



Setting Up the Data Sources

After you install Prime Analytics and set up your project, you next must define and connect to your data sources. Setting up data source connectivity is covered in the following topics

- [Adding the JDBC Driver, page 4-1](#)
- [Creating Historical Data Sources, page 4-2](#)
- [Setting Up Continuous Query Data Sources, page 4-8](#)

Adding the JDBC Driver

Before you can connect to a data source in any Prime Analytics server or client tool, you must install the appropriate database driver. Your administrator should provide you with the proper driver JAR. If not, download the JDBC driver JAR file from your database vendor or developer website. After you have the JAR, follow the instructions below to copy the JAR to the driver directories for all of the components that you want to connect to this data source.



Note

Microsoft SQL Server users frequently use an alternative, non-vendor-supported driver called JTDS. If you are adding an MSSQL data source, verify that you are installing the correct driver.

Backing Up Old drivers

Verify that no other versions of the same vendor JDBC driver are installed in the same directory. If so, back up the driver, then remove it to avoid confusion and potential class loading problems.

Installing the JDBC Drivers

To install the JDBC driver:

-
- Step 1** Copy the driver JAR file to the following directories, depending on which servers and client tools you are using:

- BI Server:
`$PA_HOME/primeanalytics/biplatform/biserver-ce/tomcat/lib/`

- Administration Console:
\$PA_HOME/primeanalytics/biplatform/administration-console/jdbc/
- Schema Workbench:
\$PA_HOME/primeanalytics/biplatform/schema-workbench/drivers/

Step 2 Once the driver JAR is installed, restart the server or client tool that you added it to.

Creating Historical Data Sources

Setting up connectivity to historical data sources is accomplished using the Data Source Wizard. To set up data source connectivity:

- [Adding a CSV File Data Source, page 4-2](#)
- [Creating an SQL Query Data Source, page 4-4](#)

Adding a CSV File Data Source

Follow the instructions below to add to a data source based on a CSV file.

- Step 1** Log into the Prime Analytics user console. For procedures, see [Logging Into the User Console, page 3-1](#).
- Step 2** From the File menu, choose **New > Data Source**.
- Step 3** In the Data Source Wizard, Data Source Name field, enter the new data source name.
- Step 4** Under Source Type, choose **CSV File**.
- Step 5** Click **Import**.
- Step 6** In the Import File dialog, click **Choose File**.
- Step 7** In the File Upload dialog box, navigate to your CSV file, then click **Open**.
- Step 8** In the Import File dialog, click **Import**.

The Import File dialog box closes and the CSV file is uploaded to the server. (If you do not want to continue, you can click **Cancel** and select different file.)

A preview of your file contents appears in the lower portion of the wizard page. If items are not aligning correctly, it is because you must specify your delimiter and enclosure types. Go on to the next step.

- Step 9** Specify the following CSV parameters:
- Delimiter—Specifies how your file contents is delimited: Comma (default), Tab, Semicolon, Space, Other (allows you to enter a different delimiter item).
 - Encoding—Specifies the CSV encoding. The default is ISO-8859-1.
 - Enclosure—Specifies the enclosure: Double Quote, Single Quote, or None.
 - First row is Header—If checked (default), the first row of your CSV file is a heading for columns in the file.

File Preview displays the first lines of your CSV file based on your delimiter, enclosure, and header selections. Column alignment indicates the delimiter and enclosure are set correctly.

- Step 10** Click **Next**.

The wizard Staging Setting panel allows you edit the final display.

Step 11 As needed, edit the final file display:

- All columns are enabled by default. If you do not want to include all of them, uncheck the ones you want to exclude. You can use Deselect All and Select All to disable and enable all columns.
- The Name and Type columns display the provided column name and type. You can edit these as necessary.
- The Source Format field displays the source format for dates and numeric values. Formats are automatically detected based on the your data source contents. The drop-down list displays alternate format options that you can select. Drop-down lists are not enabled for certain data types such as the String data type.
- Length indicates the maximum number of characters allowed in a field.
- Precision is the number of digits after a decimal point.



Note Boolean values are rendered as true or false.

- Click **Show File Contents** to see a sample of the data in your source file. Click **OK** to return to staging and make changes if necessary.

Step 12 After you complete your configuration and formatting review, click **Finish**.

A success message indicates the number of rows being loaded into the database and that a default metadata model was created.

Step 13 Click one of the following:

- Keep default model—If you want use the default model or,
- Customize model now—To customize the model.

Your CSV file data is extracted and staged in a database table. The new data source is now available for use in reports, dashboards, and analyses.

Creating Database Connections

Prime Analytics SQL and database queries require you to first create a connection to the host containing the database, including port, host name, username, and password.

To create a database connection:

Step 1 Verify that the JDBC driver is added. For information, see [Adding the JDBC Driver, page 4-1](#).

Step 2 Log into the Prime Analytics user console as an administrator user. For procedures, see [Logging Into the User Console, page 3-1](#).

Step 3 From the File menu, choose **New > Data Source**.

The Data Source Wizard appears.

Step 4 In the Data Source Name field, enter the new data source name.



Note The data source name is displayed in analytics, reports, and dashboards.

- Step 5** Under Source Type, choose **SQL Query** or **Database Table(s)**.
- Step 6** From the Connection toolbar, choose **Add Connection**.
- Step 7** In the Database Connection dialog, enter the following information:
- Connection Name—Enter a name for your connection.
 - Database Type—Two database types are available:
 - TruCQ—Prime Analytics TruCQ query database if archiving is enabled.
 - Generic—A generic database.
 - Access (not editable)—Access is enabled using JDBC drivers.
 - Host Name—Enter the name of the host where your database is located.
 - Database Name—Enter the database name.
 - Port Number—Enter the port number used to access the host.
 - User Name—Enter the username
 - Password—Enter the password.
- Step 8** After you enter the connection information, click **Test** to test it.
- If a connection is established, a Test Connection dialog indicates the connection to *[database name]* is OK. It also lists the hostname, port, and database name. If the connection was not successful, review your connection details and correct any errors.
- Step 9** Click **OK**.
- Step 10** Continue with one of the next procedures:
- [Creating an SQL Query Data Source, page 4-4](#)
 - [Creating a Database Table Data Source, page 4-5](#)

Creating an SQL Query Data Source

The following procedure tells you how to create an SQL query from an SQL database. To create it, you must be connected to the database using JDBC. You will enter an SQL statement that defines the SQL data source scope.

You can customize the data columns that are presented to users who build SQL queries. For example, you can define column names and select options that define how data is aggregated (sum, min., max., etc.). After you create the SQL query, it is available to users use in reports, dashboards, and analyses.



Note

To complete this procedure, the SQL database must be up and running.

- Step 1** Verify that you are connected to the database where you want to create a query. To create a connection, see [Creating Database Connections, page 4-3](#).
- Step 2** Log into the Prime Analytics user console as an administrator user. For procedures, see [Logging Into the User Console, page 3-1](#).
- Step 3** From the File menu, choose **New > Data Source**.
- The Data Source Wizard appears.

Step 4 In the Data Source Name field, enter the new data source name.



Note The data source name is displayed in analytics, reports, and dashboards.

Step 5 Under Source Type, choose **SQL Query**.

Step 6 Under Connection, choose the database connection for the database where you want to run your query.

Step 7 Under SQL Query, enter your SQL query. For example, to select certain columns:

```
SELECT Device, Interface, IPaddress FROM NetworkDevices
```

Or, to select all data:

```
SELECT * FROM NetworkDevices
```

Step 8 Click Data Preview to verify that your query returns data.

Step 9 Click **Finish**.

The query is added.

Creating a Database Table Data Source

You can create a reporting and/or analysis data source against one or more relational database tables. This data source type is useful for a broad range of use cases including reporting directly against an operational data store (ODS) or creating a connection for an online analytical processing (OLAP) analysis against a star schema. When you can create the database table(s) data source, two options are provided:

- **Reporting Only**—If you access data in a relational database that is operational or transactional, or which does not contain a star schema, use the Reporting Only option. It creates a data source that can be used in reports and dashboards. A metadata model is automatically generated. Users can access the data to create reports and build dashboards.
- **Reporting and Analysis**—If you access a relational database containing tables arranged in a star schema, that is, tables separated into dimensions related to a single fact-table, such as a data warehouse, use the reporting and analysis option. You can also create a reporting and analysis data source using a single table where the table acts as the single fact table and also contains dimensional information. This option creates a data source that can be used in analytics, reports, and dashboards. A Mondrian schema and metadata model are generated

The choice you make depends on the schema structure of the database tables you are accessing and the user console tools you are using. The following procedures tell you how to create a table data source for reporting only and for reporting and analysis:

- [Creating a Database Table Data Source for Reporting Only, page 4-5](#)
- [Creating a Database Table Data Source for Reporting and Analysis, page 4-7](#)

Creating a Database Table Data Source for Reporting Only

Follow the instructions below to create a database table(s) data source for reporting.

**Note**

To complete this procedure, the database must be up and running.

- Step 1** Verify that you are connected to the database table(s) where you want to create a query. To create a connection, see [Creating Database Connections, page 4-3](#).
- Step 2** Log into the Prime Analytics user console as an administrator user. For procedures, see [Logging Into the User Console, page 3-1](#).
- Step 3** From the File menu, choose **New > Data Source**.
The Data Source Wizard appears.
- Step 4** In the Data Source Name field, enter the new data source name.

**Note**

The data source name is displayed in reports and dashboards.

- Step 5** Under Source Type, choose **Database Table(s)**.
- Step 6** Under Connection, choose a database connection.
- Step 7** At the bottom of the page, click, **Reporting only**, then click **Next**.
A list of available tables retrieved from your database is displayed.
- Step 8** Under Available Tables, choose a table and click the right green arrow to move it under Selected Tables.
To choose multiple tables press **CTRL**. To remove a table, chose it and click the left green arrow.
- Step 9** When finished, click **Next**.

**Note**

While you can create a data source from a single table, the more common use is selecting and joining multiple tables.

- Step 10** In the Define Joins panel, define the table relationships:
- In the Left Table list, choose the table that you want to join. When you choose a table, key fields appear under Key Field.
 - In the Right Table list, choose the table that you want join to the table on the left.
 - Click **Create Join** to define the inner join between each table.

**Note**

The only join condition allowed in the Data Source Wizard is an inner join; references to left table and right table are not associated with other types of join conditions.

- Step 11** Click **Finish**.

A default metadata model is generated in the background for use in reports and dashboards. You can choose to keep the default model or customize the it.

Creating a Database Table Data Source for Reporting and Analysis

Follow the instructions below to create a database table(s) data source for reporting and analysis.



Note To complete this procedure, the database must be up and running.

- Step 1** Verify that you are connected to the database table(s) where you want to create a query. To create a connection, see [Creating Database Connections, page 4-3](#).
- Step 2** Log into the Prime Analytics user console as an administrator user. For procedures, see [Logging Into the User Console, page 3-1](#).
- Step 3** From the File menu, choose **New > Data Source**.
The Data Source Wizard appears.
- Step 4** In the Data Source Name field, enter the new data source name.
- Step 5** Under Source Type, choose **Database Table(s)**.
- Step 6** Under Connection, choose a database connection.
- Step 7** At the bottom of the page, click, **Reporting and Analysis**, then click **Next**.
- Step 8** Under Schema, choose the schema you want to apply to the table.
A list of available tables retrieved from your database is displayed.
- Step 9** Under Available Tables, choose a table and click the right green arrow to move it under Selected Tables.
To choose multiple tables press **CTRL**. To remove a table, chose it and click the left green arrow.
- Step 10** In the Fact Table, choose the tables you want to serve as the fact table.
- Step 11** When finished, click **Next**.



Note While you can create a data source from a single table, the more common use is selecting and joining multiple tables.

- Step 12** In the Define Joins panel, define the table relationships:
- In the Left Table list, choose the table that you want to join. When you choose a table, key fields appear under Key Field.
 - In the Right Table list, choose the table that you want join to the table on the left.
 - Click **Create Join** to define the inner join between each table.



Note The only join condition allowed in the Data Source Wizard is an inner join; references to left table and right table are not associated with other types of join conditions.

- Step 13** Click **Finish**.

A default metadata model is generated in the background for use in reports and dashboards. You can choose to keep the default model or customize it.

Setting Up Continuous Query Data Sources

Complete the following procedure to set up data sources for real-time (continuous query) data:



Note

To complete this procedure, the database must be up and running.



Note

This procedure only applies to TruCQ stream connections. To connect to an archived database table, see [Creating Database Connections, page 4-3](#).

Step 1 Verify that the TruCQ engine is running by entering the following command as the root user:

```
service trucq status
```



Note

If you switch to the root user, make sure you use the `su - root` command to ensure that the appropriate environment variables are loaded.

If the TruCQ engine is not running, you will need to start it. See [Starting and Stopping Prime Analytics Servers, page 8-30](#), for information.

Step 2 Stop the BI platform:

```
service biplatform stop
```

Step 3 Navigate to the following directory:

```
$PA_HOME/biplatform/biserver-ce/tomcat/webapps/pentaho/WEB-INF/classes
```

Step 4 Using an XML editor, open the `data.xml` file in this directory.

Under `<trucq></trucq>` tag, you will find two pre-existing connections.

Step 5 Copy the existing trucq connections and paste at the end of the xml before the `</engines>` tag, for example:

```
<trucq id="netflow">
<engine id="netflow" maxSharedConnections="5">
    <url>jdbc:truviso://localhost:5432/cqdb</url>
    <username>primea </username>
    <password></password>
    <driver>com.truviso.Driver</driver>
</engine>

</trucq>
</engines>
```

Step 6 Modify the details according to your environment:

- Trucq Id—Provide a unique engine name. This name must be unique in the file.

- Engine id—Provide a unique engine name. This is the name that will appear in the Add SQL dialog box Select a Connection list when you create the continuous query. The name must be unique in the file.
- URL—Enter the JDBC URL in the format:

```
jdbc:truviso://<trucq engine address>:<trucq port>/<trucq database>
```
- Username—Enter the database username.
- Password—Enter the database password.
- Driver—Enter the TruCQ driver name in the format: driver=com.truviso.Driver

Step 7 Save the changes.

Step 8 As the root user, start the biplatform:

```
service biplatform start
```

Step 9 Verify the new entry:

- Log into the BI platform as an administrator user. (See [Logging Into the User Console](#), page 3-1.)
- Click Manage Streaming SQL.
- Click Add SQL
- Verify the new engine ID is available the Select a Connection list.

If the new engine ID appears, continue with the next step. If not, for example, if errors appear when you click Manage Streaming SQL or your connection does not appear:

- Check the data.xml to verify the details are correct. Try connecting to the database using the psql commandline tool.
- Check the logs under \$PA_HOME/biplatform/biserver-ce/tomcat/logs/catalina.out or pentaho.log or truviso.log.
- Repeat Steps 1 through 9.

Step 10 Log into the Prime Analytics user console as an administrator user.

Step 11 If any files are open, close them to display the Prime Analytics home page.

Step 12 Under Data Source, click **Manage Streaming SQL**.



Note If you are not connected to the real time data, an error is displayed when you click Manage Streaming SQL.

Step 13 On the Real Time tab, click **Add SQL**.

Step 14 In the Add SQL dialog box, complete the following information:

- Select a Connection—Choose a connection from the available list.
- Provide CDA File Name—Enter the Community Database Access file name.
- Provide a Location for CDA—Enter the file location, or click the browse icon and navigate to the file location.
- Provide a SQL—Enter the SQL you want to use on the data, for example:

```
SELECT cq_close(*), sum(packets) as packets from \
jellybeans.netflow_v5 <slices '1 second'> where routerip=?
```



Note Verify the SQL command using the psql command line tool before entering it here.

- Description—Enter a description for your query.
- Query Type—Choose the basis for the query type:
 - Time Based
 - Row Based

Each continuous data stream has a CQTIME attribute. This attribute can be of the following types: smallint, int, bigint, timestamp without time zone, and timestamp with time zone. If CQTIME is defined as timestamp without time zone or time stamp with time zone, it is considered time-based.

- Parameter Present—If you want to add parameters, choose Yes. Otherwise, choose No.
- Parameters—If you want to add parameters, enter the following:
 - Parameter Column Name
 - Parameter Column Data Type
 - Parameter Test Value



Note The test value is only used to check the query validity. The actual value is entered when you execute the dashboard.

Step 15 When finished, click **Create**.

The query is validated. If the query has errors, messages are displayed. If this occurs, correct the error and save it again.
