

# PENTAHO - DATA SOURCES & QUERIES

[http://www.tutorialspoint.com/pentaho/data\\_sources\\_and\\_queries.htm](http://www.tutorialspoint.com/pentaho/data_sources_and_queries.htm)

Copyright © tutorialspoint.com

In this chapter, we will learn to use the Pentaho Reporting Designer by taking an example. We will create a report on the employee database to produce a quick overview of every employee. We will create our first report by adding a data source and passing queries to the Pentaho Designer.

Before using Pentaho Report Designer, create a database named **employeedb** and in that database, create a table named **employee** using the following query.

```
CREATE TABLE 'employee' (  
  'id' integer NOT NULL,  
  'name' varchar(20),  
  'designation' varchar(20),  
  'department' varchar(20),  
  'age' integer,  
  PRIMARY KEY ('id')  
)
```

Insert the following records into the table.

Id	Name	Designation	Department	age
1201	'satish'	'writer'	'Tuto_Write'	24
1202	'krishna'	'writer'	'Tuto_Write'	26
1203	'prasanth'	'php developer'	'Tuto_develop'	28
1204	'khaleel'	'php developer'	'Tuto_develop'	29
1205	'raju'	'HTML developer'	'Tuto_develop'	24
1206	'javed'	'HTML developer'	'Tuto_develop'	22
1207	'kiran'	'Proof Reader'	'Tuto_Reader'	28
1208	'pravenya'	'Proof Reader'	'Tuto_Reader'	30
1209	'mukesh'	'Proof Reader'	'Tuto_Reader'	28
1210	'sai'	'writer'	'Tuto_writer'	25
1211	'sathish'	'graphics designer'	'Tuto_designer'	26
1212	'viswani'	'graphics designer'	'Tuto_designer'	24
1213	'gopal'	'manager'	'Tuto_develop'	29
1214	'omer'	'manager'	'Tuto_writer'	32
1215	'shirjeel'	'manager'	'Tuto_Reader'	32

If you want to manipulate the data contained inside the table, the best choice is to use SQL. But if you want to create a report based on the data, Pentaho Reporting is the best option. Our task is to pass an SQL query to the Pentaho Reporting designer tool and select respective fields *which are represented in the report* and present it on the Details of the report sheet.

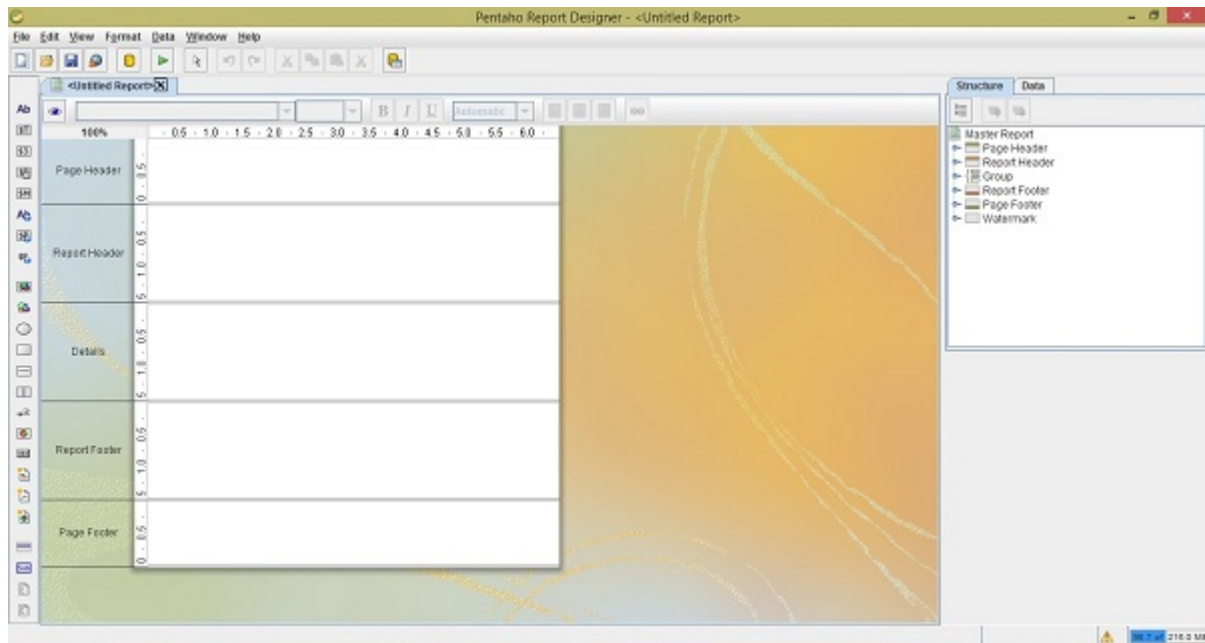
Before moving further, make sure you are well versed with all the navigation options available in Pentaho *explained in the previous chapter*. Now that we have a data source, let us proceed further and try to understand how to use Pentaho to generate a professional report.

## Steps to Generate a Report using Pentaho

Follow the steps given below to create a report from scratch without using Report Design Wizard.

### Step 1 : Create a New Report

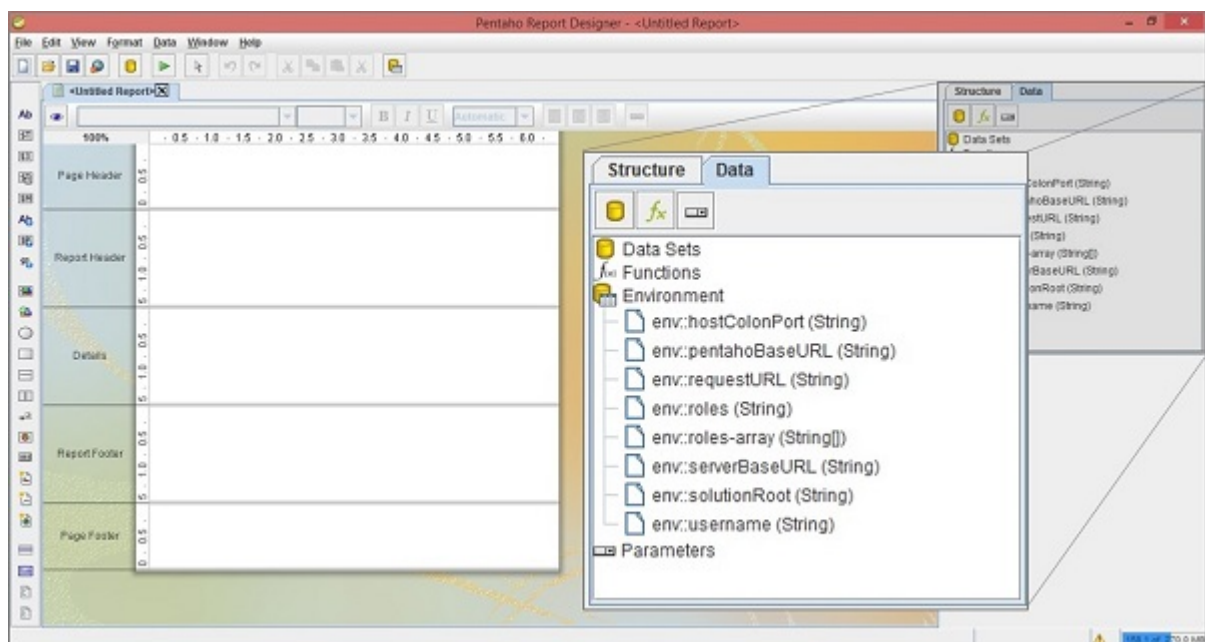
You can create a new report definition file by clicking "new report" on the welcome pane or go to "File → new".



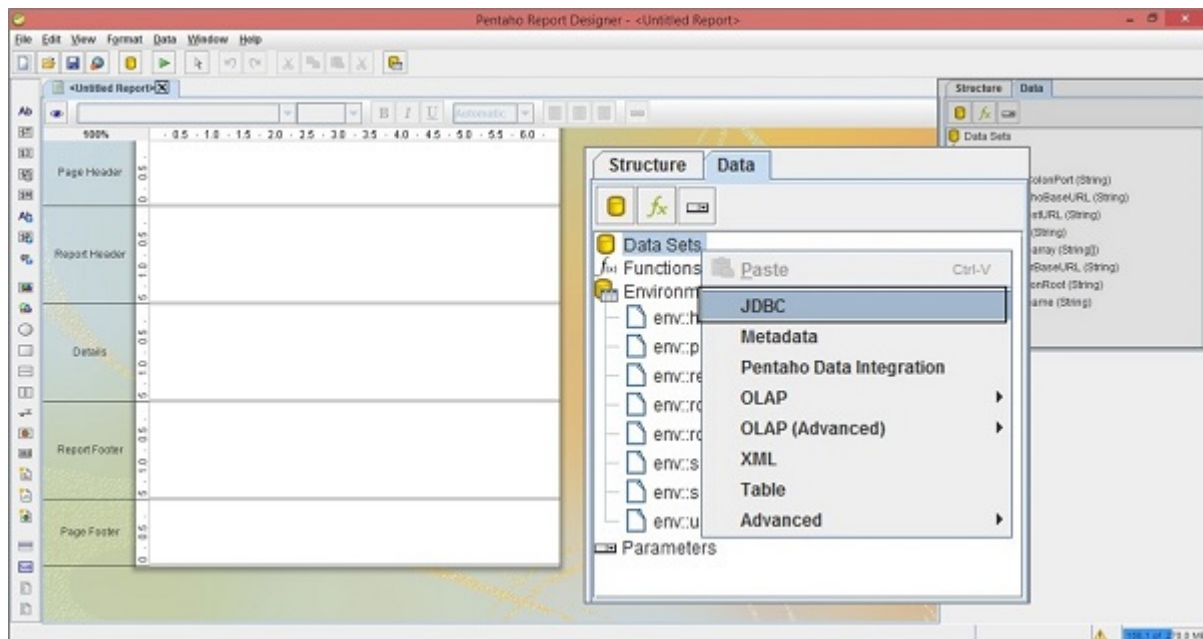
### Step 2 : Add a Data Source

The Structure Pane on the right-hand side provides a view of the visual elements of a report. The definition of the data source will be on the **Data** tab; it allows to define where the report data comes from and how this data is processed during the report processing.

A report generally displays the data that is supplied by a data source in the form of a table, whereas a report definition defines how the report is to be formatted or printed. As shown in the following screenshot, select the **Data tab** from **structure pane**.

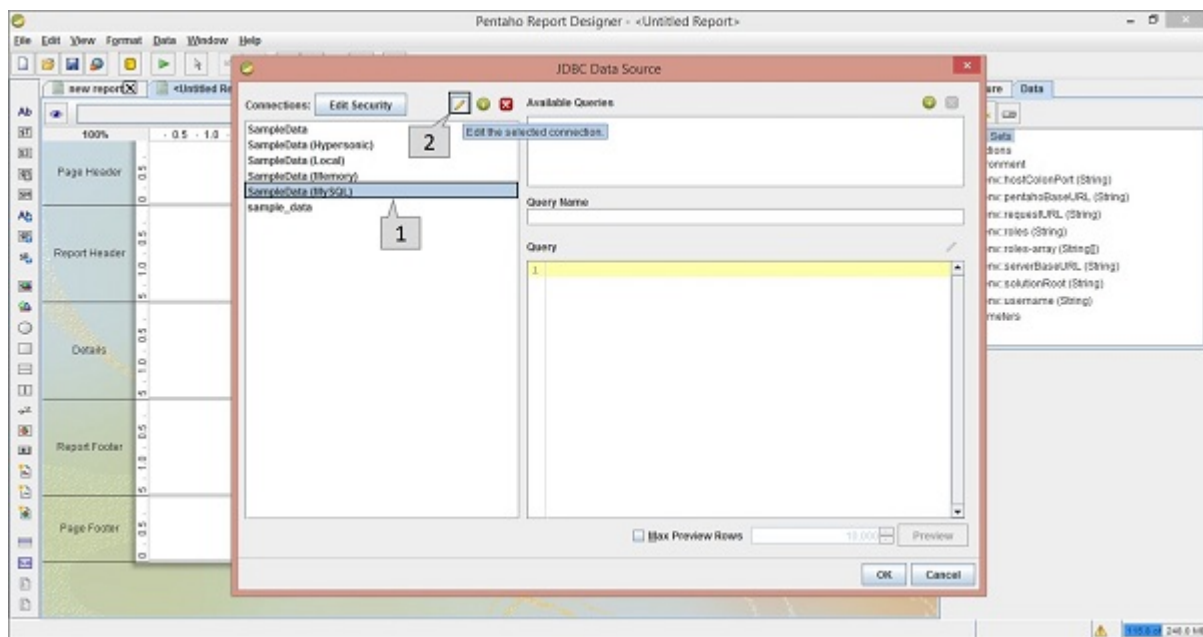


In the Data tab, right-click on the Data Sets and select JDBC to add a data source. Generally, in the list of options, you can select any other option based on the requirement. It means, if you have an XML file as your data source, then choose XML option from the list. Take a look at the following screenshot. Here we are selecting the JDBC option to add a database as data source.



After having selected the JDBC option as the data source, you will find a dialog box as shown in the following screenshot.

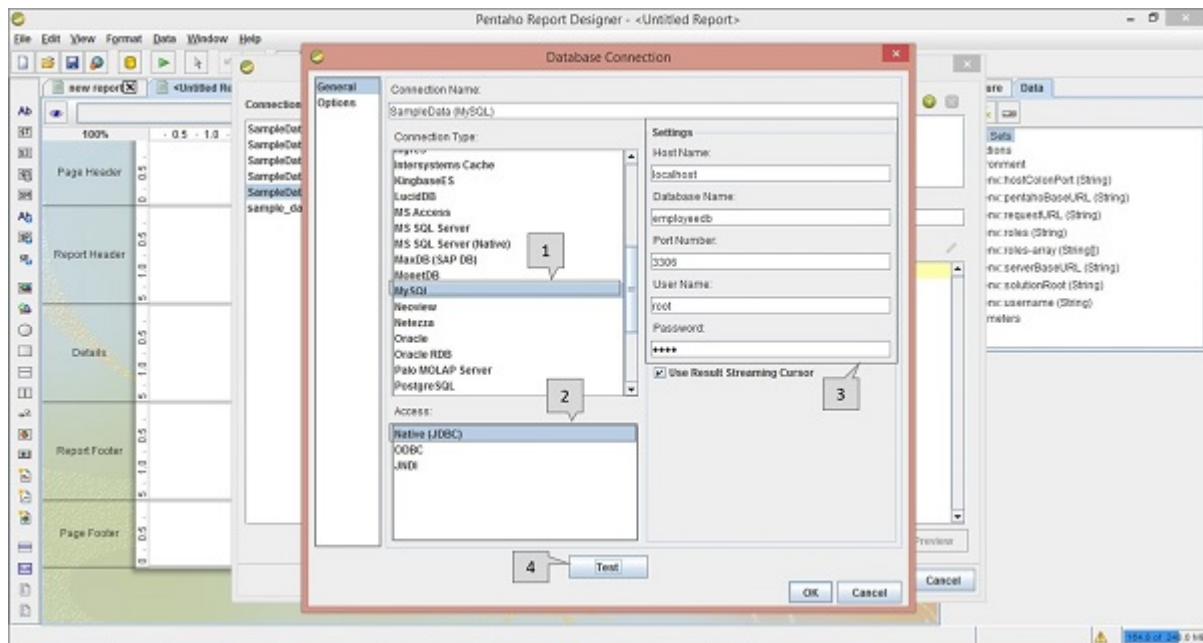
We have already chosen MySQL database for the data source, therefore we have to select the **SampleData MySQL** option in the left-side panel of the dialog box *marked as pointer "1"* in the given screenshot. Pointer "2" is meant for editing the connection statement and URL to interact with the database.



The following screenshot shows a dialog box where you can define your connection statement and the URL for the database. We need to carry out four operations on the following screen *which are highlighted using pointers*.

- In the connection type list, select **MySQL** – We have already chosen MySQL as the database *datasource*.
- In the Access list, select **Native JDBC** – Through JDBC connection, we can access the database.
- In the Settings section, we must mention the Host Name *localhost*, Database name *employeeedb*, port number 3306, username *root*, and the password *asperyousystem*.
- Test the connection statement by clicking the **Test** button.

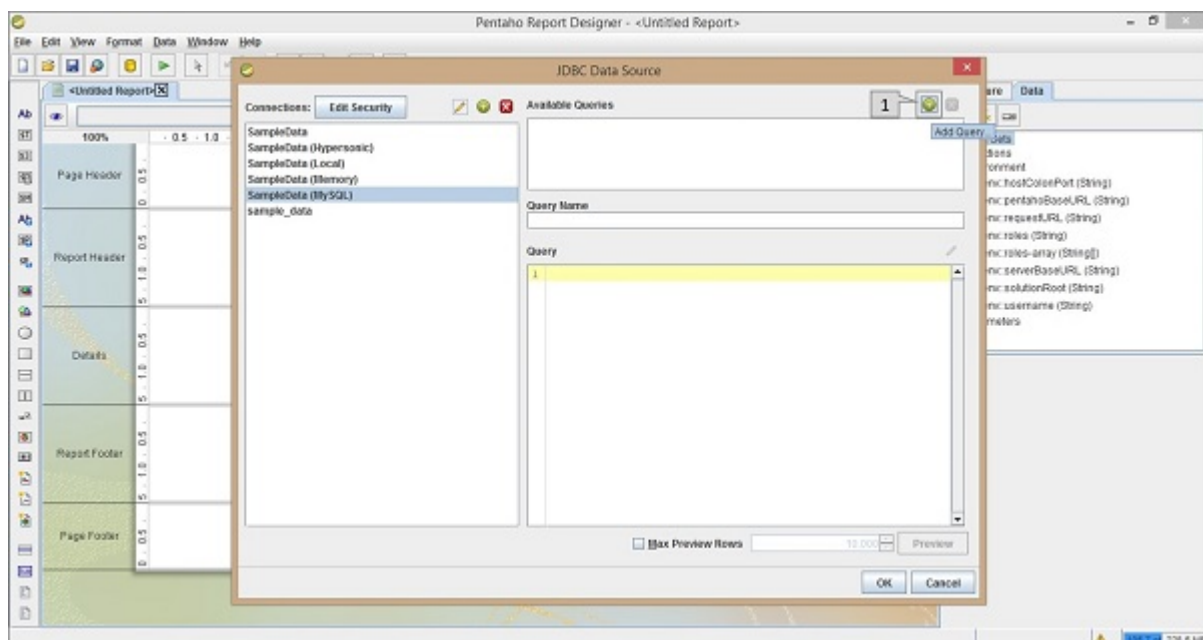
Finally, click the **OK** button to confirm the database connection.



### Step 3 : Add a Query

Take a look at the following screenshot. The dialog box presents the available saved queries available through the database connection.

- The **Available Queries** block on the right side of the dialog box displays a list of all the available queries.
- The **Query Name** block displays the selected query name which is selected in the above available queries list.
- The **Query** block displays the query statement. If no queries are available or if you want to create a new query, click the “&plus;” button which is highlighted as pointer “1” in the following screenshot.



While clicking the “&plus;” button, you can create a query by editing a name on the **Query Name** block as **select\_all\_records** and use the following query statement in the **Query** block.

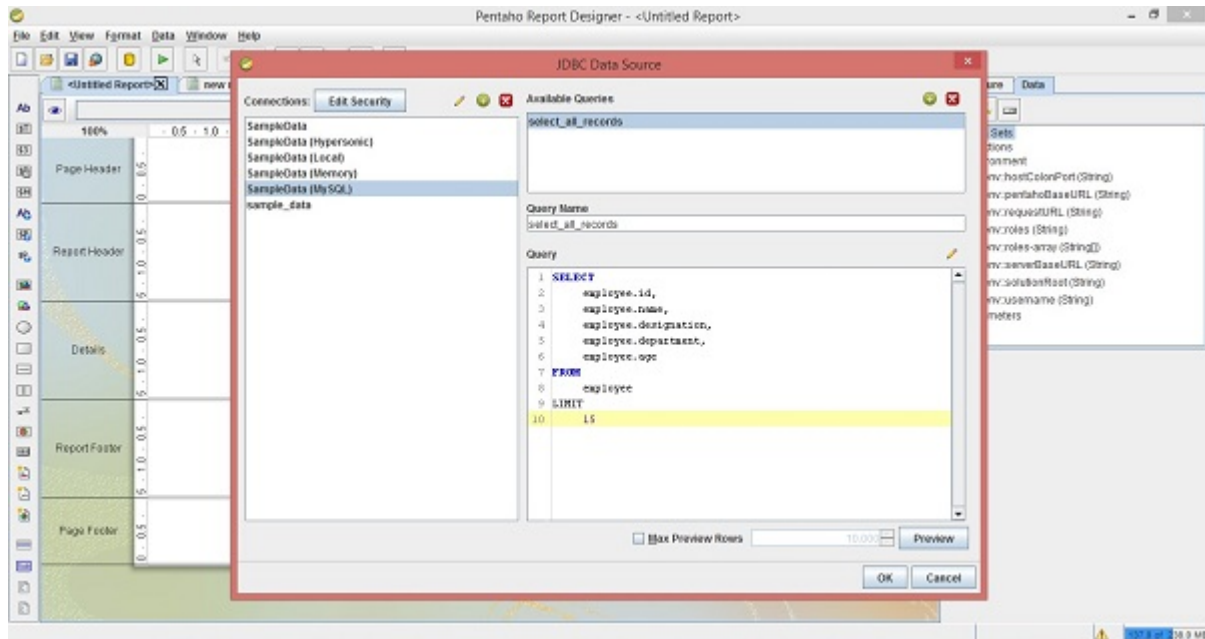
```
SELECT
  employee.id,
  employee.name,
  employee.designation,
  employee.department,
```

```

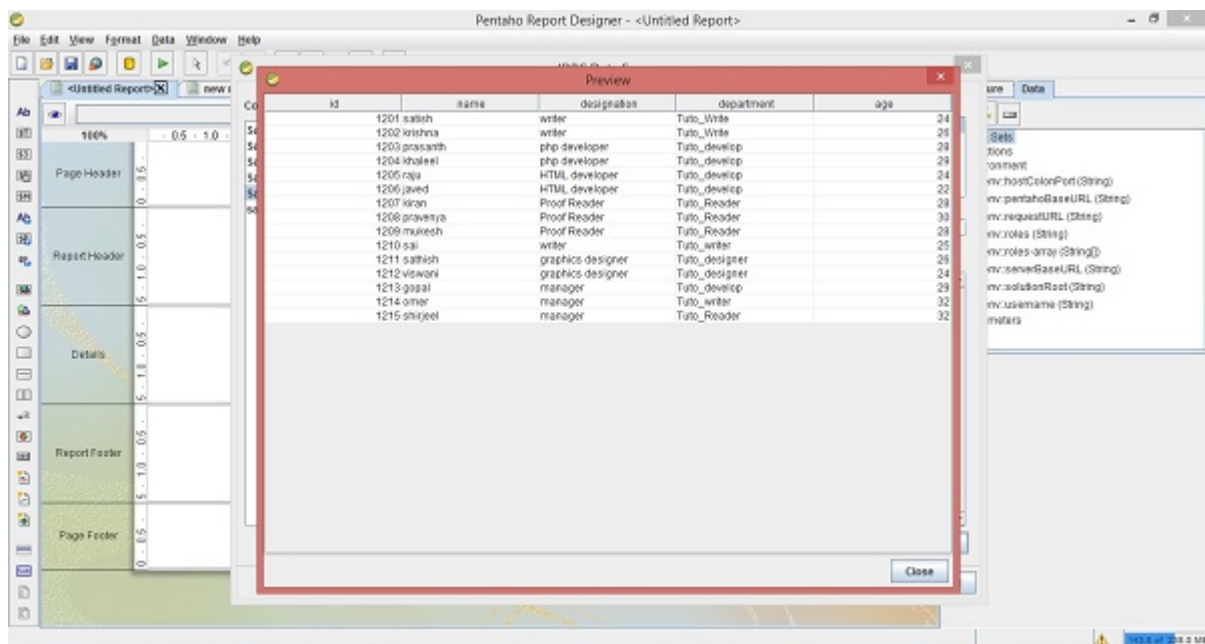
employee.age
FROM
employee
LIMIT
15

```

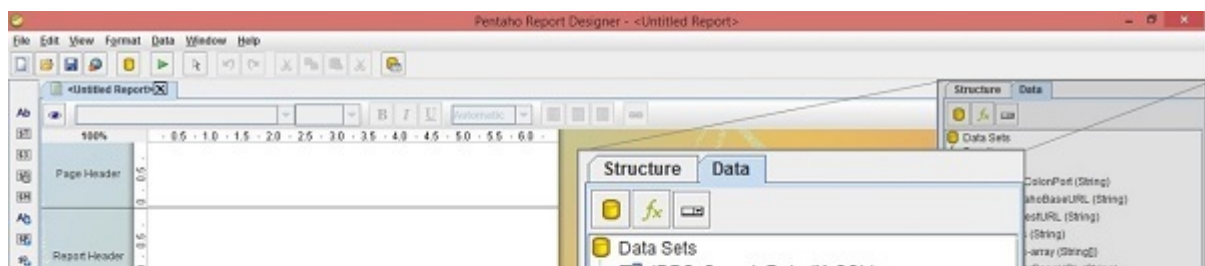
After adding the query, you should get the following dialogue box. Click the preview button.

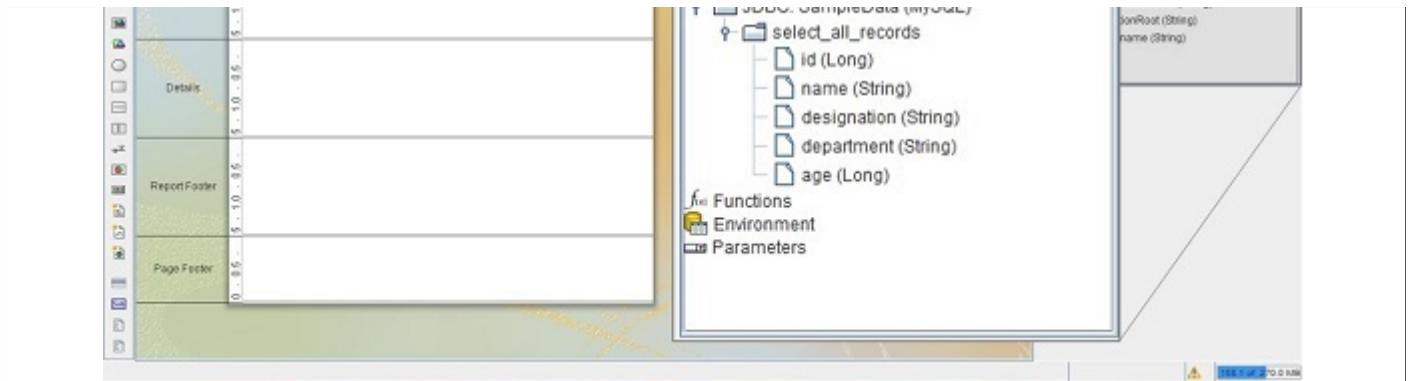


After clicking the preview button, you will find all the **employee** table records in a separate dialog box as shown in the following screenshot. Click the **close** button.



Then, click the **OK** button to submit the query. After submitting the query, you will find all the table field names and their datatypes under the query name on the right-side structure pane, as shown in the following screenshot. Here, the maximized box is the structure pane which is placed on the right side of the screen.





We have so far added a data source and a query to the Pentaho Reporting Designer. Now, we have to add elements into the workspace to create a report. The same example is extended to the next chapter "**Reporting Elements**"

Loading [Mathjax]/jax/output/HTML-CSS/jax.js