINE** field, type in any of the product lines from the sample data: Classic Cars, Motorcycles, Vintage Cars, Ships, Trains, or Planes, then click **Ok**.

    You can use your browser's back button to revisit this screen and enter a different product line if you wish.

You now have a dynamic report that BI Platform users can run at any time, with updated data, that concentrates on specific product lines. The conditional formatting is still set to give Classic Cars a red background, and the pie chart is configured to compare cancelled order costs among all product lines, so there is still some adjustment to do to put this report into production. However, at this point in your evaluation, you've seen most of Report Designer's key features; if you'd like to go back into Report Designer and make more changes to the cancelled orders report, feel free. This guide, however, will proceed to a walkthrough of the BI Platform's ad-hoc reporting capabilities.

Ad hoc Reporting Tutorial

You must be logged into the Pentaho User Console before continuing.

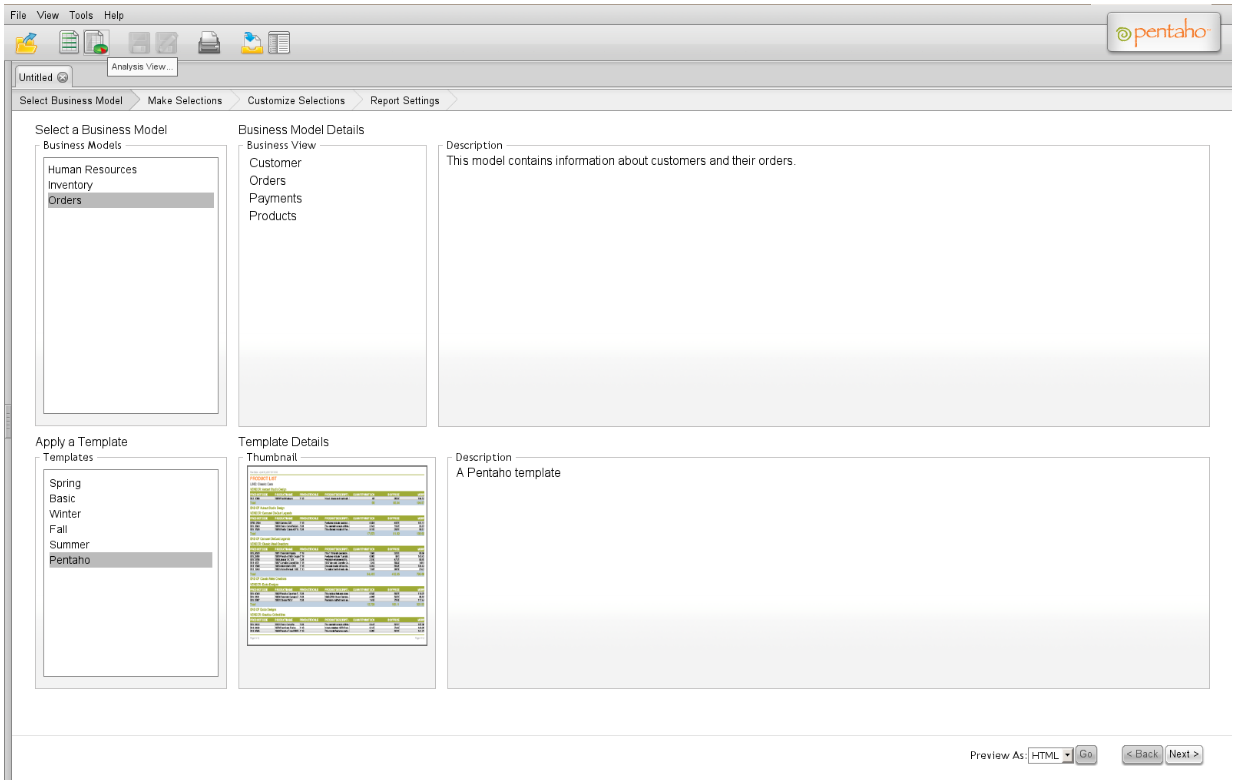
This walkthrough shows you how to create a simple, template-based report that shows which territory generates the most sales.

1. Click the **Create New Report** button in the middle of the Pentaho User Console screen.

   The ad hoc query wizard will start.
2. In the first step of the wizard, select **Orders** in the **Business Model Details** pane.

   A business model is another term for data set.
3. In the **Apply a Template** field, select a predefined report template that appeals to you.

   A thumbnail preview of the template will appear in the **Template Details** field. A template specifies a variety of properties in the report that affect its appearance, like font size and background colors for various report elements.



4. Click **Next**.
5. In the **Available Items** list, click the **Territory** business column and drag it to the upper right into the **Level 1 box**.

   This will determine how the data is grouped.
6. Drag and drop the **Amount** and **Buy Price** into the **Details** box on the right.

   This determines which fields to display for the given groups.
7. Click **Go** to preview how these new items have affected the report, then close the preview tab when you're done.
8. Click **Next**.
9. Click the **Territory** item in the **Groups** list.

   A list of general options will appear on the right.
10. Click **Center**.

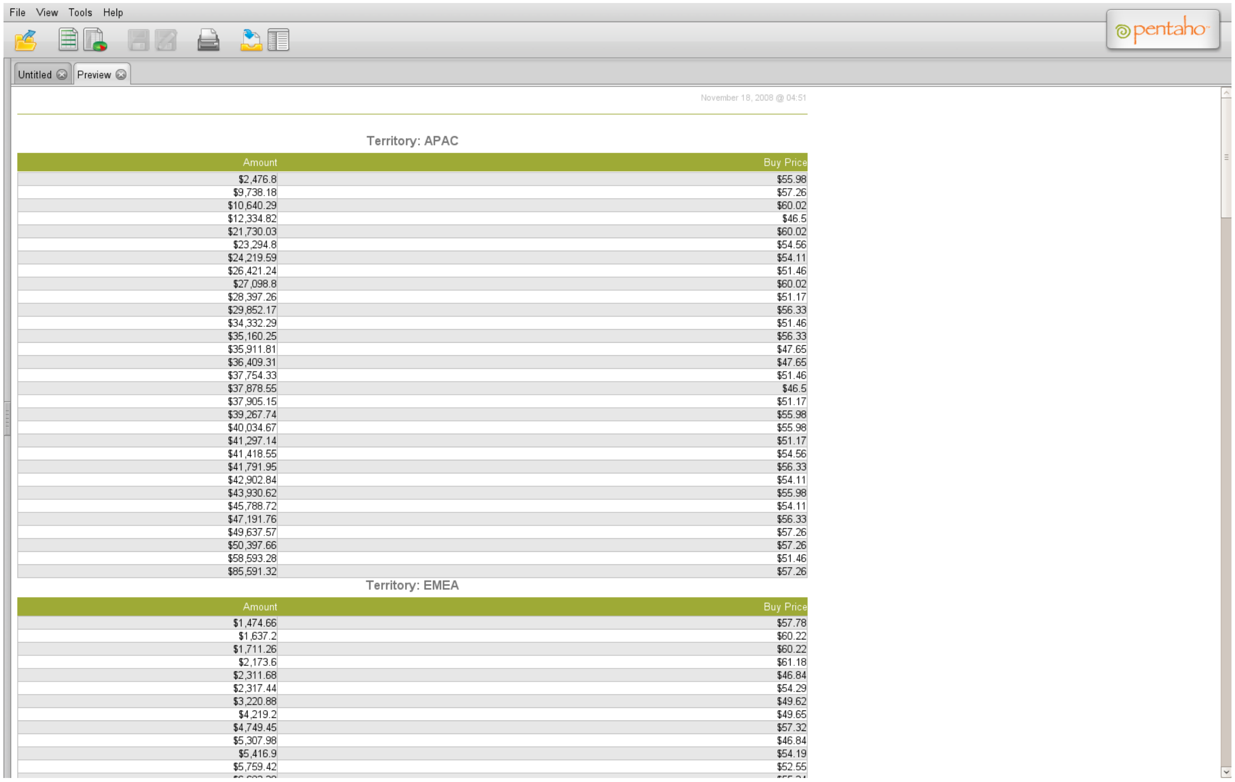
    This will center the territory name above each table, making it easier to read.
11. Click **Amount**, then click **Add** in the **Sort Detail Columns** area on the right.

    This will sort the sales amounts from lowest to highest.
12. Click **Go** to test the new change, or **Next** to continue to the next part of the wizard.
13. To set the header, footer, description, paper type, and page orientation, change the on-screen values for these elements accordingly.

    PDF is the only output type that has a concept of a page, so the **Page** portion of the **Header** and **Footer** sections only applies to PDFs.
14. Click the blue **Save** button in the top toolbar to save your report. In the ensuing file dialog, navigate to the location you want to save the report to, and type in a filename for the report.

You can continue to modify your report after it's been saved; just click **Save** to update the report file after you've made changes.

You now have a report that shows how much revenue is coming from each sales territory, and the itemized price of each purchased product. As you can see, ad hoc reporting is quicker and simpler than Report Designer, but doesn't offer nearly the same level of design detail, nor does it have advanced reporting features like conditional formatting or parameterization.



Analysis View Tutorial

Analysis views are similar to reports, except they're designed to be totally interactive and dynamic, whereas reports tend to be static or minimally interactive after they're created. Analysis views allow you to dynamically explore your data and drill down into it to discover previously hidden details. In this example, you'll try to find out which product line is responsible for the highest number of cancelled orders.

1. In the **File** menu, select the **New** sub-menu, then click **New Analysis View...**.

   This is one of several ways to create a new analysis view; all methods lead to the same screen in the Pentaho User Console.
2. Select the **SteelWheels** schema and **SteelWheels Sales** cube from the drop-down lists.
3. Click **OK** to continue.

   An analysis view will open in a new tab.
4. Open the **OLAP Navigator** by clicking the cube icon in the top button bar – it's the third from the left.

   A new table will appear above the analysis view fields. The default basis for comparison is **Measures**, though that is not very useful for finding returned products.

5. In the **Rows** section, click the two-tone square next to **Product**. This will move it up to the **Columns** section.
6. In the **Columns** section, click the funnel icon next to **Measures**. This will move it down to the **Filters** section.
7. Click **OK** to modify the analysis view.
8. Click the **+** next to **All Products** to drill down into it.
9. Click the **+** next to **All Status Types** to drill down into it.

   Clearly the **ships** category has the most cancelled orders.
10. Drill down further into ships by opening the OLAP Navigator again.
11. Click **Product** in the **Columns** section.
12. Un-check all products except **Ships**.
13. Click **OK**.
14. Click **Order Status**.
15. Un-check all statuses except **Cancelled**.
16. Click **OK**, then **OK** again to close the OLAP Navigator and refine the analysis view.
17. Click the **+** next to **Ships** to show its constituent product lines.
18. Notice that the **Carousel DieCast Legends** series has the highest number of cancelled orders.

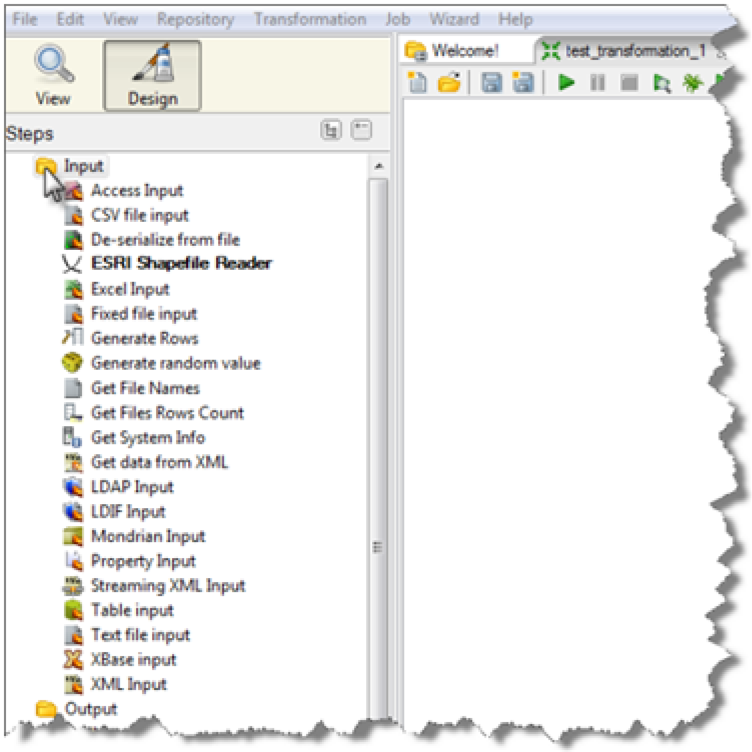
You now know which product line has the most cancelled orders.

Building a simple input-output transformation

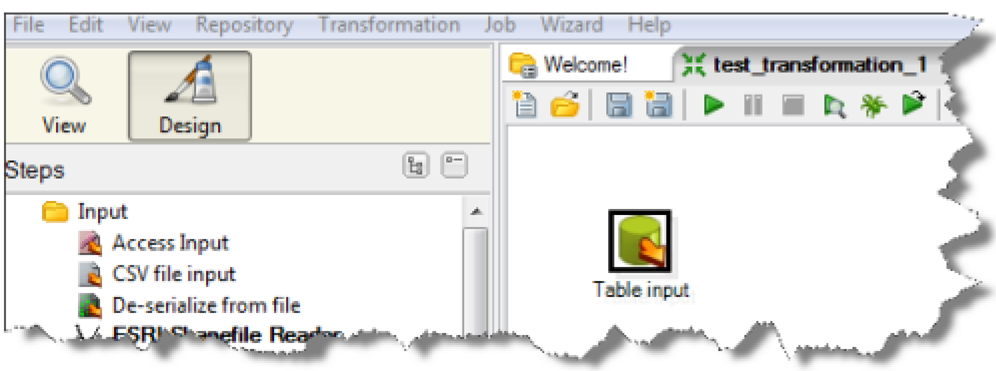
You must have a database connection to complete this exercise.

This exercise walks you through the process of building a transformation that uses input from a database table (a list of customers), applies an SQL statement, and outputs the data stream to a plain text file. This exercise is useful for testing and tuning transformations to see how the fields are passed in the data stream; it also shows you how to examine the results of each step of a transformation, as well as the final output.

1. In the left pane, click **Design**, then click **Input** to expand the folder and view the input step options.

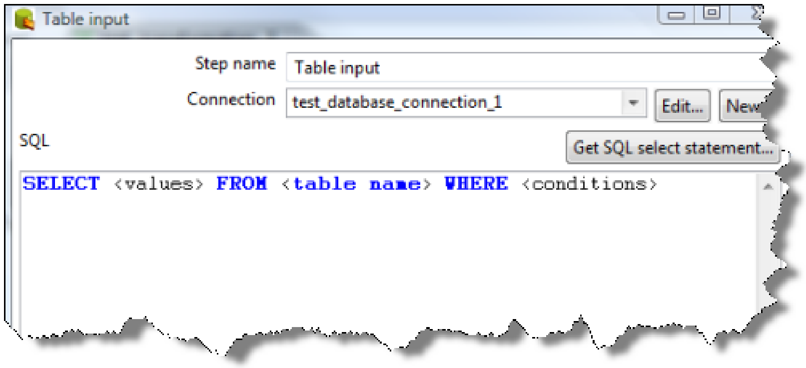


2. Click **Table Input** and drag the cursor to the right pane.



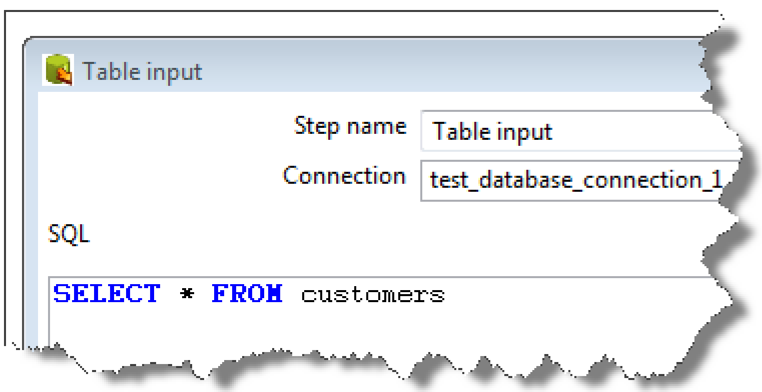
A **Table Input** step is added to the transformation. This step reads information from the database using the database connection and SQL.

3. Double-click **Table Input** to display a dialog box that allows you to view and edit the SQL statement.

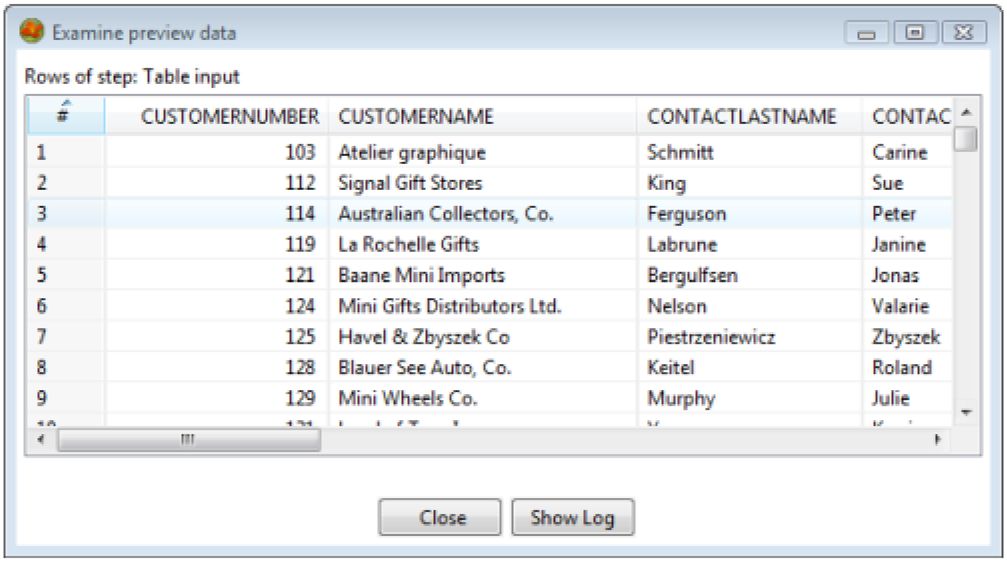


As a general rule, it is a good idea to maximize this window. This is a relatively simple example, but for longer statements, maximizing the window ensures that you see all parts of the statement. You must edit the following parameters to specify the table to use as input, the values get from the table, and (optionally) the conditions:

(SELECT) Values:	Specific value(s) or an asterisk * to indicate all values.
(FROM) Table Name:	Name of the database table, for example **customers**.
(WHERE) Conditions:	Specific condition(s), or may be deleted as in this example.

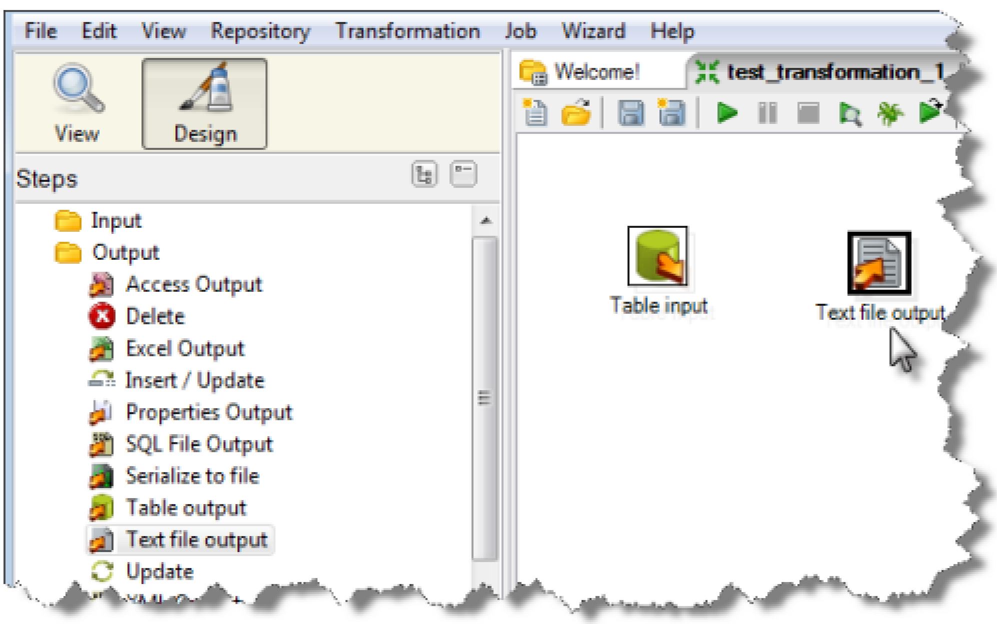


4. Click **Preview** to view the data stream flowing from the **Table Input** step. A dialog box appears that allows you to specify the number of rows you want to preview.



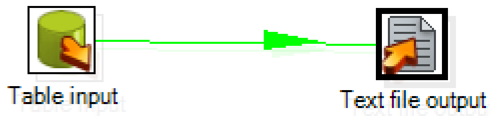
Preview shows you what is in the data stream coming into the step. If there are errors, a log file appears that includes a record of the errors to assist in troubleshooting.

5. In the left pane, click the **Output** folder to expand it and view the output step options. Click **Text File Output** and drag the cursor to the right pane.



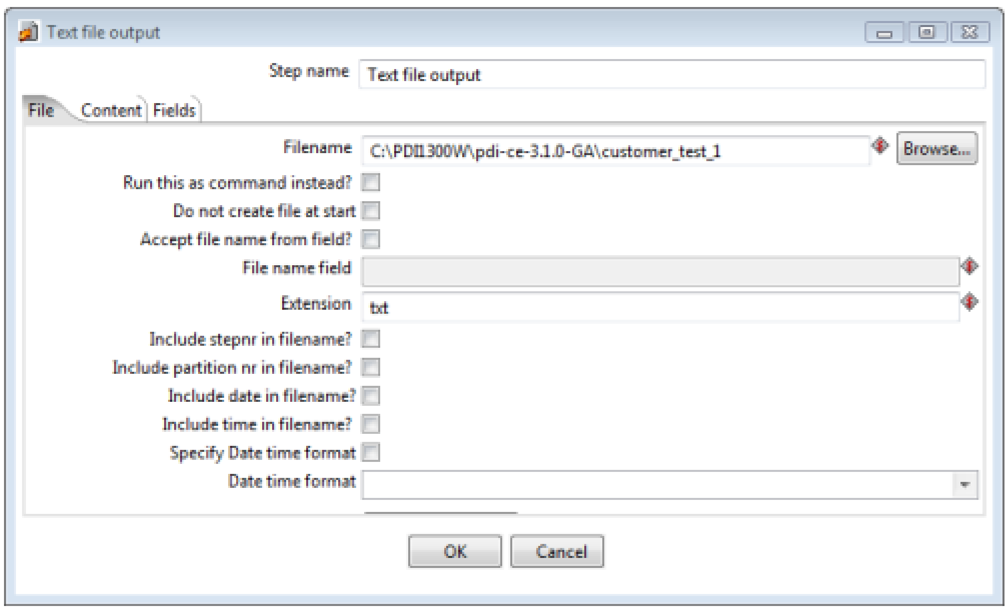
A **Text File Output** step is added to the transformation. This step sends the fields in the incoming data stream to a text file.

6. Click **Table Input**, then press <Shift> and drag the cursor to **Text File Output**.

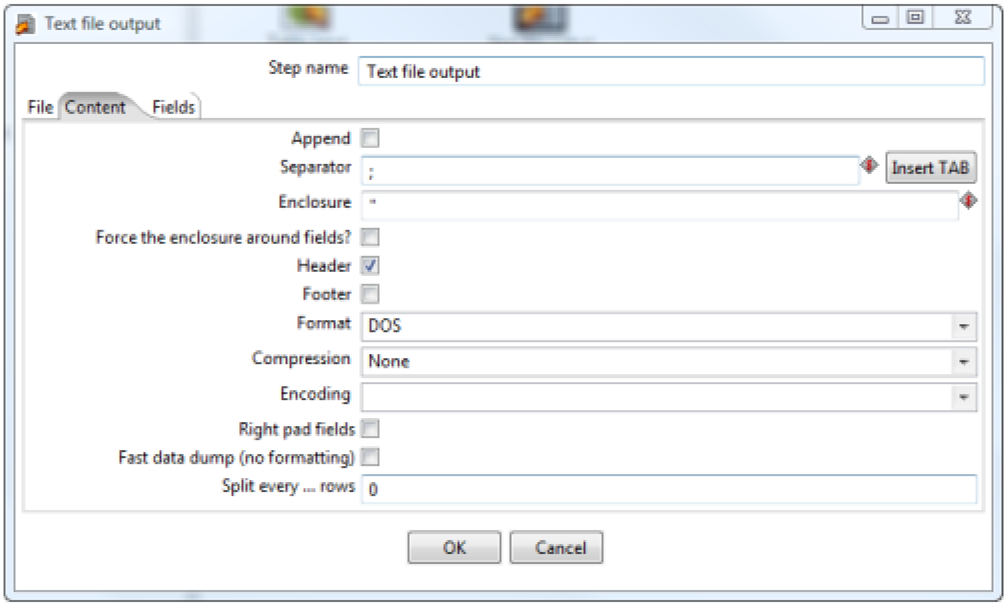


This adds a hop (data connection) between the two steps and displays as an arrow connecting the steps. The hop represents the data stream flowing from the **Table Input** step to the **Text File Output** step.

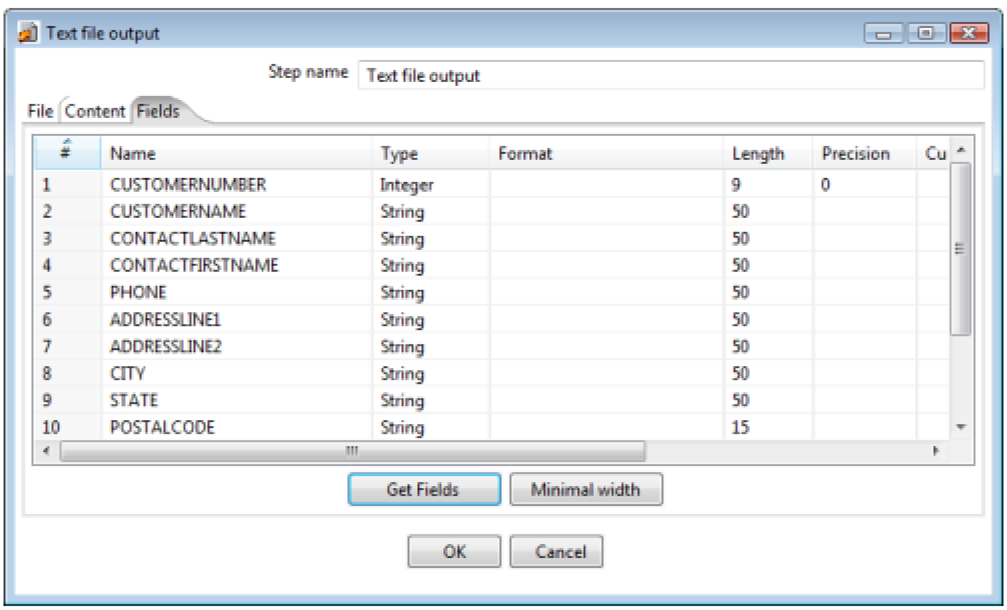
7. Double-click the **Text File Output** step and type a path/file name in the **File name** box.



The **Content** tab allows you to specify optional settings such as the field separator and the format (DOS or UNIX).

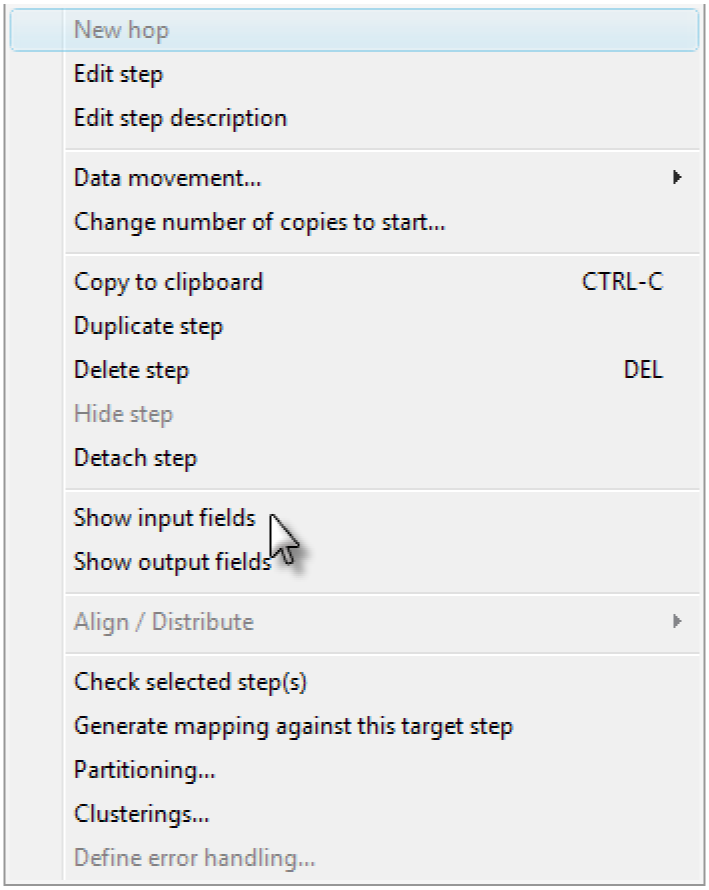


8. Click the Fields tab, then click **Get Fields**.

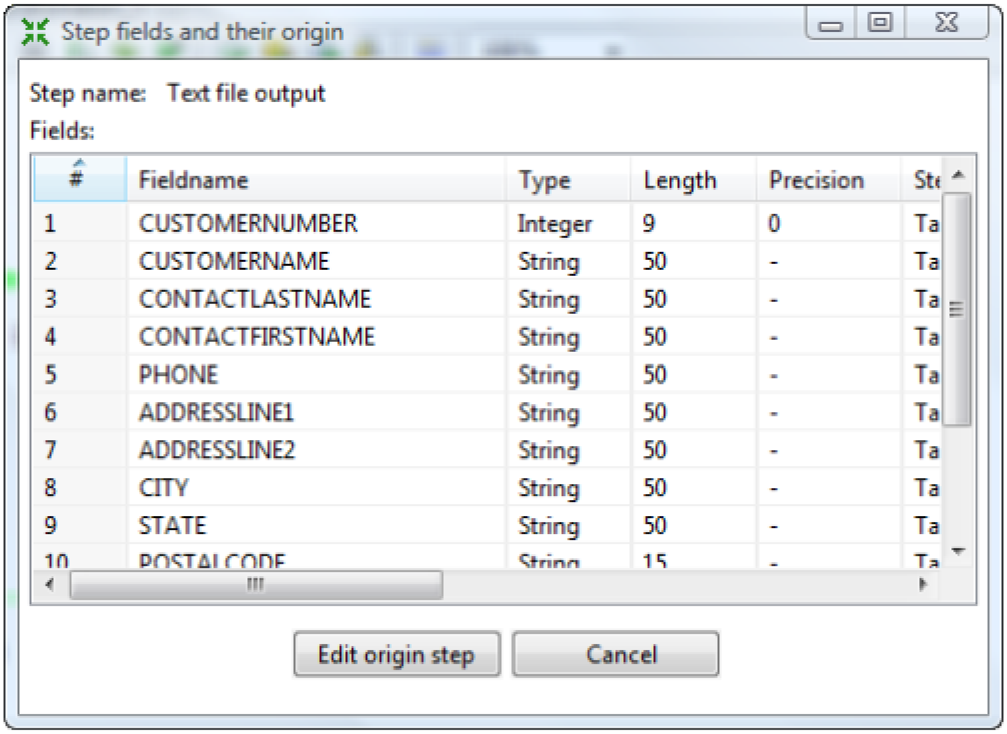


Note: The file extension (in this case, .txt) is filled in for you.
The Fields tab displays all the fields in the data stream that is flowing through the hop from **Table Input** step to the **Text File Output** step, and this data will in turn flow into the text file you defined. You can also type fields into this tab to add them to the data stream.

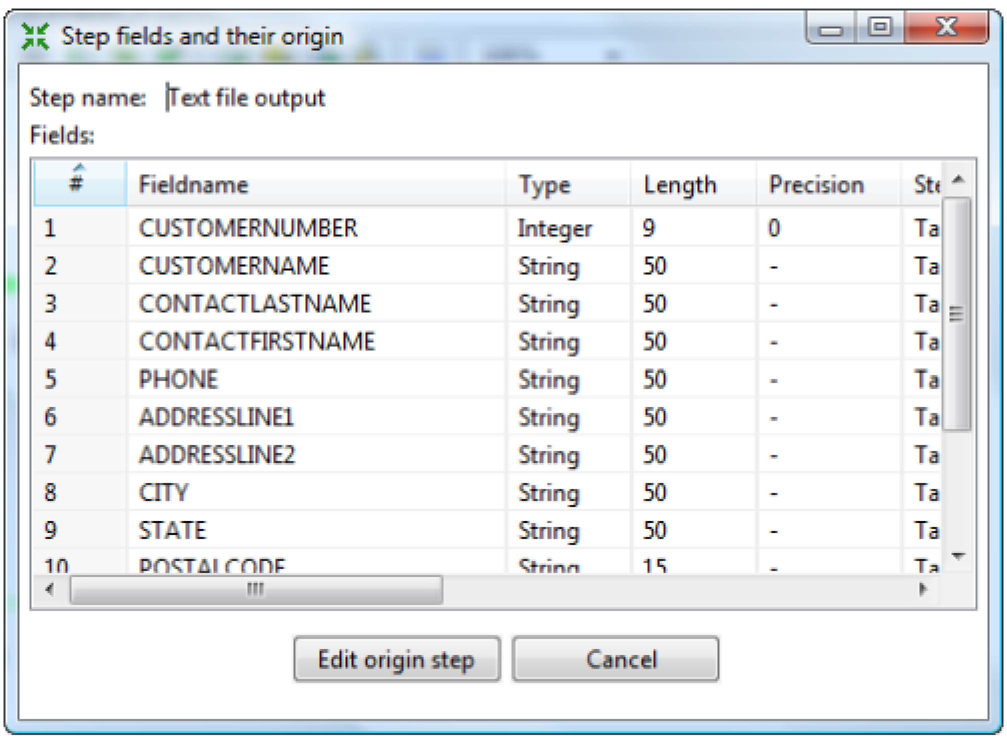
9. Right-click **Text File Output** to display the context menu, then click **Show Input Fields**.



This dialog box displays all the fields in the data input stream and their transformations. In this simple example, there are no intermediate transformation steps, for example filters, but in a transformation with intermediate steps, the details display here.



10. On the context menu, click **Show Output Fields**.

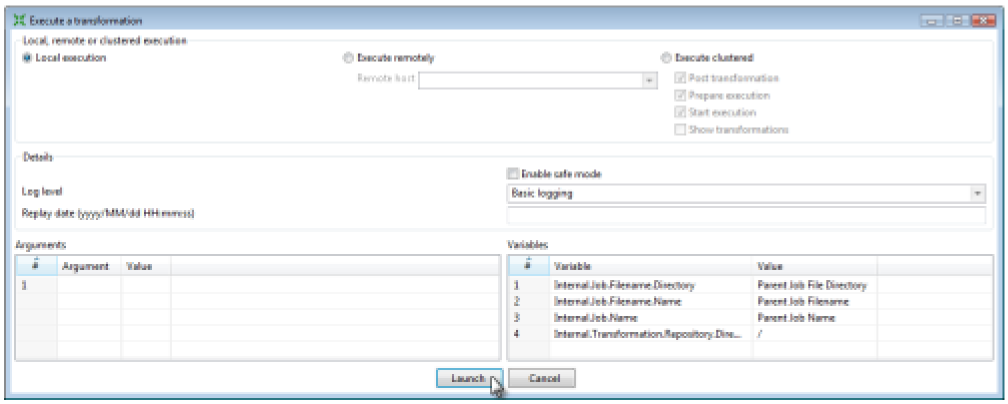


This dialog box displays all the fields in the data output stream and their transformations. In this simple example, there are no intermediate transformation steps, for example filters, but in a transformation with intermediate steps, the details display here. Together, the **Show Input Fields** and **Show Output Fields** dialog boxes display all the detail information on every field that is entering and leaving a step. This is important when you are designing, tuning, and troubleshooting a transformation.

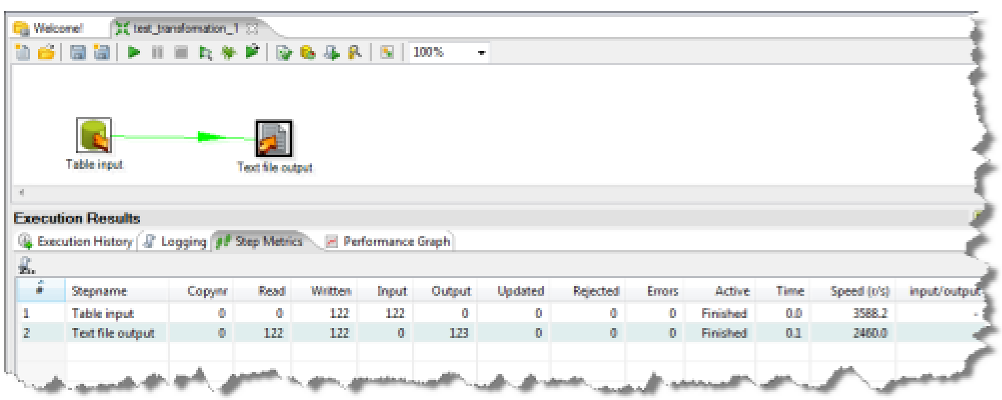
11. Click **File** > **Save**, then click .

You must save the transformation before you can run it.

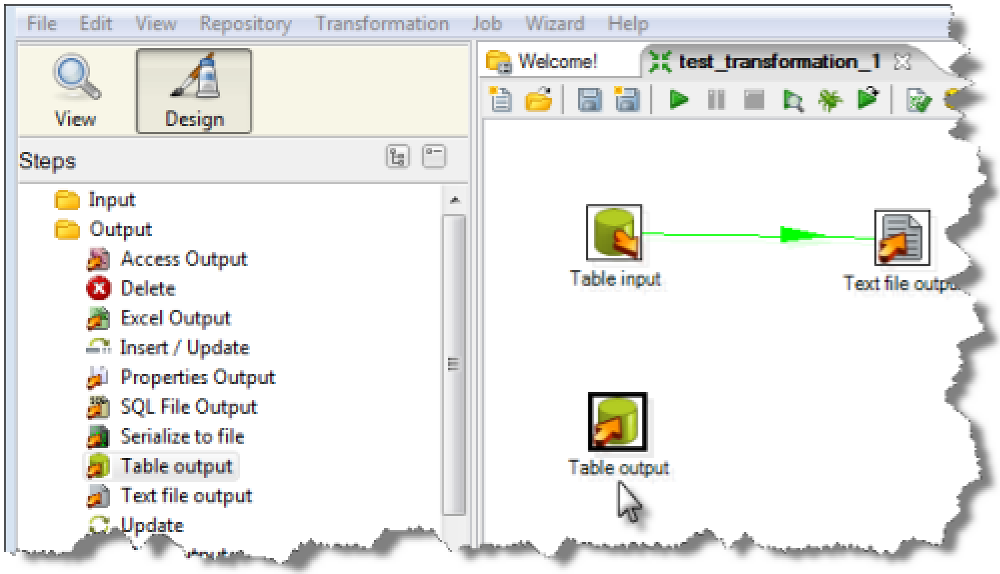
12. Click **Launch**.



The transformation runs and the results display in a pane in the lower portion of the window.



13. As an alternate to the Text File Output, you can use a Table Output, which inserts fields into a database table.



You can double-click **Table Output** to display the Table Output dialog box, where you can specify settings for the output.