Table of Contents

Tutorial: Excel Spreadsheets ......................................................................................................................... 1

Create a New Project ........................................................................................................................................... 1

Connect and Extract Tables from Spreadsheets .............................................................................................. 2

Prepare the Sources ............................................................................................................................................. 4
  Merge Columns from Two Excel Sheets into One Prepared Source ............................................................ 5
  Rename a Column in the Source .................................................................................................................... 10
  Merge Two Excel Sheets ............................................................................................................................... 12

Relate the Sources ............................................................................................................................................... 15

Create a Chart Report in Visualizer .................................................................................................................. 17

Put the Chart and a KPI on a Dashboard ........................................................................................................... 24
Tutorial: Excel Spreadsheets

This product tutorial steps you through using Pronto to merge two Excel spreadsheets into a prepared source, relate a prepared source and a raw source, and create reports in Visualizer. This also shows the different use cases for a merge transform in the Prepare page versus a relationship between sources in the Relate page.

The scenario involves a main spreadsheet that represents a corporate database, based on the classic Northwind demo. It covers products, orders, suppliers, customers, and employees for an imaginary food distributor. A warehouse manager has compiled a spreadsheet of additional data about their regional customers for promotional purposes, including contact details. The manager wants a report showing their local customers, a count of their orders, and what promos were run.

As with all Birst Professional Edition projects, first you connect and extract the data. The spreadsheets for this tutorial are available at CorpData - Excel Tutorial File and MyPromos - Excel Tutorial file on the Community.

- Create a New Project
- Connect and Extract Tables from Spreadsheets
- Prepare the Sources
- Relate the Sources
- Create a Chart Report in Visualizer
- Put the Chart and a KPI on a Dashboard

Create a New Project

In Pronto, create a new project.
Connect and Extract Tables from Spreadsheets

On the Connect page, either drag-and-drop or use the Browse Files button to upload both the CorpData and MyPromos spreadsheets.

A preview shows the spreadsheet tabs as separate data sources, and the columns in each. This is a preview of the tables and columns, the data is not yet in Birst.
In the preview you select the tables and columns you want to pull into Birst. For this tutorial, select all of the tables and all of the columns in both spreadsheets, and click Done.

**Tip:** It’s not always necessary to select all the tables and columns, if you know that you only need to use particular ones.

Then click Extract Imported Files.
Pronto shows a progress bar, and then a message when the imported files have been extracted into Birst.

**Prepare the Sources**

You can think of a prepared source as a customized raw source. Customizations at the source level provide report-friendly changes such as improved column titles, merged columns, or text replacements. These changes are called transformations, or transforms for short. A prepared source can also be based on multiple sources, such as when you want to merge tables together.
When you create a prepared source you can reuse it in multiple reports.

In this tutorial there are two prepared sources, one that is a merge of tables from two sources, and another that is a merge of tables from one source.

**Merge Columns from Two Excel Sheets into One Prepared Source**

In the Prepare page, locate the CorpData Customers table. (This was the Customers tab in the original spreadsheet. Hover over the list items on the left and use the popup to find the file, and click on it.

Make sure it's title turns blue, to indicate that it is selected.

Hover to the right of the item in the list to see and then click Create a New Prepared Source.

Name the source and click the checkmark.

Pronto creates the new prepared source and it is ready for you to edit it.

In our tutorial scenario, the manager wants to put additional customer info into the same table as the corporate customer table. This means that the columns in the local spreadsheet need to be
appended and the rows matched up with the corporate information. On the Prepare page this you accomplish this using the Merge transform.

The next step uses the Merge transform on the Prepare page.

In the new prepared source, click Add Step in the pipeline.

Select the Merge transform.

The Merge dialog opens, with the prepared source shown on the left.
Then select the MyPromos_xlsx MyCustomers table on the right. You may have to scroll to see it.

To create a merge, and to make sure the rows and columns end up in the right organization, a shared or key column needs to be identified. This is called a "join column". The column titles don't have to match, but the data should be pretty much the same. In this example, both spreadsheets used the same CustomerID keys. Keys are meant to be unique, so if two spreadsheets use the same keys, they can be merged together in a logical way. It is important that at least one of the two columns has only one row for each unique key, otherwise the data may not match up correctly.
Tip: Part of deciding how to use the join depends on what rows the final prepared source should contain. This is controlled by the type of join. For example, the default inner join will match the rows between the left and right tables, and return only the ones that contain the same CustomerID in both columns. An outer join would return all the rows from each table, a right join would result in all the rows from the right and any any rows that match the same CustomerID from the left, and a left joint would result in all the rows from the left table and any rows that match the same CustomerID from the right table.

In this scenario, the manager doesn't really care about all the rows that are not from the local customers, and does care about getting all the columns from the corporate data. So an inner join will be the best choice.
Now you can refine the final set of columns for the prepared source. For example, both CustomerID columns are not needed, only one, so it can be deselected from the second source. The column will still be there behind the scenes because it is required for making the join work, consistently. But it doesn’t need to be displayed. Unchecking the column causes it to not be displayed in the final prepared source.

In addition to deselecting the CustomerID from the table on the right so that it is not displayed, it makes sense to deselect the CompanyName and ContactName, since both of those are also in the main corporate table. The Email and Twitter columns are the ones that the manager wants to add.

Now the merge is defined and it’s time to click Apply. Pronto compares the two tables and merges them based on the join, and hides the deselected columns. It returns the results, in this case the 5
local customers with all the data from corporate, augmented by the email and Twitter columns provided by the manager. Pronto also shows the merge step in the pipeline below the data.

<table>
<thead>
<tr>
<th>Fax</th>
<th>Latitude</th>
<th>Longitude</th>
<th>SalesRegion</th>
<th>Email</th>
<th>Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>55.</td>
<td>44.028521</td>
<td>-123.072898</td>
<td>Am West</td>
<td><a href="mailto:Howie@GLFoods.co">Howie@GLFoods.co</a>...</td>
<td>SnyderMark...</td>
</tr>
<tr>
<td>87.</td>
<td>(503) 555-237..</td>
<td>45.564861</td>
<td>-117.917434</td>
<td>Yoshi@HungryCoyo...</td>
<td>YoshiStore</td>
</tr>
<tr>
<td>61.</td>
<td>45.513786</td>
<td>-122.675375</td>
<td>Am West</td>
<td><a href="mailto:liz@TBC.com">liz@TBC.com</a></td>
<td>LizTBC</td>
</tr>
<tr>
<td>25.</td>
<td>(206) 555-217..</td>
<td>47.681488</td>
<td>-122.208735</td>
<td>nagy@THGourmet...</td>
<td>GourmetNa...</td>
</tr>
<tr>
<td>11.</td>
<td>(206) 555-411..</td>
<td>47.599974</td>
<td>-122.314166</td>
<td>Karl@WhiteClover...</td>
<td>KarlClover</td>
</tr>
</tbody>
</table>

**Rename a Column in the Source**

The CompanyName column is likely to be used in many reports, and in this scenario, the manager would prefer to title it the Customer Name. By changing the title in the prepared source, the change is consistent for all future reports. You would not have to change the title for each report.
Scroll to the CompanyName column and select it. The Transforms menu highlights all the customizations that could be used for that column.

Select Rename, type in the new title, and click Apply.

The preview for that column turns green to indicate that a transformation has been applied to it, and the step displays in the pipeline.
Merge Two Excel Sheets

In this scenario, the manager wants to show the descriptions of the local promos in the final reports; however, the local spreadsheet doesn't have all the information on one sheet. So the next step is to create another merge between two sheets and create the MyPromos prepared source.

Using the same process as before, select the MyPromos_xlsx MyPromos and MyPromos_xlsx PromoCustomers sources, and merge them using the PromoID field as the inner join. Don't forget to
hide the second PromoID column.

Click Apply and the prepared source looks like the screenshot below.
Click Publish and Birst creates the two new prepared sources based on the definitions you created.
Relate the Sources

Now that the base customer data is appended with the additional info and the rows are limited to the manager’s local customers, the new Prepared_Customers source can be related to another corporate table. The corporate table that tracks Orders has an OrderID column that can be used to count the number of orders per customer.

Go to the Relate page.

Select the Prepared_Customers prepared source, and the spreadsheet CorpData_xlsx Orders table raw source. Pronto displays the columns of each table.
The same concept of a join applies here. Click CustomerID in Prepared_Customers and the same in Orders to have Pronto relate the CustomerIDs to each other. The default join type, an inner join, is good.

Then deselect Orders and select Prepared_MyPromos to create another relationship using the CustomerID key.
An interesting question arises: how is relating different from merging? They both deal with two tables that each have the same key column, and they both use joins to define the rows in the resulting table. The short answer is that you use a relationship when you need all the data from each table, and the data must be logically related, via the key, so that the rows make sense. A merge is more appropriate when you want to simplify the data sources that will be available for reports, by combining multiple tables (sheets) into one.

In the Relate page you don't have to save anything. When you are done, you can go straight to Visualizer to make a report.

**Create a Chart Report in Visualizer**

Click the Main menu button to get to Visualizer button.
Once in Visualizer, from the Measures section double-click the OrderID.

Click OrderID when it is listed in the gray Chart Builder column, to open its settings.
Set the OrderID aggregation to Count Distinct.

Now the OrderID column represents the count of distinct Orders, so you can also change the label by double-clicking its title.

Click Subject Area to get to the Attributes button.
From the Attributes section scroll down and double-click Prepared_Customers.

Double-click Customer Name for the Category.

At this point there is a column chart showing the Count of Orders by Customer Name.
Click Back in the Attributes section.

Then double-click Prepared_MyPromos and double-click Promo Description.

Visualizer adds it as a Color section to the chart.
The chart shows which promos were used for which local customers, and the count of orders for each promo per customer.

Save the report and give it a name.

You can make other changes to the report, such as changing its chart type.
Also notice that when the cursor hovers over a section, a popup provides the details.
If you like your changes, click Save to save it again.

There are many more features and things you can do with Visualizer. For more ideas see the Visualizer Help.

**Put the Chart and a KPI on a Dashboard**

Dashboards are how you display and distribute one or more of the charts or tables you create in Visualizer. For this scenario, the manager wants to share the dashboard with the home office via an email.

Click the main menu to get to the Dashboards button.
An empty dashboard opens and you can start by renaming it.

Then click the center button to add the report. Select the report and click Done.
You can move and resize the report on the dashboard.

Another way to show information on a dashboard is to use a Key Performance Indicator (KPI). In this scenario, the manager wants a quick count of the total number of orders. To use a KPI, first select its menu option.

KPIs are for measures, so they are listed in the KPI Builder. Navigate to the OrderID and changes its aggregation to Count Distinct.
Notice how a preview of the KPI displays on the left.
You can change style settings for the measure number.

And you can change the title name and its style.
Click Save and place the KPI on the dashboard.

Click Publish to save the dashboard.

There are many more features and things you can do with dashboards. For more ideas see the Dashboards 2.0 Help.