Data Analysis and Integration

Lab 11: Reporting

Note: In order to do this lab, you need to have successfully completed the previous labs.

In this lab, we will use Pentaho Report Designer (PRD) to create business reports from a data source, such as a database or a data warehouse.

Creating a connection to a database

1. Open a new terminal and navigate to the folder: **~/Pentaho/report-designer**

2. Start the Pentaho Report Designer (PRD) with: **./report-designer.sh**

3. In the menu **File**, click **New**.

4. On the left side, you will see the canvas that is used to design the report. On the right side, there are two tabs: **Structure** and **Data**. Switch to the **Data** tab.

5. Right click **Data Sets**, and select **JDBC**. The **JDBC Data Source** window will open. We will be working in the **Data Source** tab.

6. Next to **Connections**, click the "**+**" button. The familiar **Database Connection** dialog will open.

7. Configure the database connection as follows:
   - Connection Name: **steelwheels**
   - Connection Type: **MySQL**
   - Access: **Native (JDBC)**
   - Host Name: **localhost**
   - Database Name: **steelwheels**
   - Port Number: **3306**
   - User Name: **aid**
   - Password: **aid**

8. Press the **Test** button to test the connection, and close the **Database Connection** dialog with **OK**.

9. Back in the **JDBC Data Source** window, under **Connections**, select the **steelwheels** connection that you have just created.

10. On the right side, next to **Available Queries**, press the "**+**" button. A new query named **Query 1** will be created.

11. In **Query Name**, rename the query to **Products**.

12. Below, in the **Static Query** tab, click the edit button  to edit the query. The **SQL Query Designer** will open.

13. Drag the **products** table from the lower-left pane to the right pane.

14. Close the **SQL Query Designer** with **OK**. The result is that the **Static Query** tab is now filled in with an **SQL** query to the products table.

15. Click the **Preview** button to check that the data can be retrieved from the database.

16. Close the **Preview** window, and close the **JDBC Data Source** dialog with **OK**.

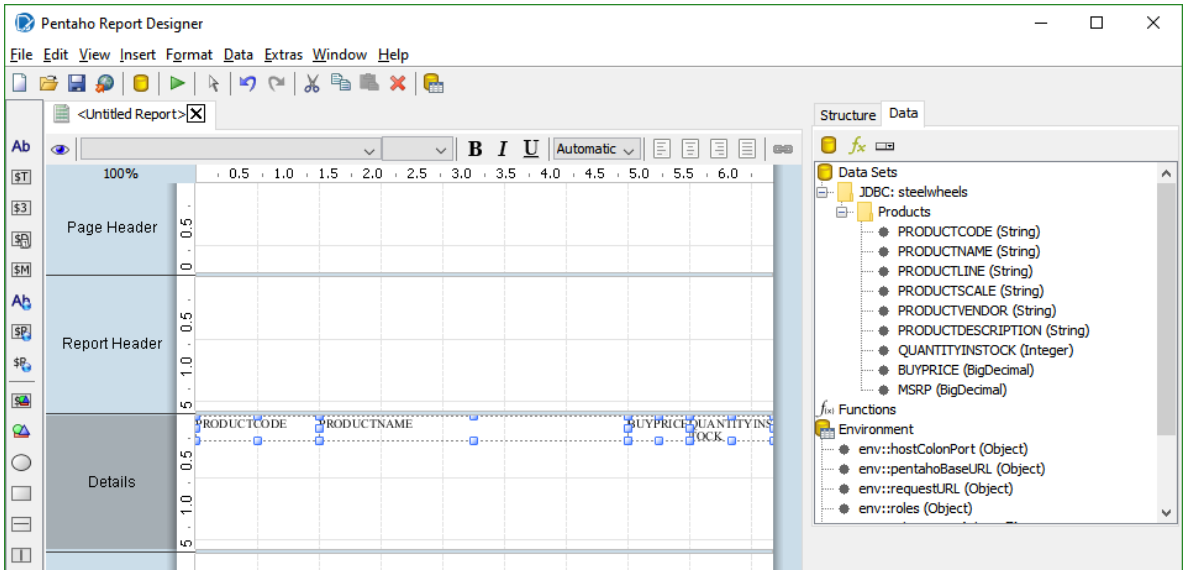
17. The **Data** tab should now display the fields of the query that we have just configured.

Bringing the data into the report

18. From the **Data** tab, drag the **PRODUCTCODE** to the report, and place it in the **Details** section.

19. Drag also the **PRODUCTNAME**, the **BUYPRICE**, and the **QUANTITYINSTOCK** to the Details section.

20. Arrange and resize these elements in the **Details** section, so that they look similar to the following figure.



21. Select the **BUYPRICE** element in the **Details** section.

22. On the right pane, switch from the **Data** tab to the **Structure** tab.

23. Below the **Structure** tab, there are two new tabs: **Style** and **Attributes**. These two tabs can be used to configure the properties of the selected element.

24. Switch to the **Attributes** tab.

25. In **format**, select the following format for the BUYPRICE element: **¤ #,##0.00;(¤ #,##0.00)**

26. Now select the **QUANTITYINSTOCK** field in the **Details** section.

27. In **format**, select the following format for the QUANTITYINSTOCK element: **#,##0**

28. At the top-left corner, click the preview button to preview the report.

29. You will see a listing of all the products in the database. This listing has several pages. You can switch between pages using the arrow buttons at the top.

30. At the top-left corner, click the edit button to switch back to the designer view.

Embellishing the report

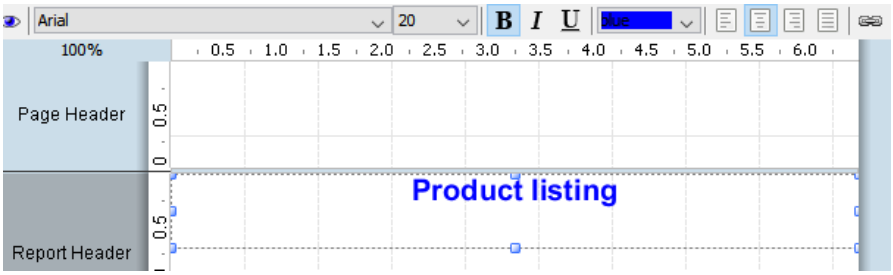
31. From the toolbar on the left, drag a **label** and drop it on the **Report Header** section.

32. Resize the label to fill the whole width of the report, and about half the height of the **Report Header** section.

33. Center the text in the label by clicking the appropriate button in the toolbar.

34. Double-click the label and replace its text with **Product listing**. Press **Enter** when done.

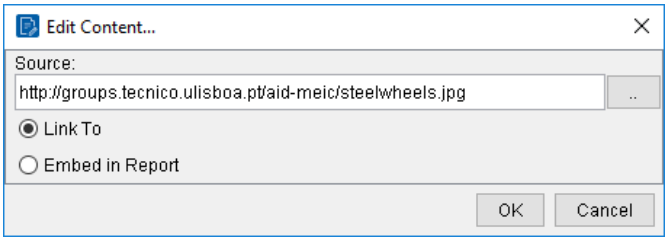
35. Using the options in the toolbar, select a **font family** of your liking (e.g. **Arial**), increase the **font size** to 20, make it **bold**, and change its color to **blue**.



36. From the toolbar on the left, drag an image element and place it on the **Page Header** section.

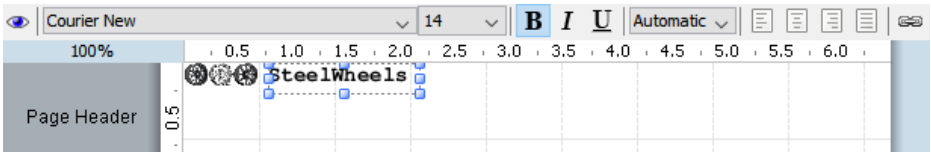
37. Double-click on the image element and specify the following URL as the source of the image:
    http://groups.tecnico.ulisboa.pt/aid-meic/steelwheels.jpg

    Note: The dialog may seem to contain a large textbox, but in fact it is just a single-line textbox together with a browse button.



38. Next to the image, place a **label** with the text **SteelWheels**.

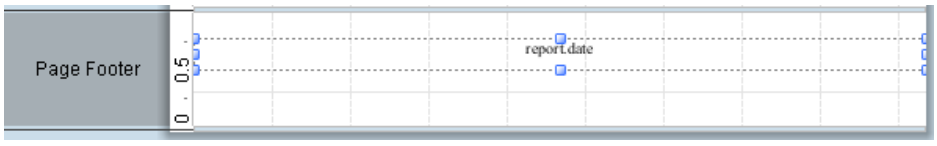
39. Change the font family and the font size of the label to match the size of the image.



40. From the toolbar on the left, drag a **date-field** and drop it on the **Page Footer** section.

41. Resize the date-field to occupy the whole width and center-justify the text.

42. Double-click the date-field and select **report.date**.



43. Press the **Run** button on the toolbar and select **PDF**. You should see your report in PDF format.

Grouping the products by product line

44. On the right pane, in the **Structure** tab, expand **Group** and click on the **Group Header**.

45. Below, in the **Attributes** tab, change the **hide-on-canvas** property to **false**.
    Note: The Group Header section will appear in the canvas.

46. In the **Structure** tab, right-click on **Group** and select **Edit Group**.

47. In the **Edit Group** dialog, select **PRODUCTLINE** on the left and click the arrow button to add it to the **Selected Fields**.

48. Close the **Edit Group** dialog with **OK**

49. On the right side, change from the **Structure** tab to the **Data** tab.

50. Drag **PRODUCTLINE** to the **Group Header** section in the report.

51. Align the **PRODUCTLINE** element with the top-left corner of the **Group Header** section, and expand it horizontally to occupy the full width of the report.

52. Using the options in the toolbar, select a **font family** of your liking, increase the **font size** to 14, make it **bold**.

53. While keeping the element selected, change ', below, to the **Style** tab.

54. In the **v-align** property, select **MIDDLE**.

55. In the **text-color** property, click on the value and then on the ellipsis button (...) and a color dialog will open.

56. Change to the **Swatches** tab, choose white, and press **OK**.

57. In the **bg-color** (background color) property, click on the value and then on the ellipsis button (...) and the color dialog will open again.

58. Change to the **Swatches** tab, choose a dark blue color, and press **OK**.

59. Click the preview button  to preview the report.

60. Navigate through the pages of the report using the arrow buttons at the top:


61. Check that the products are now organized (grouped) by category.

62. Click the edit button  to switch back to the designer view.

63. Press the **Run** button  on the toolbar and select **PDF**. You should see your report in PDF format.

64. Go back to Pentaho Report Designer. In the menu **File**, select **Save As**, and save your report in a file **product-listing.prpt**

Querying a data warehouse

65. In the menu **File**, click **New**.

66. In the **Data** tab, right-click **Data Sets** and select **OLAP > Pentaho Analysis**.

67. In the **Data Source** tab, next to **Pentaho Analysis Schema File**, click **Browse**.

68. Browse to the **steelwheels_dw.xml** file that you created in a previous lab, using Pentaho Schema Workbench (PSW).

69. Next to **Connections**, click the "**+**" button to add a new connection.

70. Configure the connection as follows:
    - Connection Name: **steelwheels_dw** (the data warehouse)
    - Connection Type: **MySQL**
    - Access: **Native (JDBC)**
    - Host Name: **localhost**
    - Database Name: **steelwheels_dw** (the data warehouse)
    - Port Number: **3306**
    - User Name: **aid**
    - Password: **aid**

71. Press the **Test** button to test the connection, and close the dialog with **OK**.

72. On the left side, under **Connections**, select the connection **steelwheels_dw** that you have just created.

73. On the right side, next to **Available Queries**, press the "**+**" button to add a new query.

74. In **Query Name**, rename the query to **SalesByCountry**.

75. In the **Static Query** tab, write the following MDX query:

```
SELECT Measures.Sales ON COLUMNS,
       NON EMPTY ORDER(Customer.Country.Members, Measures.Sales, DESC) ON ROWS
FROM Orders
```

76. Click the **Preview** button to check that the data can be retrieved from the data warehouse.

77. Close the **Preview** window, and close the **Pentaho Analysis Data Source** dialog with **OK**.

78. The **Data** tab should now display the fields of the query that we have just configured.

Designing the report

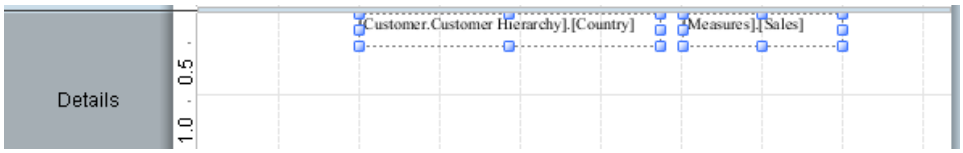
79. Copy and paste the **Page Header** elements from the previous report (**product-listing.prpt**) into the **Page Header** for this new report.
Note: Click the Page Header section before pasting.

80. Also, copy and paste the **Report Header** elements from the previous report (**product-listing.prpt**) into the **Report Header** for this new report
Note: Click the Report Header section before pasting.

81. Change the label text from **Product listing** to **Sales by country**. Press **Enter** when done.

82. Drag the following elements from the **Data** tab into the **Details** section of the report:
    - **[Customer.Customer Hierarchy].[Country]**
    - **[Measures].[Sales]**

83. Place the two elements side by side, align them to the top, and resize them to look similar to the following figure.



84. Preview the report to check that the data is being displayed correctly.

85. Click the edit button to switch back to the designer view.

86. From the toolbar on the left, drag a **chart** and drop it on the **Report Header** section.

87. Double-click the chart. The **Edit Chart** window will open.

88. On the left pane, change the **horizontal** property to **True**.

89. Scroll down and change the **show-legend** property to **False**.

90. On the right pane, for **category-column** select **[Customer.Customer Hierarchy].[Country]**.

91. For **value-columns**, click on the value and on the ellipsis button (...). The **Edit Array** window will open.

92. Select **[Measures].[Sales]** and click the arrow button to add it to **Selected Items**.

93. Close the **Edit Array** window with **OK**, and close the **Edit Chart** window with **OK**.

94. Preview the report to check that the chart is being displayed. 

95. Click the edit button  to switch back to the designer view.

96. Save the report as **sales-by-country.prpt**

Exercises

Now that you have understood how Pentaho Report Designer works, you should be able to perform the following task on your own:

97. Create a report with a pie chart and a listing of sales by product line.
    Use the following MDX query:

```
WITH MEMBER Measures.SalesPercent AS
            (Measures.Sales, Product.[Product Line].CurrentMember) /
            (Measures.Sales, Product.[Product Line].CurrentMember.Parent),
            FORMAT_STRING = '#0.00%'
SELECT {Measures.Sales, Measures.SalesPercent} ON COLUMNS,
        NON EMPTY ORDER(Product.[Product Line].Members, Measures.Sales, DESC) ON
ROWS
FROM Orders
```

    Configure the pie chart as follows:
    - In value-column select [Measures].[Sales]
    - In series-by-field select [Product.Product Hierarchy].[Product Line]

    In a listing below the chart, include the three columns: product line, sales, and sales percentage. Set the format of sales percentage to match the format string in the query above.

    Take a screenshot with the report in preview mode.