## Version 5.3



# Getting Acquainted with Pentaho Components

# Copyright Page

This document supports Pentaho Business Analytics Suite 5.3 GA and Pentaho Data Integration 5.3 GA, documentation revision January 15th, 2015, copyright © 2015 Pentaho Corporation. No part may be reprinted without written permission from Pentaho Corporation. All trademarks are the property of their respective owners.

Help and Support Resources

To view the most up-to-date help content, visit https://help.pentaho.com.

If you do not find answers to your questions here, please contact your Pentaho technical support representative.

Support-related questions should be submitted through the Pentaho Customer Support Portal at http://support.pentaho.com.

For information about how to purchase support or enable an additional named support contact, please contact your sales representative, or send an email to sales@pentaho.com.

For information about instructor-led training, visit http://www.pentaho.com/training.

Liability Limits and Warranty Disclaimer

The author(s) of this document have used their best efforts in preparing the content and the programs contained in it. These efforts include the development, research, and testing of the theories and programs to determine their effectiveness. The author and publisher make no warranty of any kind, express or implied, with regard to these programs or the documentation contained in this book.

The author(s) and Pentaho shall not be liable in the event of incidental or consequential damages in connection with, or arising out of, the furnishing, performance, or use of the programs, associated instructions, and/or claims.

Trademarks

The trademarks, logos, and service marks ("Marks") displayed on this website are the property of Pentaho Corporation or third party owners of such Marks. You are not permitted to use, copy, or imitate the Mark, in whole or in part, without the prior written consent of Pentaho Corporation or such third party. Trademarks of Pentaho Corporation include, but are not limited, to "Pentaho", its products, services and the Pentaho logo.

Trademarked names may appear throughout this website. Rather than list the names and entities that own the trademarks or inserting a trademark symbol with each mention of the trademarked name, Pentaho Corporation states that it is using the names for editorial purposes only and to the benefit of the trademark owner, with no intention of infringing upon that trademark.

Third-Party Open Source Software

For a listing of open source software used by each Pentaho component, navigate to the folder that contains the Pentaho component. Within that folder, locate a folder named licenses. The licenses folder contains HTML.files that list the names of open source software, their licenses, and required attributions.

Contact Us

Global Headquarters Pentaho Corporation Citadel International, Suite 460

5950 Hazeltine National Drive Orlando, FL 32822

Phone: +1 407 812-OPEN (6736)

Fax: +1 407 517-4575

http://www.pentaho.com

Sales Inquiries: sales@pentaho.com

# Introducing Pentaho Business Analytics

The Pentaho Business Analytics (BA) Suite is a single, comprehensive platform from which you can access, integrate, manipulate, visualize, and analyze your data. Whether data is stored in a flat file, relational database, Hadoop, NoSQL database, analytic database, social media streams, operational stores, or in the cloud, the Pentaho BA Suite helps you discover, analyze, and visualize data to find the answers you need -- even if you have no coding experience. If you can program, you can customize reports, queries, transformations, or extend functionality using our extensive API.

- **Empower Your Organization To Explore, Discover, and Innovate:** Pentaho's intuitively designed tools help business analysts, data analysts, data scientists, and others quickly connect to, explore, discover, and visualize data. Pentaho's comprehensive training and technical support, professional services, pre-sales support, and professional documentation suite provide many avenues for learning, problem-solving, and other assistance.

- **Harness the Power of Big Data:** Pentaho Data Integration provides visual tools that help you define and analyze data sets. Pentaho's technology agnostic design allows you to connect to many types of data sources including Hadoop. Interface with Cloudera, Intel, MapR, Hortonworks, Cassandra, Splunk, Impala, MongoDB, DataStax, Hive and other Big Data technologies. Process large data sets across clustered computers with Pentaho MapReduce, which you can run even if you have no programming experience. Beecause the Pentaho BA Suite is part of a comprehensive platform that encompasses business analytics and data integration, you can create reports that feature big data, and use technologies such as Weka to perform predictive analytics.

- **Visualize and Create Reports in Minutes:** A full complement of reporting and modeling tools are at your fingertips. Intuitive interfaces that include drag-and-drop functionality help even novice users rapidly create print-ready reports from a variety of data sources. Tools range from Interactive Reports and Analyzer, which allow you to create quick reports to answer immediate business questions to Dashboard Designer, which allows you to create many reports at once. You can also use Pentaho Report Designer to create highly-detailed, professional, pixel perfect reports.

- **Visually Integrate and Connect to Data:** With the Pentaho BA Suite you can extract and integrate data from multiple sources, with or without data warehouses or data marts. Our easy-to-use data wizard helps you quickly create and test connections. Use the Spoon visual design tool to transform data into structures that fit your reporting and analytical needs and to enrich the data. Use Pentaho Data Integration tools to clean data and make it consistent. Use instaview to visualize data immediately.

- **Deploy on Windows, Linux, Mac ... Or Go Mobile:** Pentaho BA Suite can be installed on Windows, Linux, Mac, and OS X mobile platforms. Install all components on a single computer or distribute them across your network. Or, create clusters to support enterprise-level implementations.

The Pentaho BA Suite consists of BA and Data Integration (DI) components.

- BA Components help you create compelling visualizations, reports, and dashboards, based on data models that you design.
- DI Components allows you to connect to and extract data from diverse data sources such as relational databases, NoSQL databases, and Hadoop. DI allows provides a visual interface that you can use to transform your data to support business analytics.

# Business Analytics

Pentaho Business Analytics (BA) components are comprised of a server, design tools, and plugins that use your data to provide valuable insight into business trends and performance. BA Components are divided into three categories:

## BA Server and User Console

The BA Server, which is the heart of the Pentaho BA Suite, processes reporting, analysis, and dashboard content. The BA Server hosts the centralized BA Repository for secure sharing of all BA data solutions. It also provides scheduling and audit functionality. The BA server is managed through its web-based tool, the User Console. Use the User Console to create business analytics content, display and schedule reports, and manage Pentaho security.

**Updated:** Thu, 29 Jan 2015 15:57:01 GMT
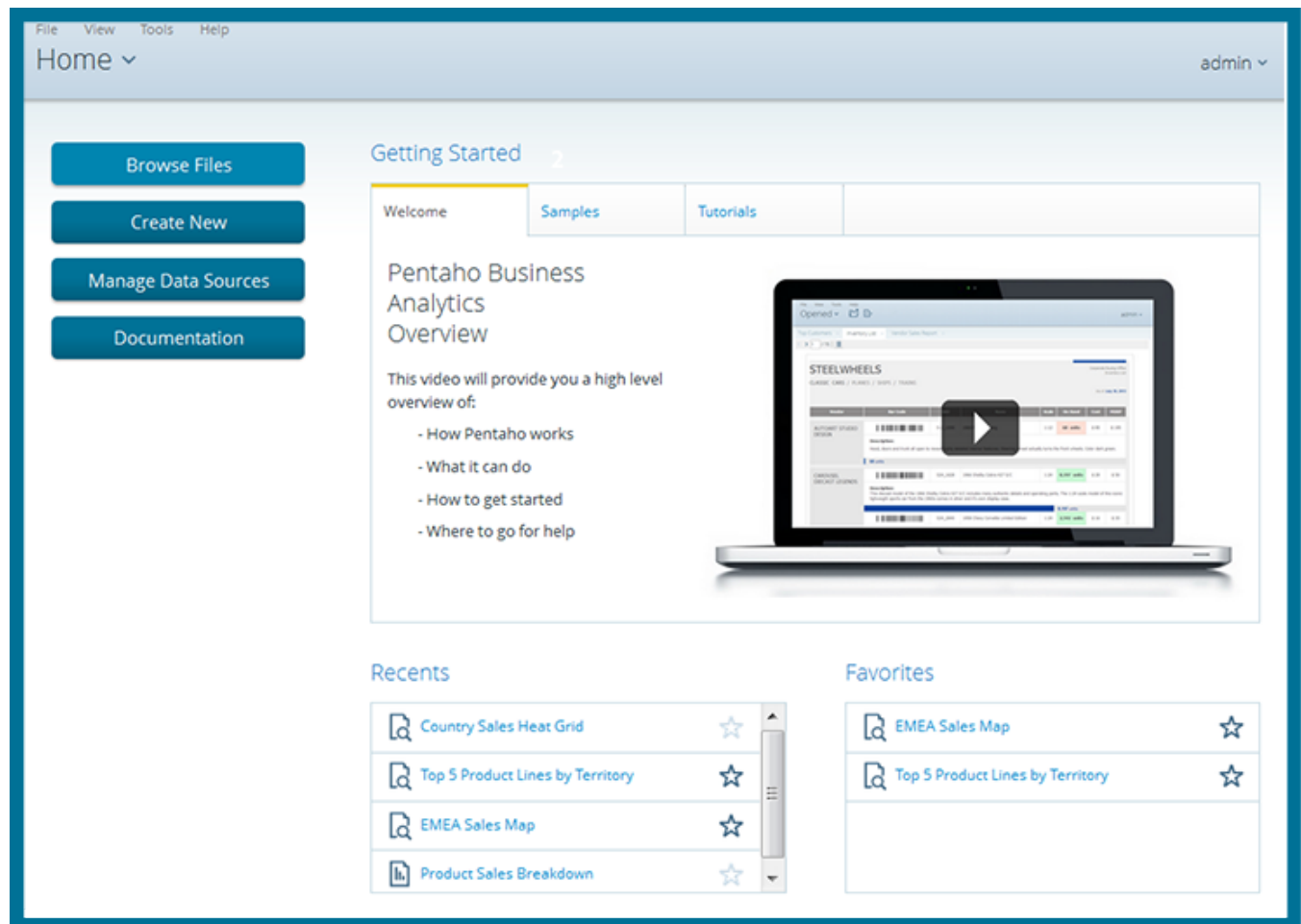
# Web-Based Design Tools and Plugins

BA Web-based components include a host of reporting tools including:

- Report Designer
- Analyzer
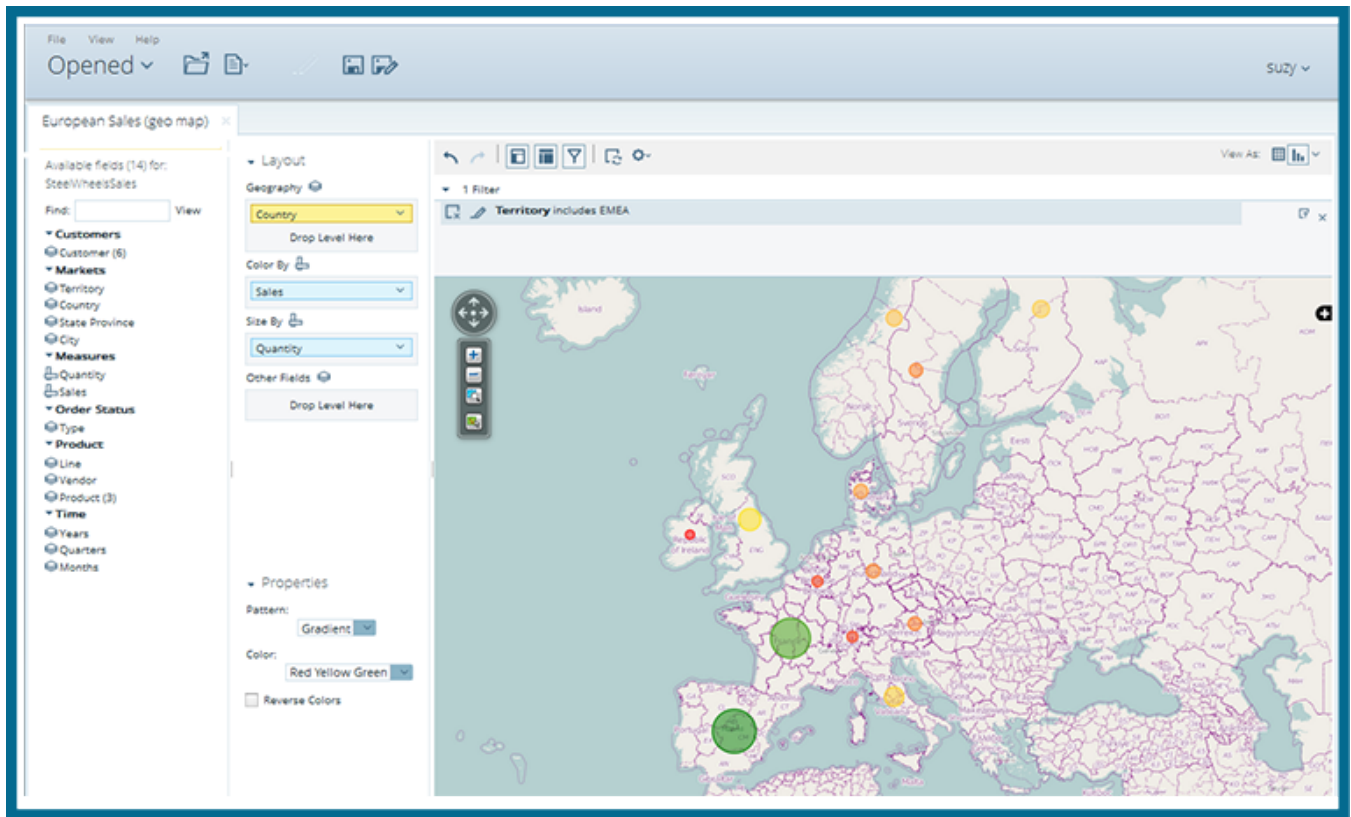- Interactive Reporting
- Dashboard Designer
- Mobile

## Report Designer

Report Designer is a sophisticated report creation tool that allows you to generate highly-detailed, pixel-perfect reports using virtually any data source.
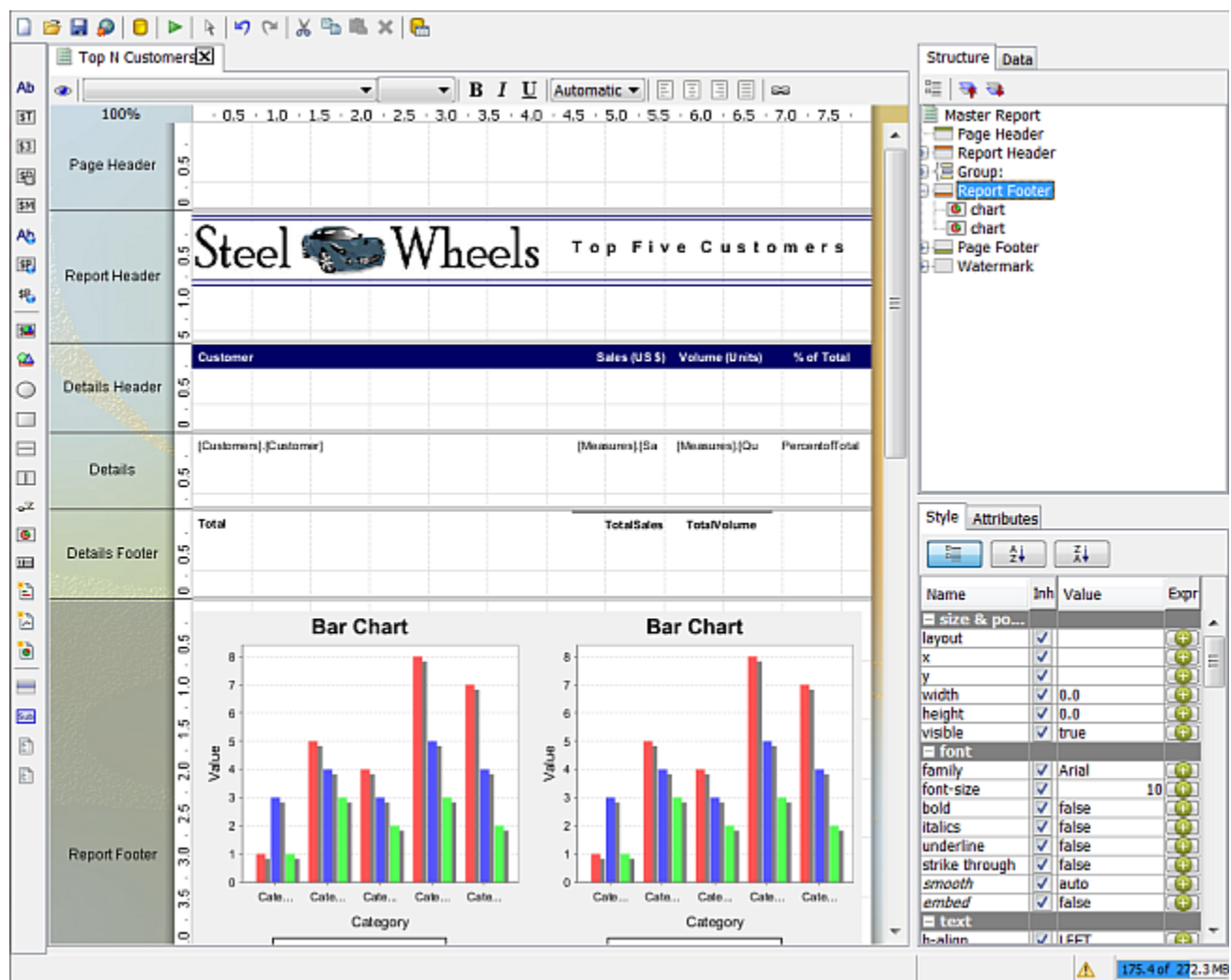
# Analyzer

Analyzer helps you filter and visualize data to make informed business decisions. You can use Analyzer to filter data, add query parameters, configure drill-down links, apply conditional formatting, and to generate hyperlinks. You can also create geographic, scatter chart, heat grid, and multi-chart visualizations.



# Interactive Reporting

Create template-based, on-demand reports in an intuitive, drag-and-drop environment. Quickly add elements to your report and format them any way you like.

## Dashboard Designer

Choose layout templates, themes, and content to create visually attractive dashboards that help decision makers recognize critical knowledge at a glance. Embed a wide variety of content in the web-based display, including Interactive Reports, Analyzer visualizations, graphics, charts, web pages, even collaborative content. Mobile and desktop users can personalize dynamic dashboards to reflect their preferences.
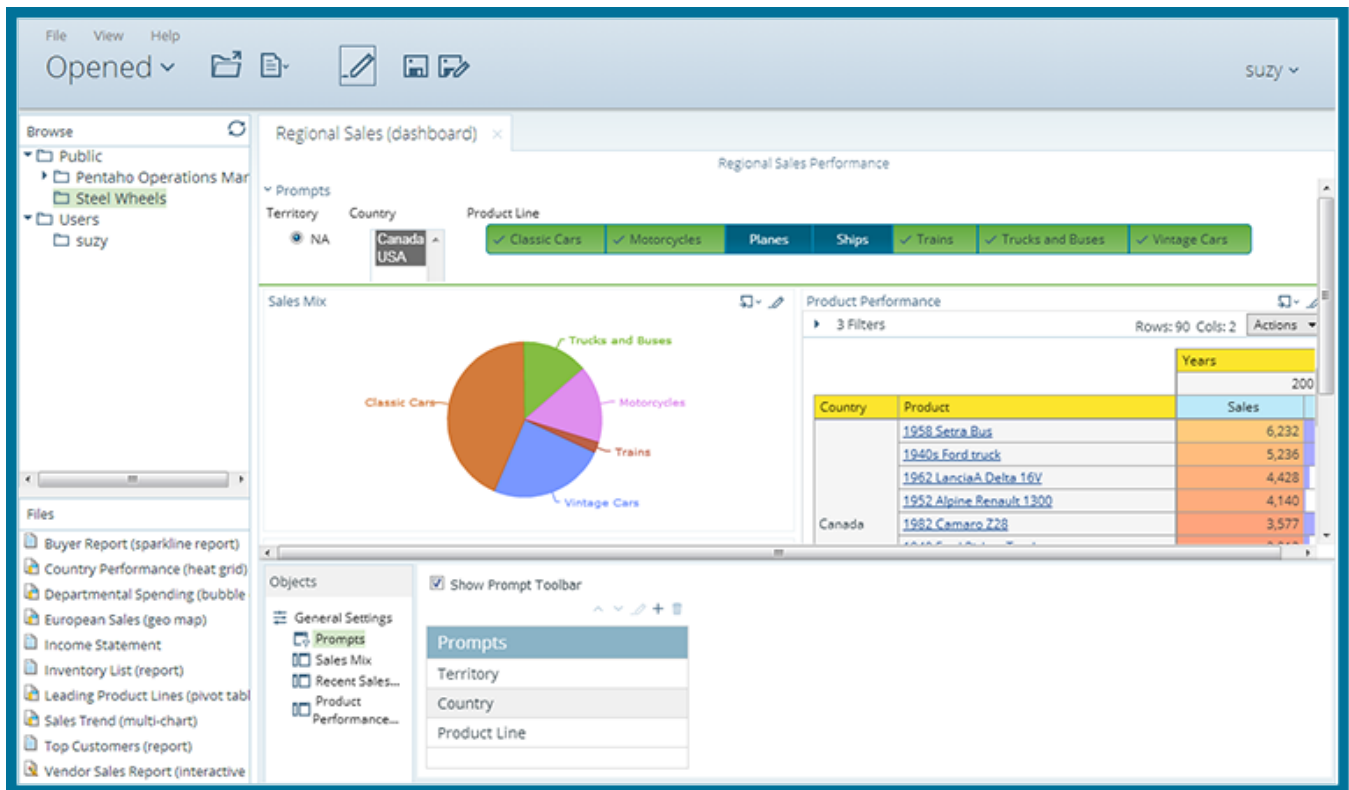
## Mobile

Get immediate access to powerful business analysis, even when you are on the run, with Mobile. Mobile is an innovative application built specifically for the Apple iPad. View and edit Analyzer reports or review reports made by Dashboard Designer, Interactive Reporting and Report Designer.
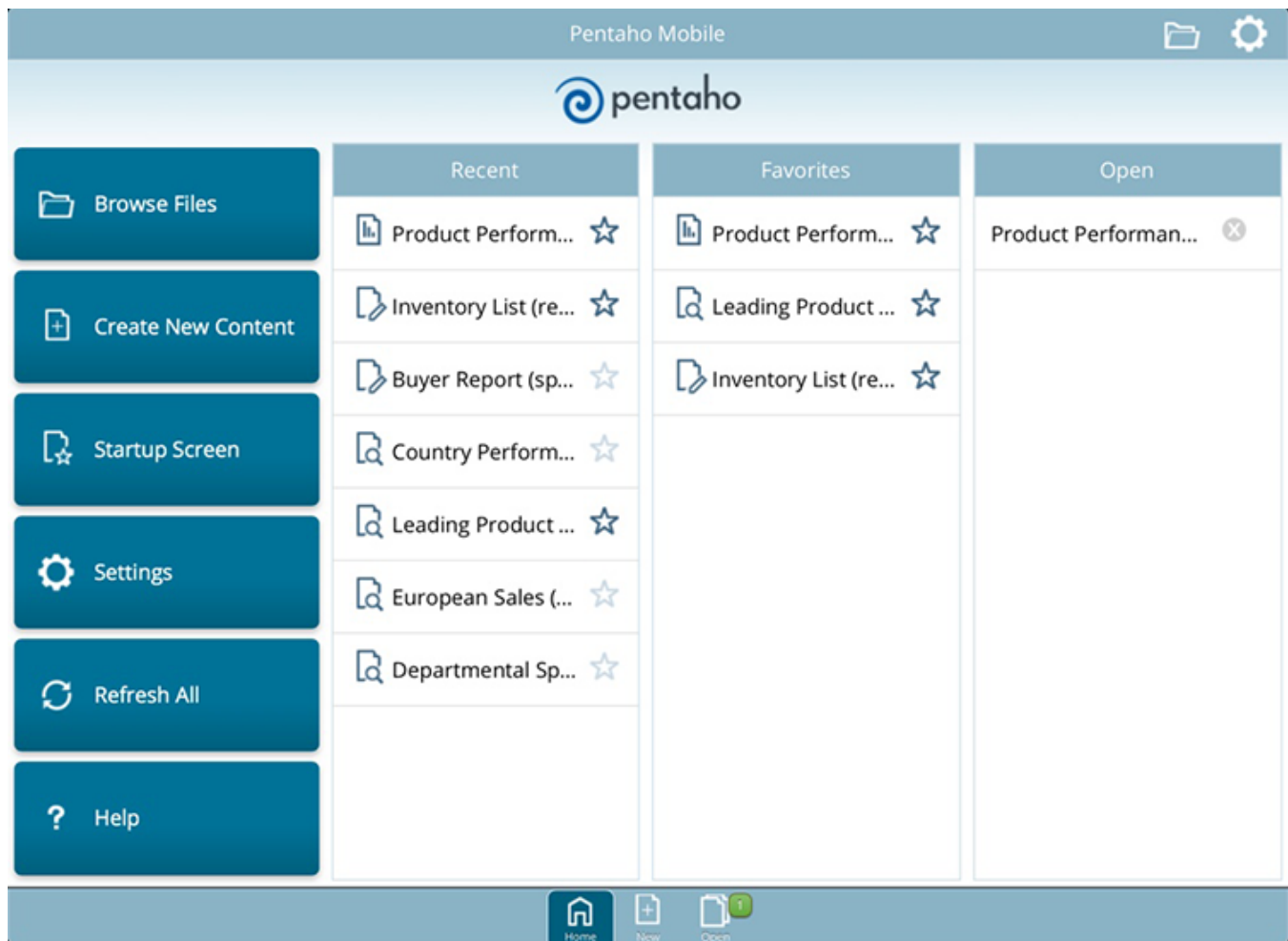
# Client-Based Design Tools

Client-based tools include:

- Aggregation Designer
- Metadata Editor
- Schema Workbench

## Aggregation Designer

Aggregation Designer provides you with a simple interface that allows you to create aggregate tables from levels within the dimensions you specify. Based on these selections, the Aggregation Designer generates the Data Definition Language (DDL) for creating the aggregate tables, the Data Manipulation Language (DML) for populating them, and an updated Mondrian schema which references the new aggregate tables. Use this tool to improve the performance of your Pentaho Analysis (Mondrian) OLAP cubes.

## Metadata Editor

The Metadata Editor simplifies report building. Use the Metadata Editor to build Pentaho metadata domains and models. A Pentaho Metadata Model maps the physical structure of your database into a logical business

model. These mappings are stored in a centralized metadata repository and allow administrators to create business-language definitions for database tables, set security parameters, localize data, add calculations, and to format report data.

## Schema Workbench

Schema Workbench allows you to edit and create multidimensional (Mondrian) models. Use this tool to create multidimensional models graphically or define them by hand-coding XML files.

# Next Steps

The BA Server, design tools, and plugins are designed to be used in tandem with PDI. Check out the PDI components to learn more.

# Data Integration Components

Pentaho Data Integration components are comprised of a server, a design tool, command line utilities, and plugins that you can use to manipulate your data. DI Components are divided into three categories:  Data Integration Server, Design Tools and Command Line Utilities, and Plugins.

## Data Integration Server

The DI Server runs centrally stored transformations and jobs. The DI Server also hosts the DI repository and processing engine, provides a service layer for security and authentication, and allows scheduling.

Manage the DI Server through its related tool, [Spoon](#).

## Design Tools and Command Line Utilities

Spoon is the only DI design tool component.

### Spoon

The DI Server is a core component that executes data integration jobs and transformations using the Pentaho Data Integration Engine. It also provides the services allowing you to schedule and monitor scheduled activities. Drag job entries and transformation steps onto the Spoon canvas, or choose from a rich library of more than 300 pre-built transformation steps and job entries to create a series of data integration processing instructions.
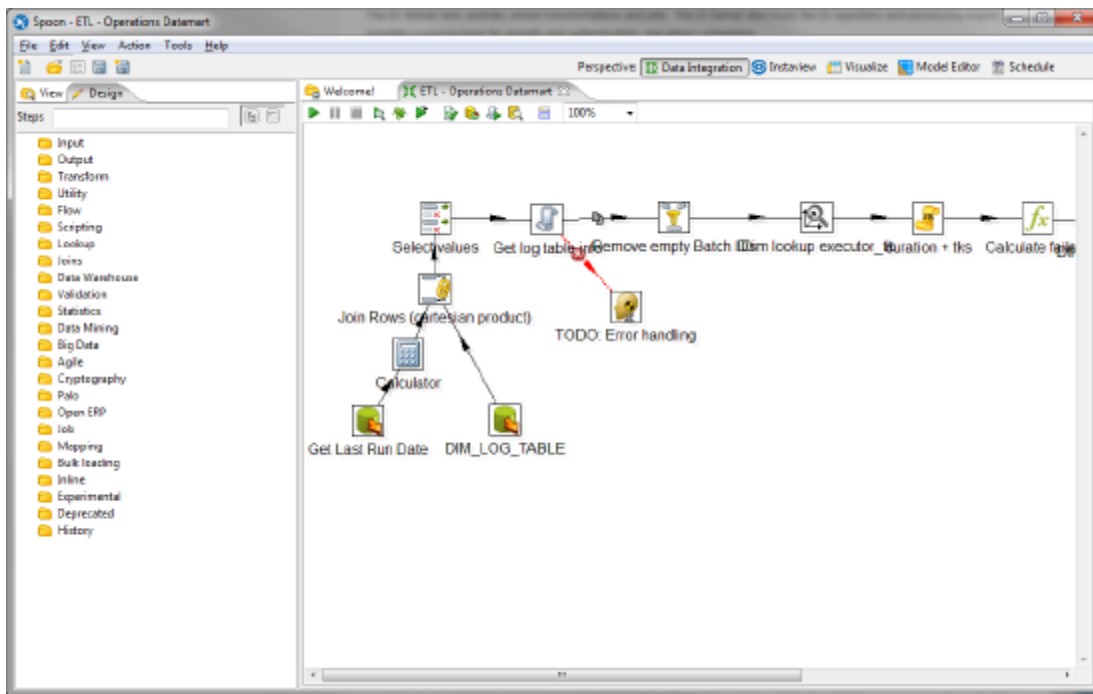
## Kitchen, Pan, and Carte

Use Pan to execute PDI transformations, which represent a data stream through a set of independent tasks. Kitchen can orchestrate PDI jobs, which contain transformations and other job entries as part of a larger business process. Use Carte to set up dedicated, remote PDI servers, so you can coordinate jobs across a collection of clustered computers, and execute transformations within a cluster of Carte cluster nodes.

# Plugins and Command Line Utilities

PDI has hundreds of plugins that can be accessed from the marketplace. But there are two plugins that are installed by default:

- Instaview
- Agile BI

## Instaview

Instaview is software that allows you to use templates to manage the complexities of data access and preparation. Instaview automatically generates transform and metadata models, executes them, and allows you to visualize the results.
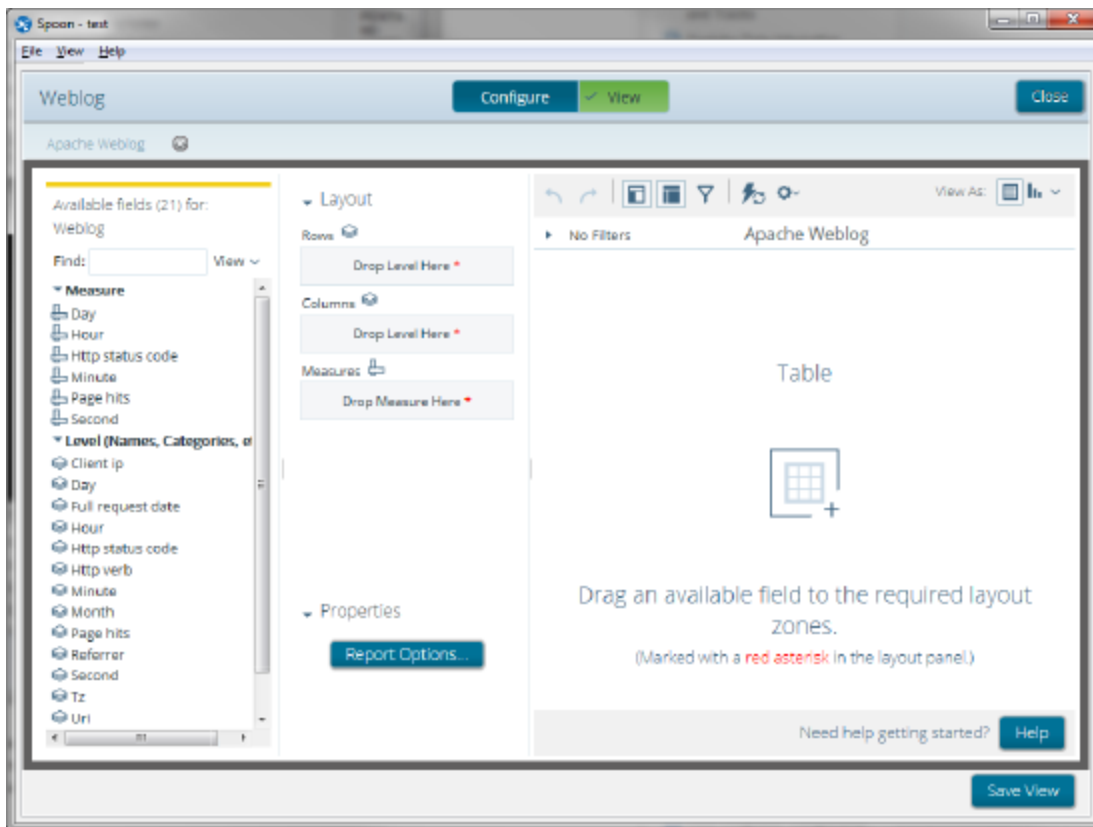
## Agile BI

Agile BI is an accelerated development approach, which links Spoon's data integration to Analyzer and its visualizations. You can immediately see analyzed results as you change your data mart or data models, without leaving Spoon. This technique helps data design professionals and business users work together to rapidly resolve business analysis concerns.

# Next Steps

Now that you are familiar with components, use our workflows to evaluate the Pentaho BA Suite.

**Updated:** Thu, 29 Jan 2015 15:57:01 GMT