PRACTICAL FILE

OF

Data Warehousing

&

ETL Technologies - Lab

(CA130)

**SUBMITTED TO :- SUBMITTED BY :**

Mrs. Deepika Chaudhary Divya Sarna

1810992717

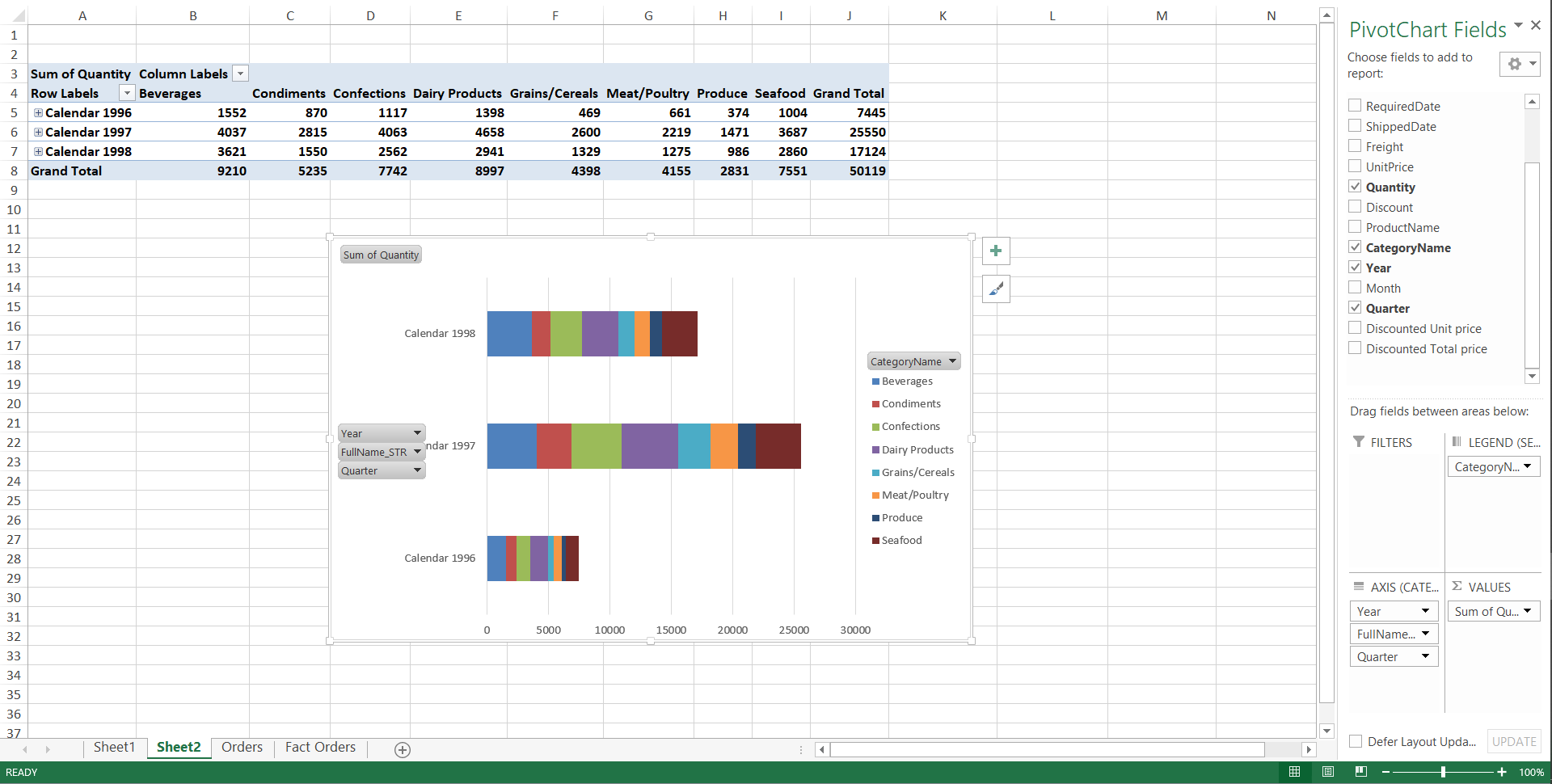
BCA – 5Z-UCA

**Experiment 1:** How to illustrate data analysis using Pivot Table and Charts and Use of functions (count, Average, Product etc.) and Exercise questions.

**Solution)**

Steps:

1. Collect data in excel sheet on which we are going to analyze.
2. Select the data and under “Insert” tab click on pivot table and pivot chart.
3. Now click on “OK” and a new existing window will appear consisting of Pivot chart fields.
4. Drag and drop the attributes from there into Axis, values and legends respectively.
5. The graph will automatically configure according to the data. Now the data will be more organized and easy to analyze.



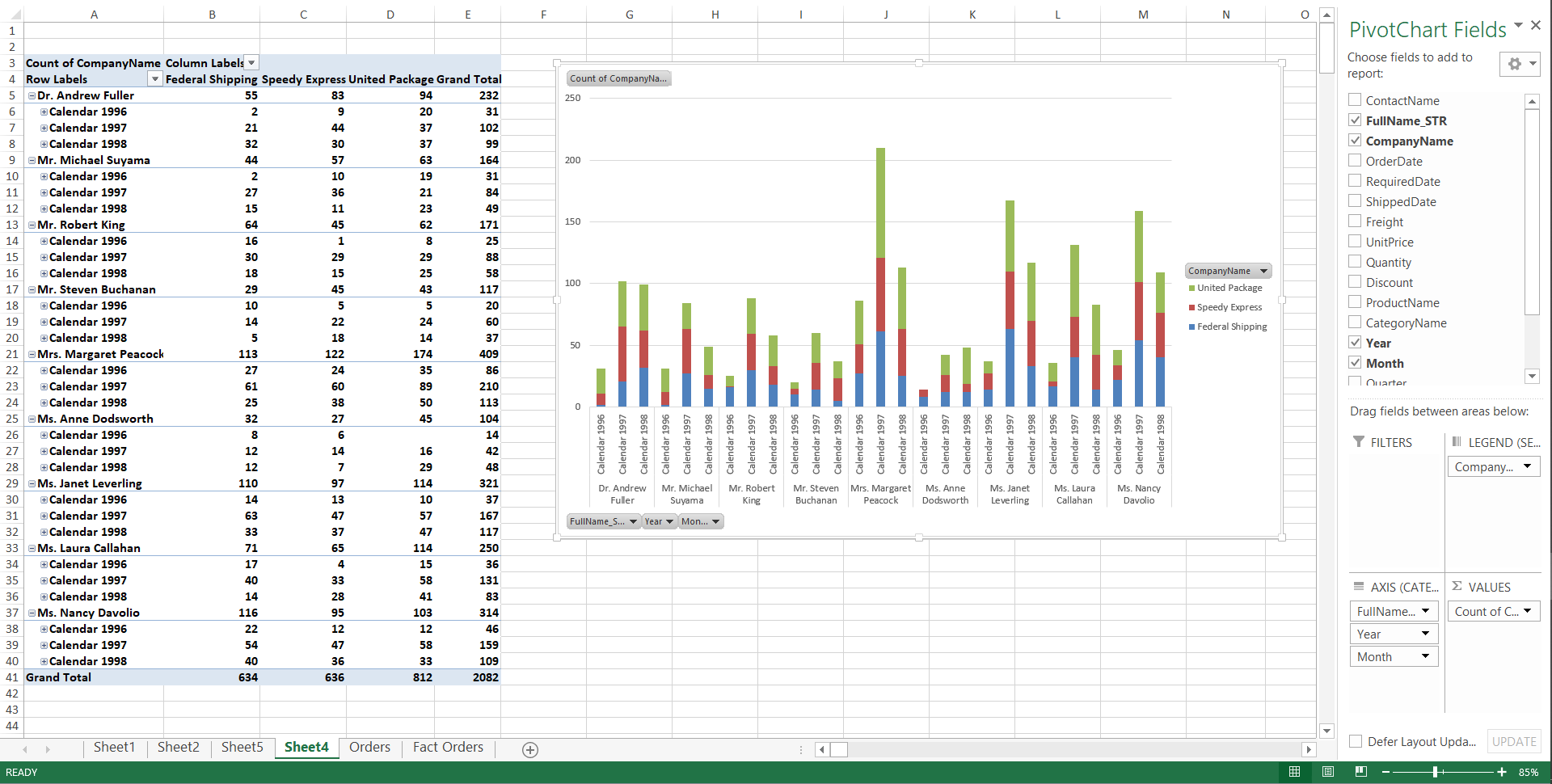
**Experiment 2:** Use of functions such as refresh, filters , drill down option of pivot tale and Exercise questions on Pivot table and chart implementation.

**Solution)**

Steps:

1. Collect the data in excel sheet on which we are going to analyze.
2. Select the data and under “Insert” tab click on Pivot table and pivot chart.
3. Now click on “OK” and a new or existing window will appear consisting of Pivot chart fields.
4. Drag and drop the attributes from there into Axis, Values and Legends respectively.
5. **Functions:** To use this function, click on “Value field settings” and select options under “summarize by”. Same goes for product, max, min, etc.
6. **Drill down:** Left click on the selected data value and select “show details”. It will open a drill down dashboard in new sheet.

Pivot chart exercise to find the average speed of a roller coaster from analyzing data.



**Experiment 3:** How to perform slicing and dicing operation in MS Excel.

**Solution)**

Steps:

1) Collect the data in excel sheet on which we are going to analyze.

2) Select the data and under “Insert” tab click on Pivot table and pivot chart.

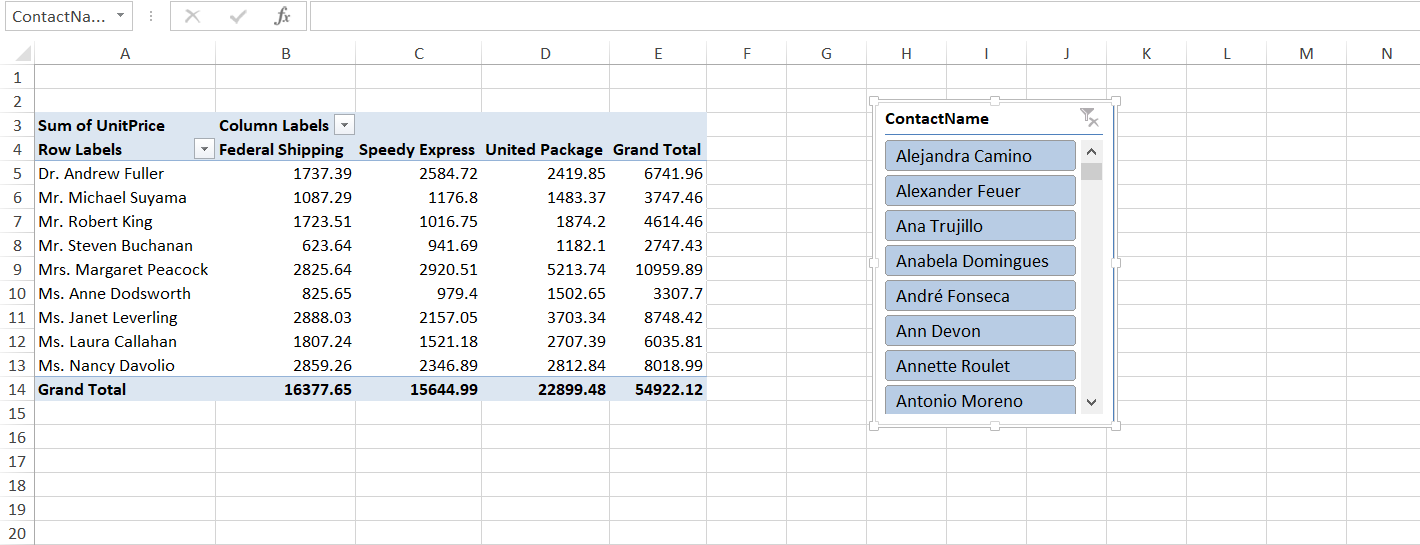
3) Now click on “OK” and a new or existing window will appear consisting of Pivot chart fields.

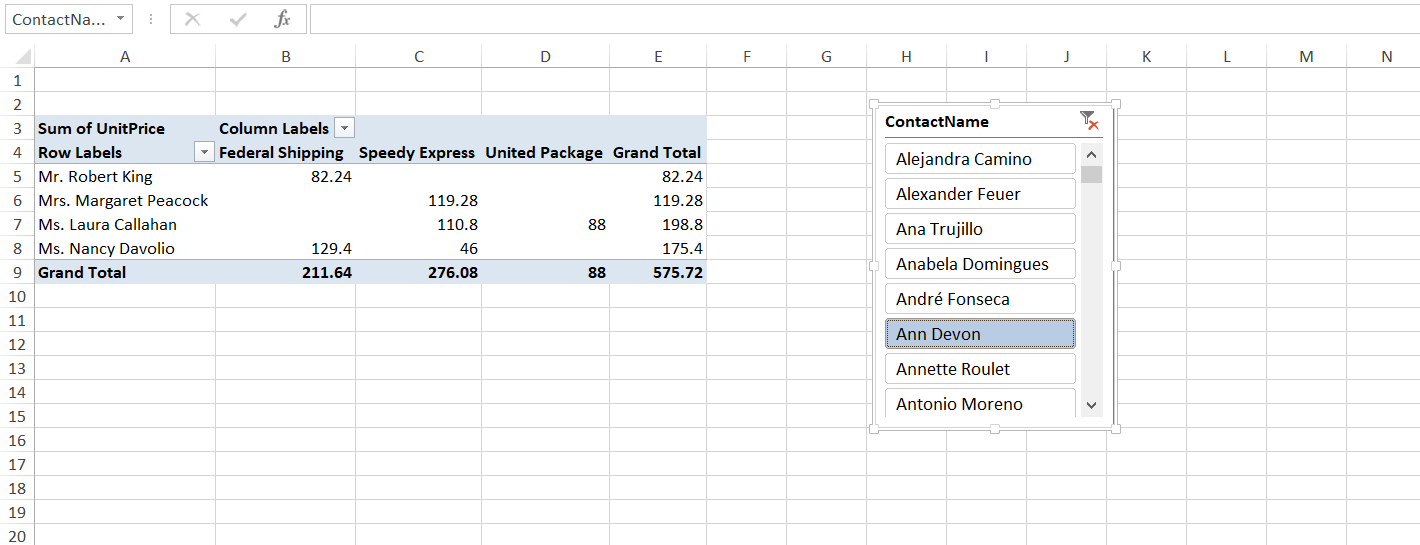
4) After arranging the data in pivot data fields, click on any cell in the pivot table.

5) Under “Insert” Tab click on Insert slicer and select the desired field.

6) Slicer option will give a button like dashboard to slice and dice the data according to user requirement.

7) It makes loads of data easy to analyze by categories.





**Experiment 4: How to perform the VLook UpHLook Up operation in MS Excel**

**Solution)** Steps for VLOOKUP:

1)To implement VLOOKUP, get the data on which we want to use VLOOKUP.

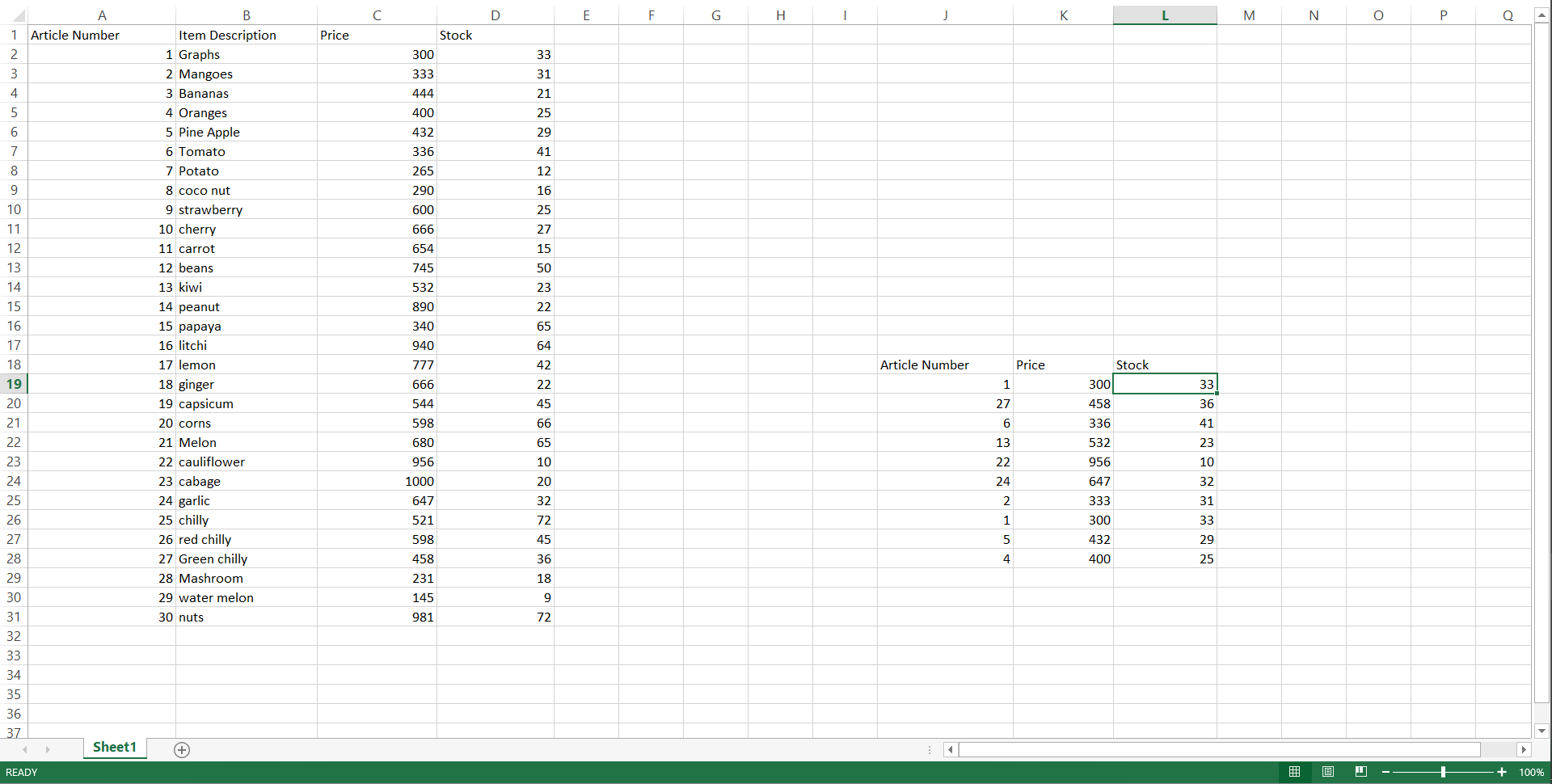
2) From the data, write the fields name in the side sheet and write their article number as it’s shown in the output.

3) Use the VLOOKUP formula in the cell where you want to fetch the data. **VLOOKUP(lookupvalue,table\_array,col\_index\_num,Range Lookup)**

4) **Specify the data which you want VLOOKUP to use for search in table\_array.**

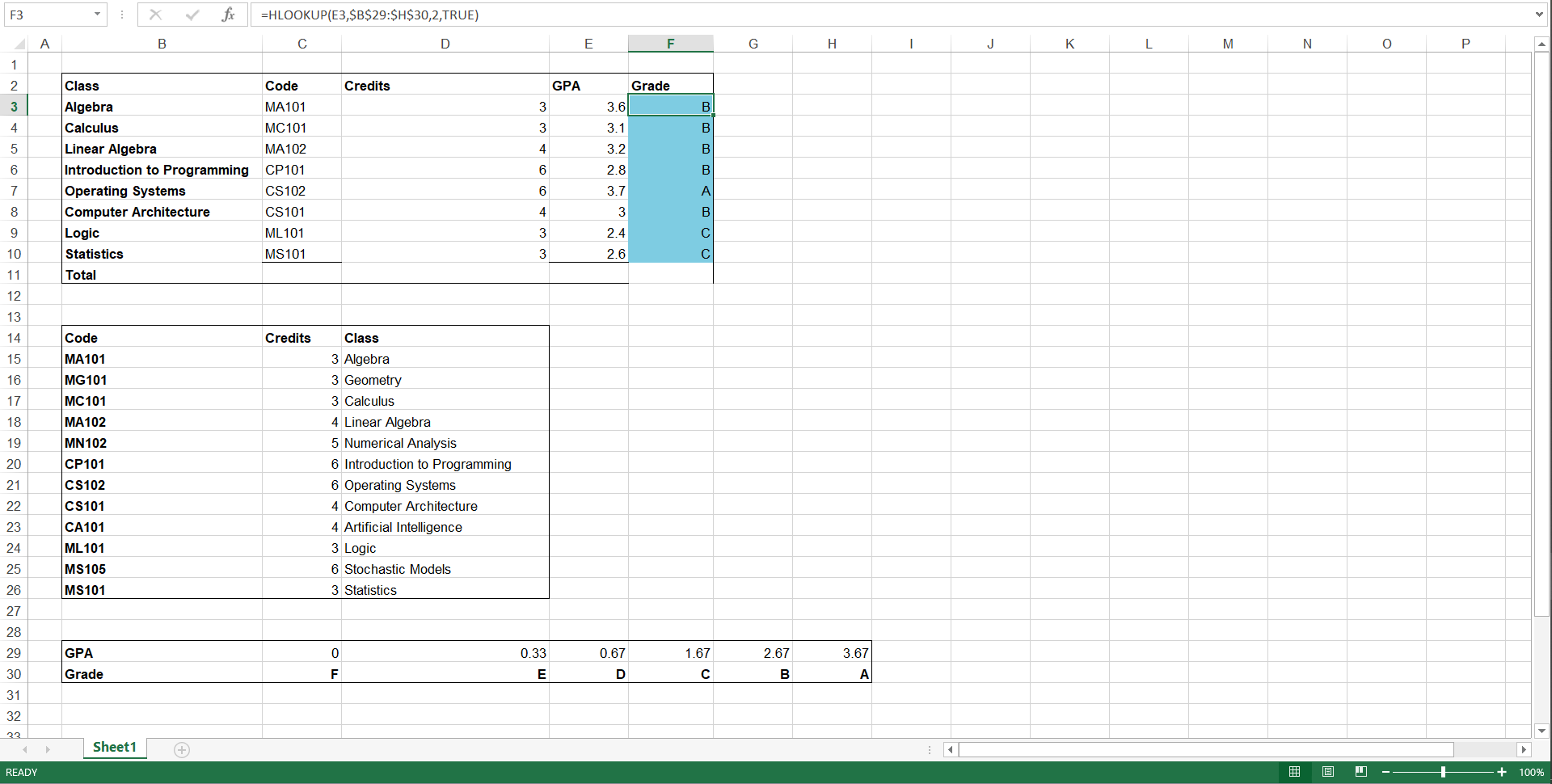
**5) Specify column number and range value. False for exact match and True for Approximate match.**

**6) Press Enter to find the result. To fetch more than one result just copy the above result and paste in other cells.**



**Steps for HLOOKUP:**

The steps are same as of VLOOKUP. HLOOKUP works on horizontal cells.

****

**Experiment 5:** How to Install and Run Pentaho Data Integration IDE and Weka Installation

**Solution)** Pentaho is business intelligence(BI) software that provides data integration, OLAP services, reporting, information dashboard, data mining and ETL capabilities.

Steps to install Pentaho by Hitachi Vantara:

1) To install and run Pentaho Data Integration, we must have JDK installed in our system.

2) Now download the Pentaho setup, run the installer.

3) Make sure the PATH is set to JAVA in environment variables.

4) Now under data-integration folder, click on “set-pentaho-env.bat” and run the batch file.

5) After setting the environment, run the batch file “spoon.bat”, it will open Pentaho Data integration software.

**Experiment 6:** Create a Simple PDI Transformation for transforming data from:

1. From Text File to Excel file.
2. From Excel file to Text File
3. From Text to XML File

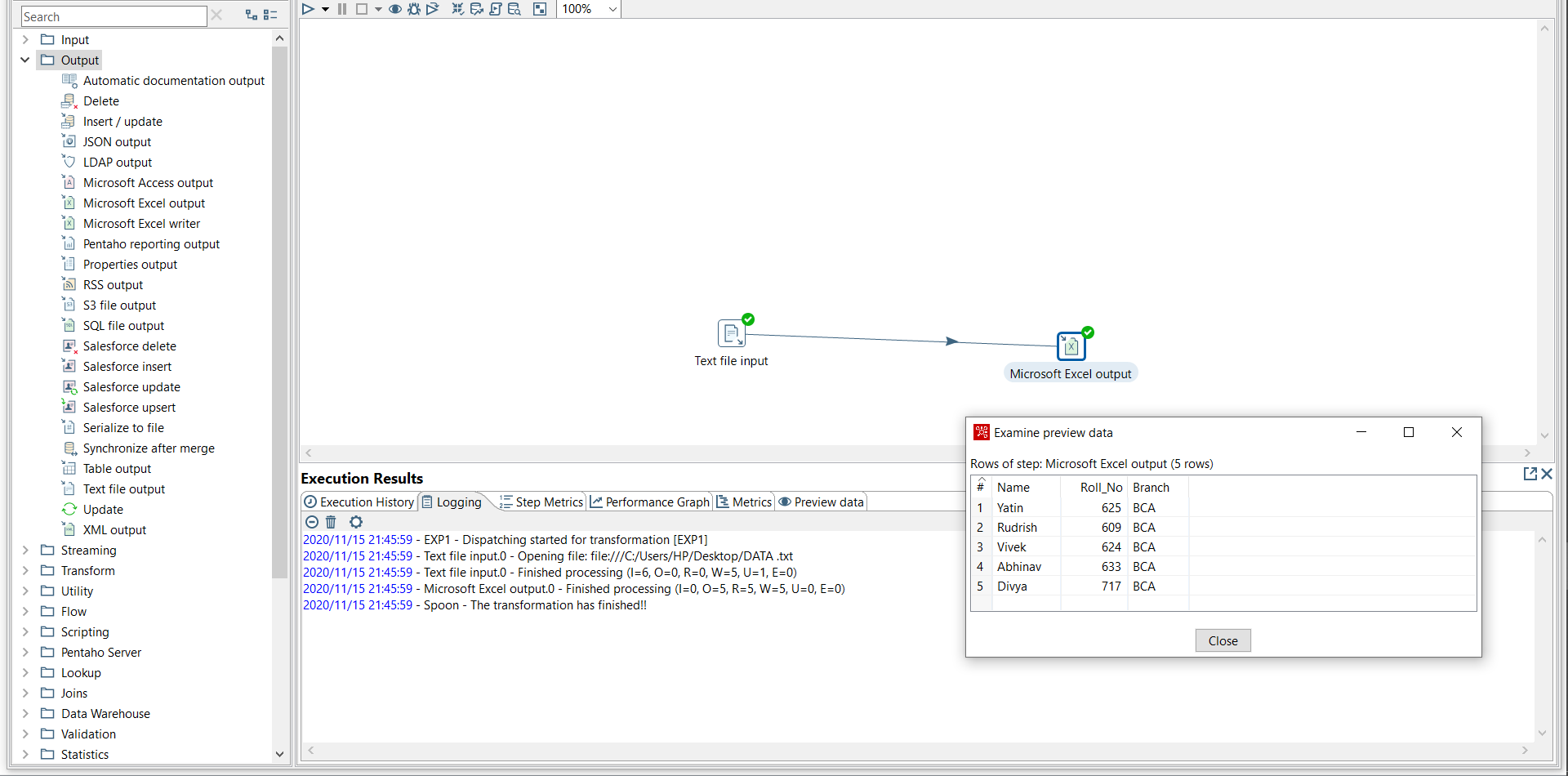
**a) From Text file to Excel File**

**Solution-a)**

Steps for text to excel:

1. Click on Input tab and select Text input, drag and drop it to the area.
2. Click on Output tab and select Excel output, drag and drop it to the area.
3. Select the input text, double click on it and add the source file, get the fields.
4. Draw the Hop from Input to the output file, in the excel output save the location and get the fields.
5. Now click on Run and see the preview.

**Output:**



**b) From Excel to text file**

**solution)**

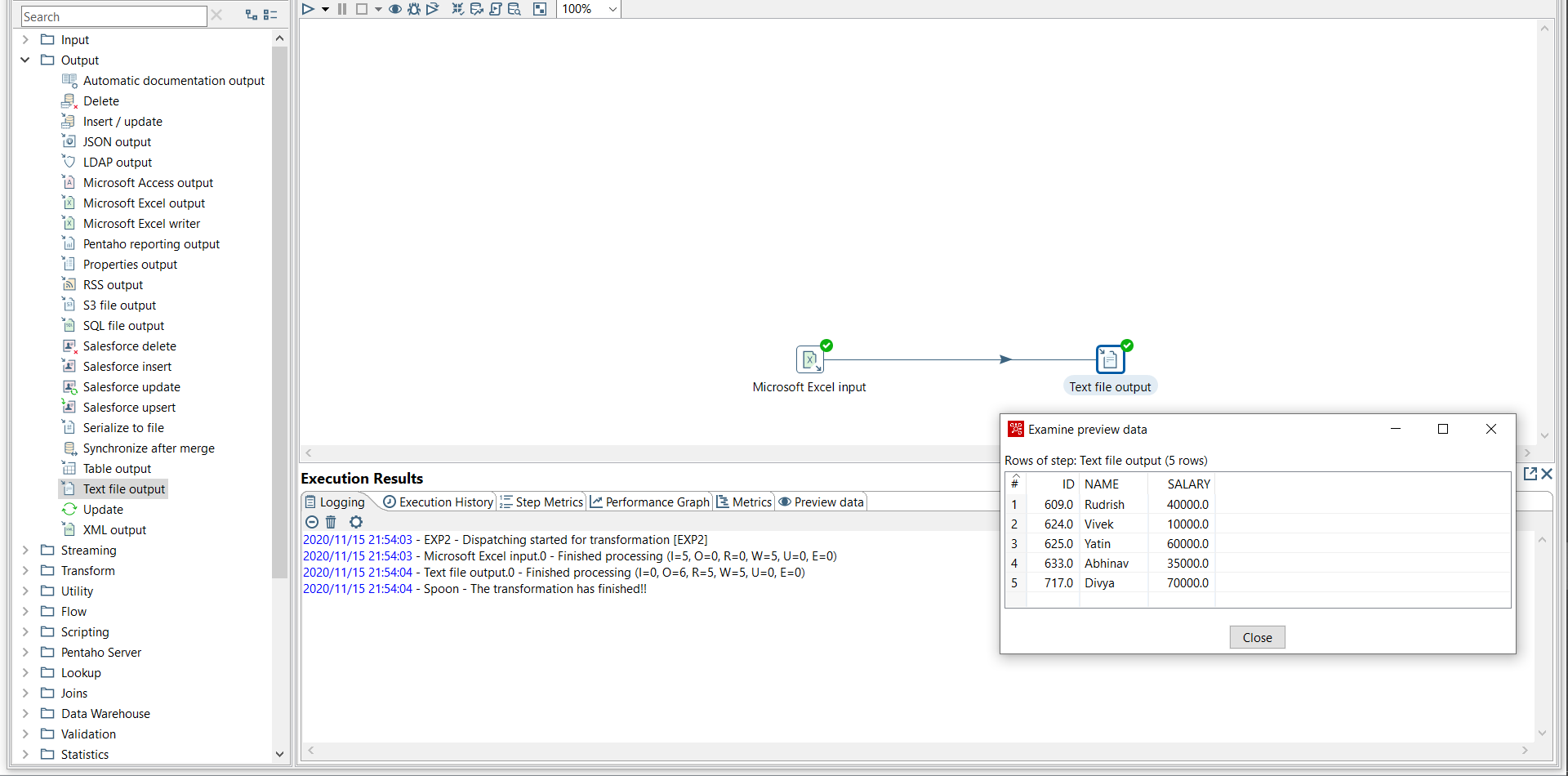
1.Click on Input tab and select Excel input, drag and drop it to the area.

2.Click on Output tab and select text output, drag and drop it to the area.

3. Select the input excel, double click on it and add the source file, get the fields

4. Draw the Hop from Input to the output file, in the text output save the location and get the fields.

5. Now click on Run and see the preview.



**Experiment 7:**

1. Create a Simple PDI Transformation for transforming data from Multiple Excel File to One Excel File.
2. Create a Simple PDI Transformation for transforming data from multiple text File to One Text File.
3. Create a Simple PDI Transformation for transforming data from Multiple text File to One Excel File

**Solution**

**a)** Steps:

1) Drag and Drop Microsoft Input files and place in the area.

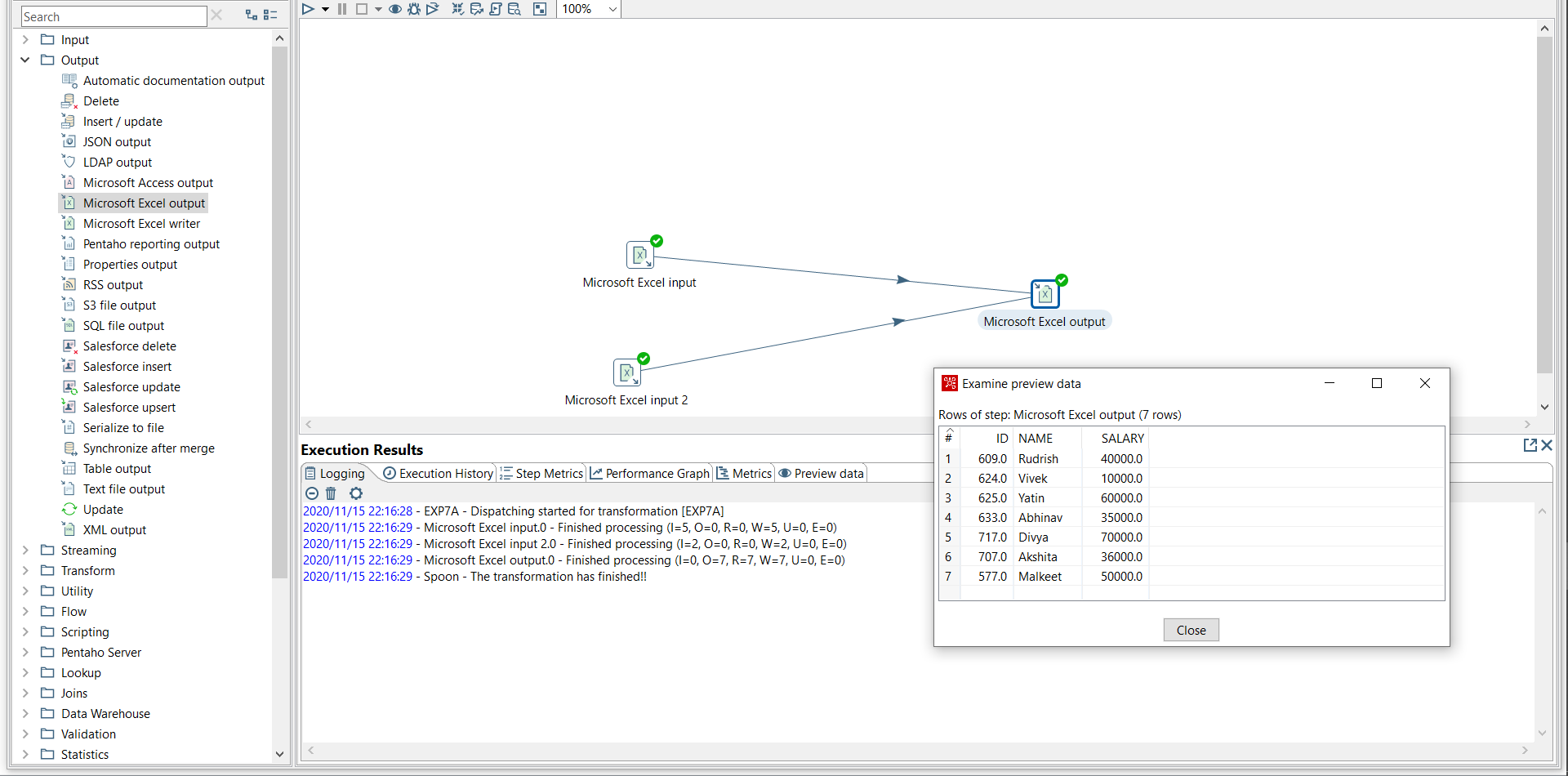
2) Add the data source files in them and get sheets, fields.

3) Drag and drop the Excel output file and place in the area.

4) Add Hop from input to output files and get sheet field , preview rows and save location.

5) Click on run and preview the output.

**Output:**

****

**b) Create a Simple PDI Transformation for transforming data from Multiple text file to One text File.**

**Solution b)** Steps:

1) Drag and Drop Text Input files and place in the area.

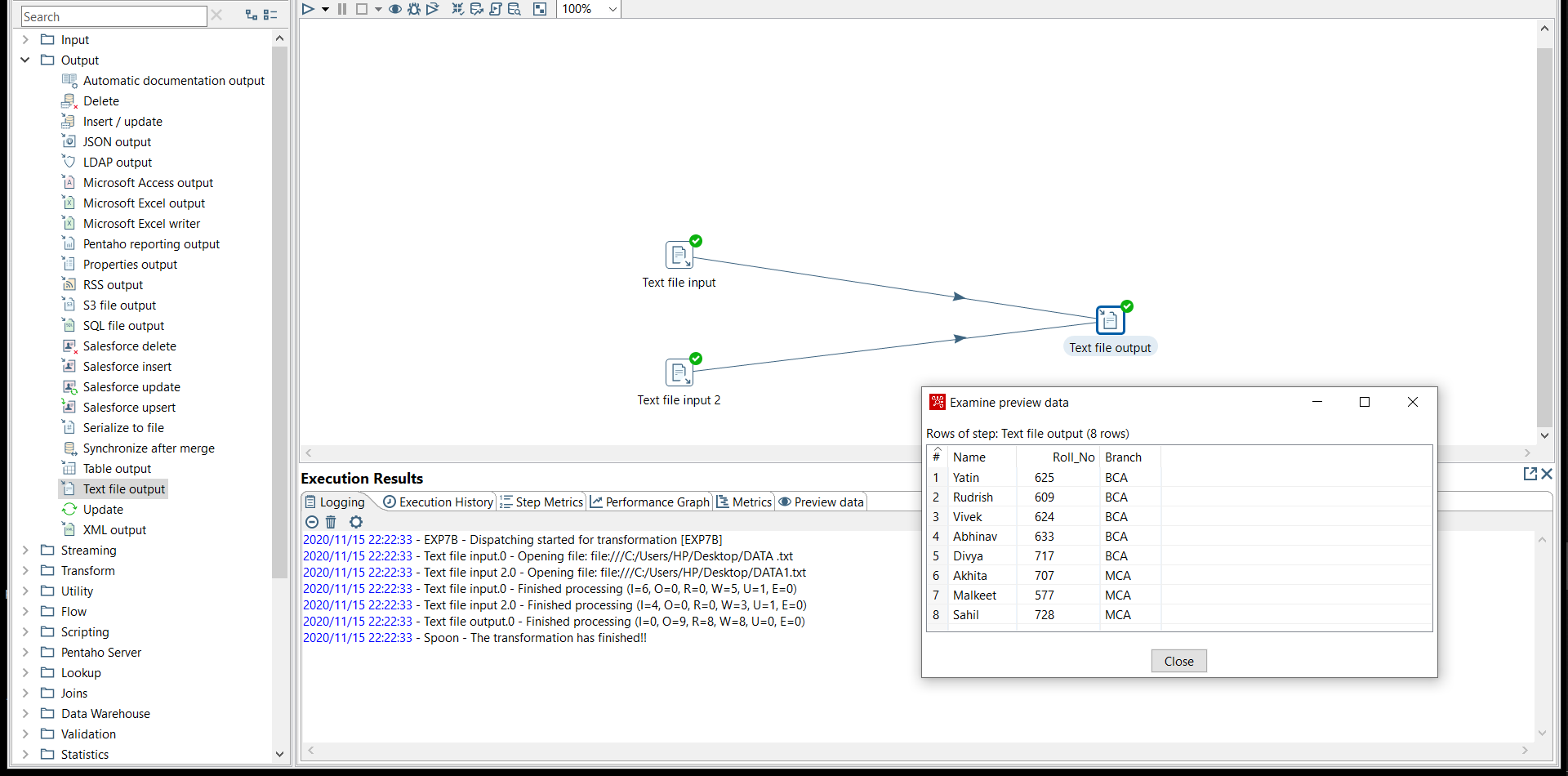
2) Add the data source files in them and get sheets, fields.

3) Drag and drop the Text output file and place in the area.

4) Add Hop from input to output files and get sheet field , preview rows and save location.

5) Click on run and preview the output.

**Output:**

****

**c) Create a Simple PDI Transformation for transforming data from Multiple text file to One Excel File.**

**Solution)**

1) Drag and Drop Text Input files and place in the area.

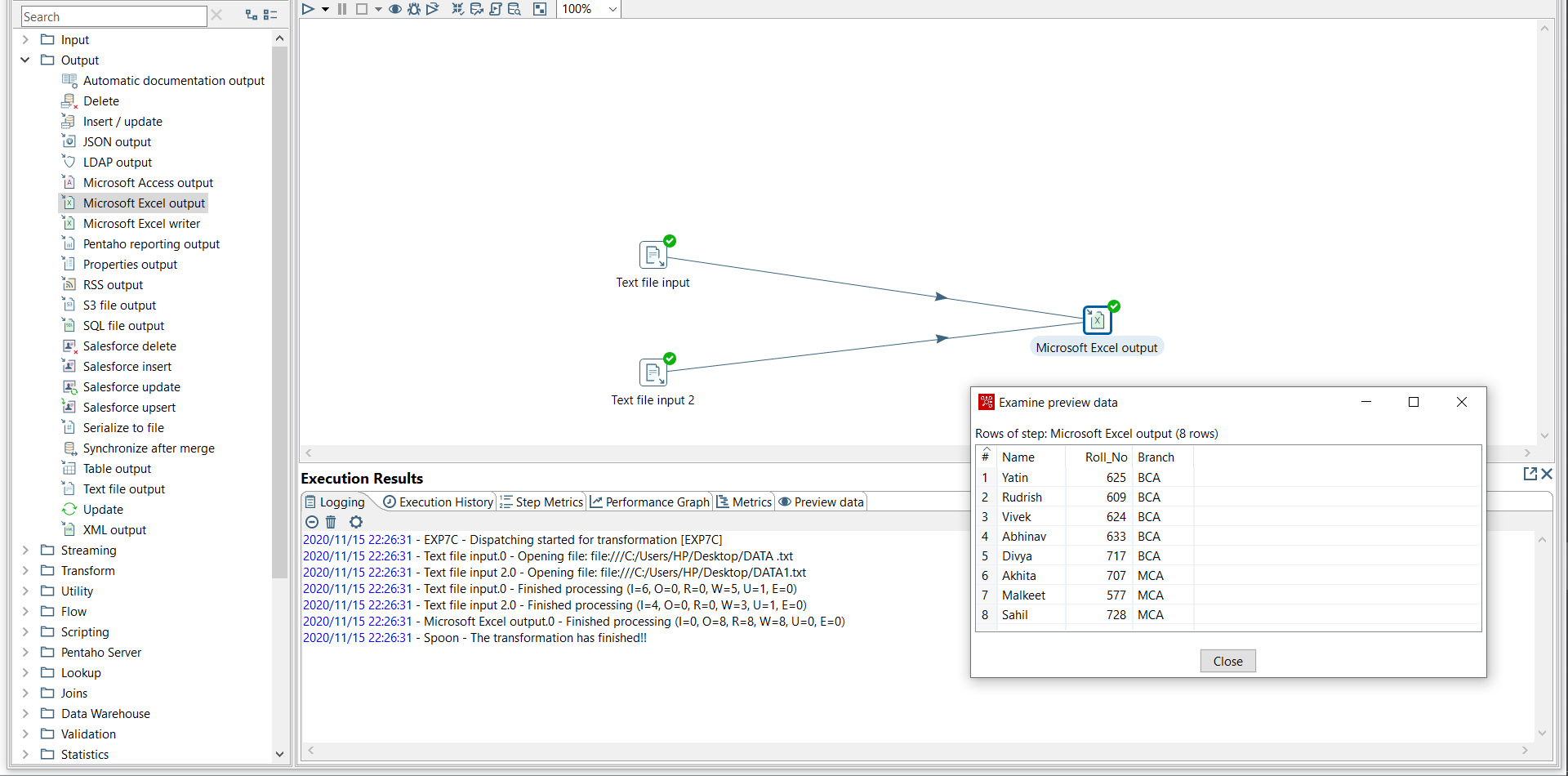
2) Add the data source files in them and get sheets, fields.

3) Drag and drop the Text output file and place in the area.

4) Add Hop from input to output files and get sheet field , preview rows and save location.

5) Click on run and preview the output.

**Output:**



**Experiment 8:**

1. Use of Filter rows for dividing Source data to two destinations.
2. Distribution of data from one excel file to text file and excel file
3. Distribution of data from one Text file to Multiple excel files.

**a) Use of Filter rows for dividing Source data to two destinations.**

**Sol a)** Steps:

1) Drag and Drop Text Input files and place in the area.

2) Add the data source files in them and get sheets,fields.

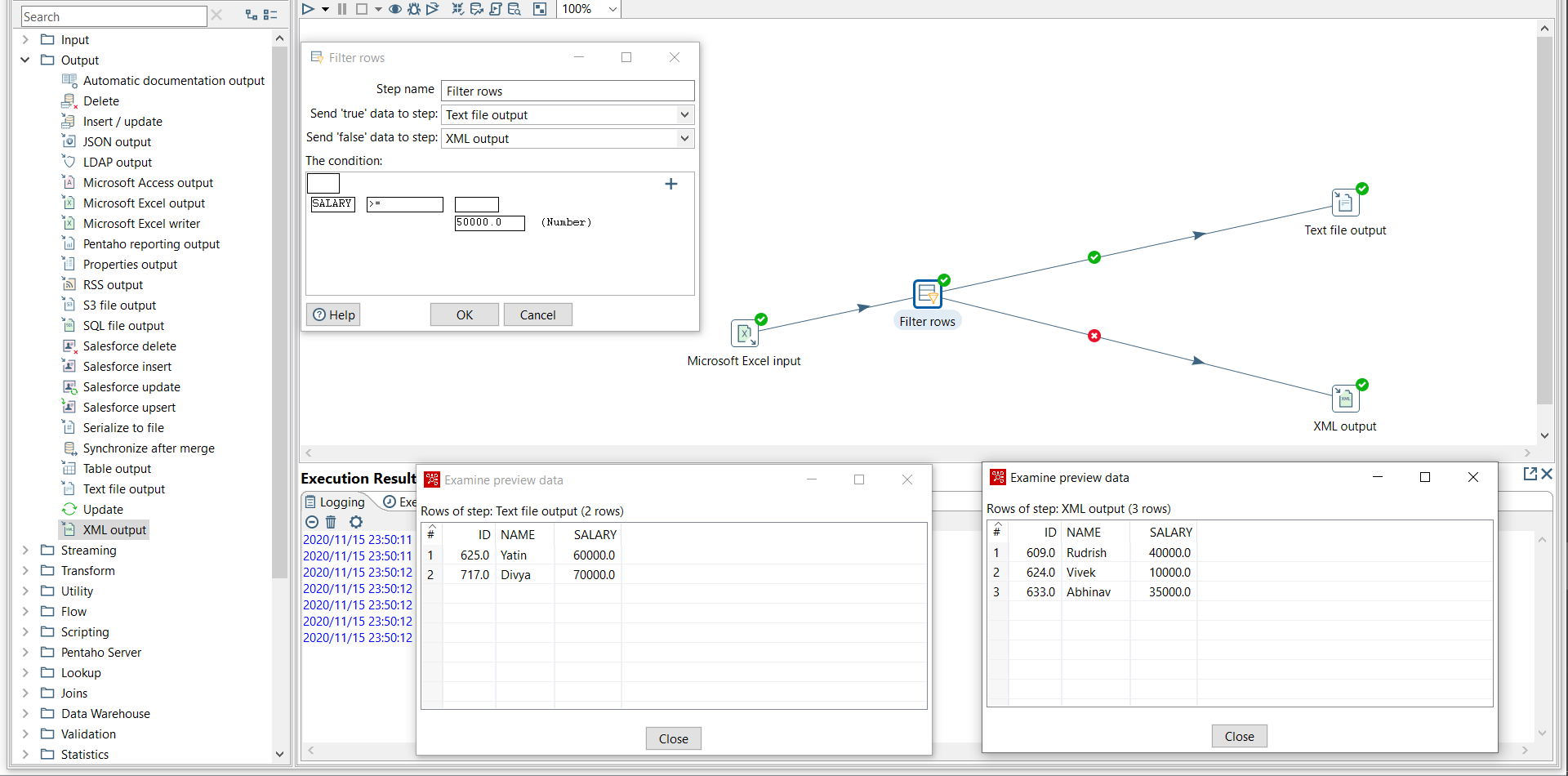
3) Drag and drop the Text output file and place in the area.

4) Add filter rows and divert the true and false result to the output files

5) Add Hop from input to output files.

6) Click on run and preview the output.

**Output:**

****

**b) Distribution of data from one excel file to text file and excel file.**

**Solution) Steps:**

1) Drag and drop the Excel input file and Excel output and Text Output.

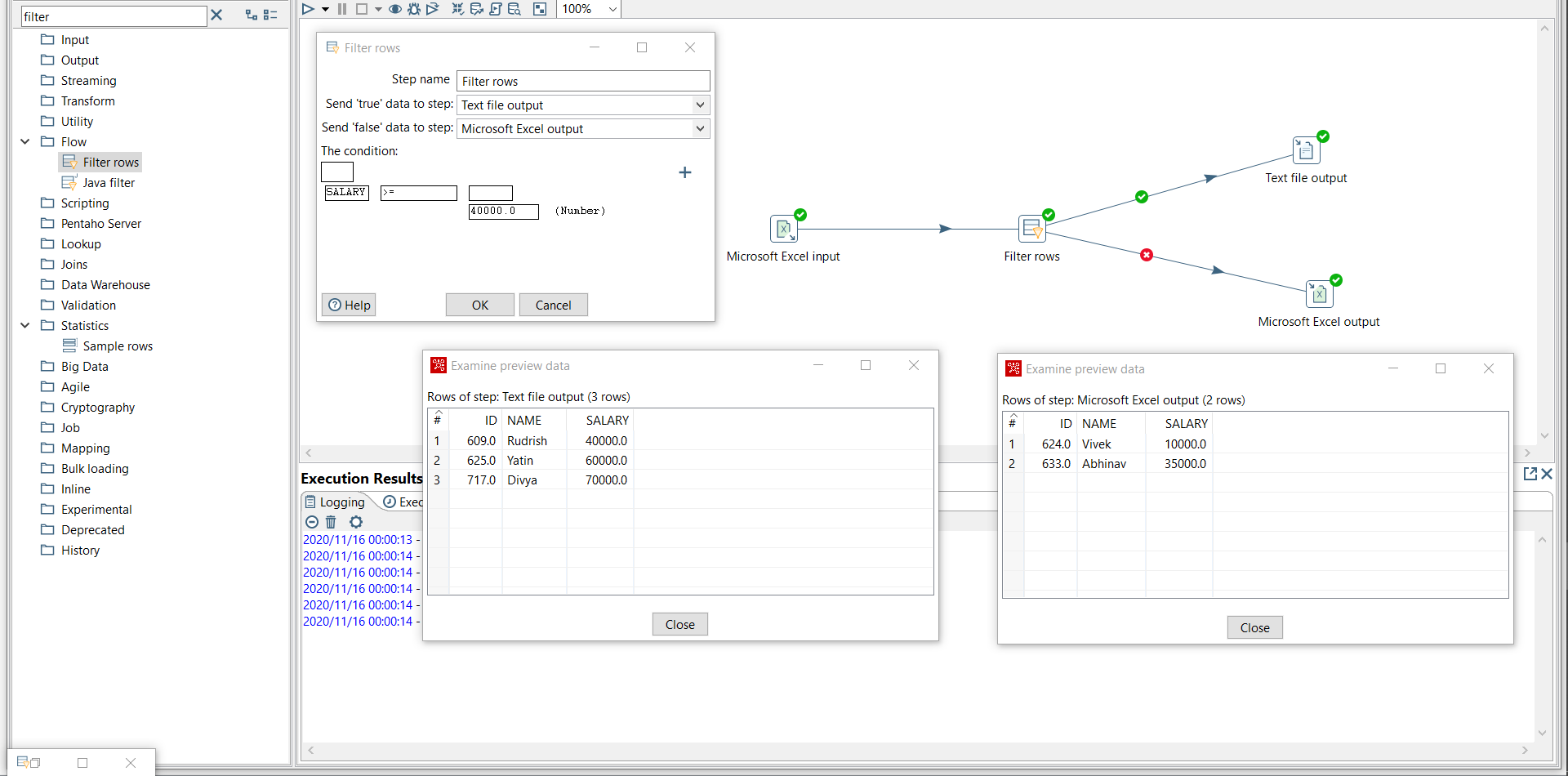
2) Add the Excel file in input , get fields and draw a hop to the files.

3) After adding hope to the output files, a message popup will appear, click on “Distribute” to each of the output files.

4) Now, Click on “Get fields” in both the outputs. Click on save.

5) Run the transformation and see the output.

**Output:**

****

**c) Distribute data from one Text file to multiple Excel files.**

**Solution) Steps:**

1) Drag and drop the Text input file and Excel output files.

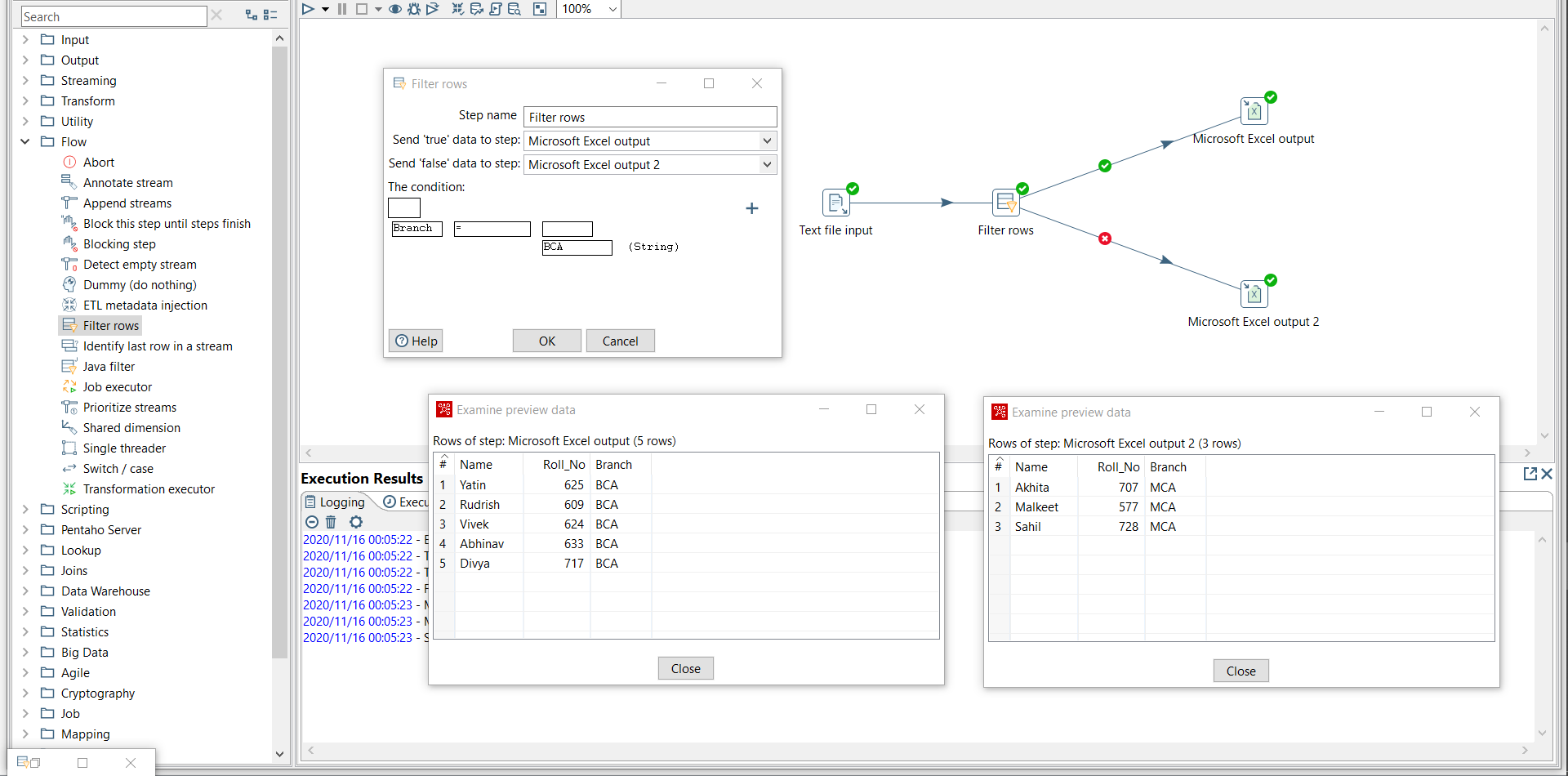
2) Add the Text file in input , get fields and draw a hop to the files.

3) After adding hope to the output files, a message popup will appear, click on “Distribute” to each of the output files.

4) Now, Click on “Get fields” in both the outputs. Click on save.

5) Run the transformation and see the output.

**Output:**



**Experiment 9:** Create Database connection in PantaHo with MySQL and Create a Simple PDI Transformation for transforming data from:

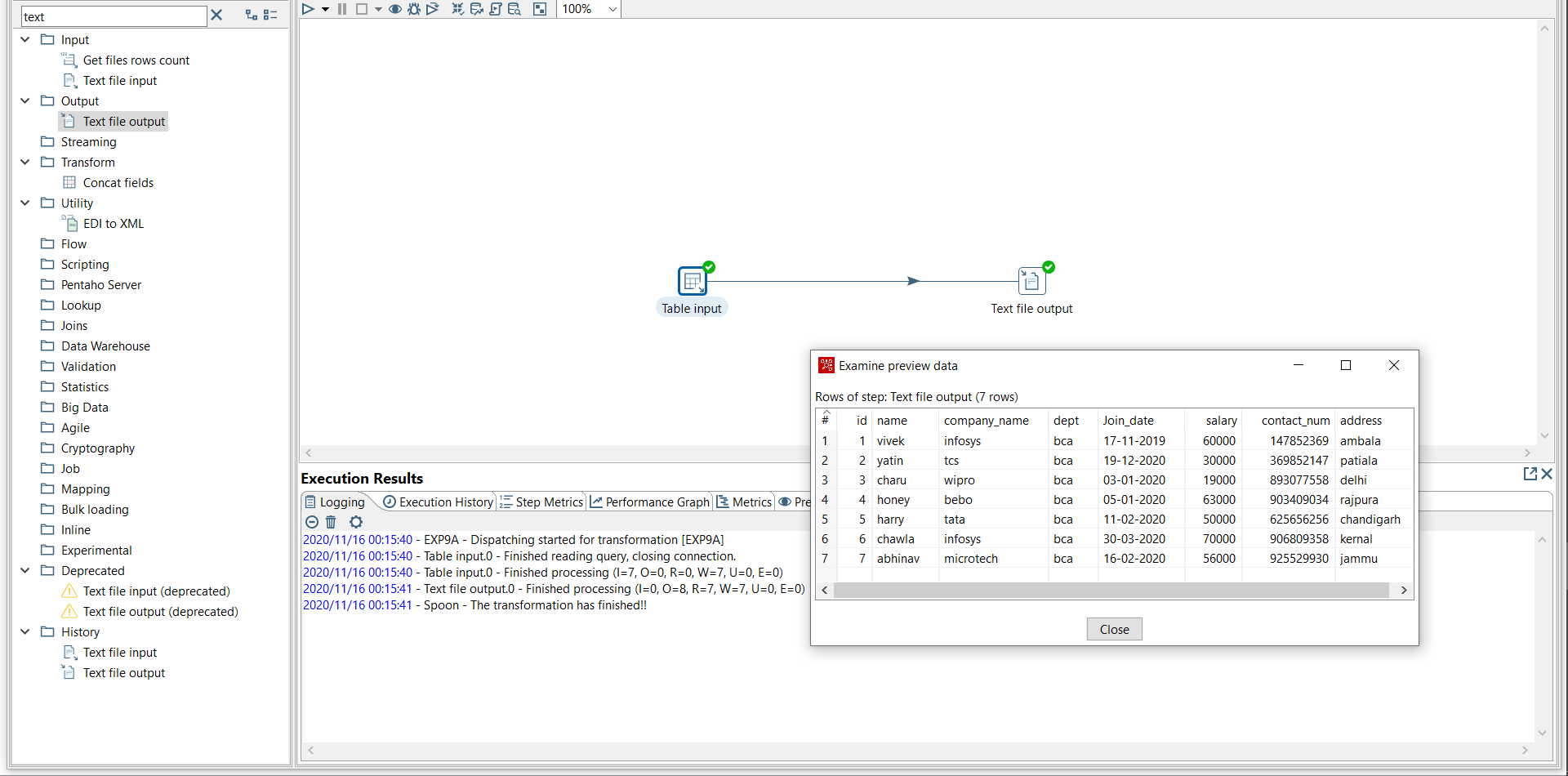
1. From Database table to Text File.
2. From Database table to Excel file

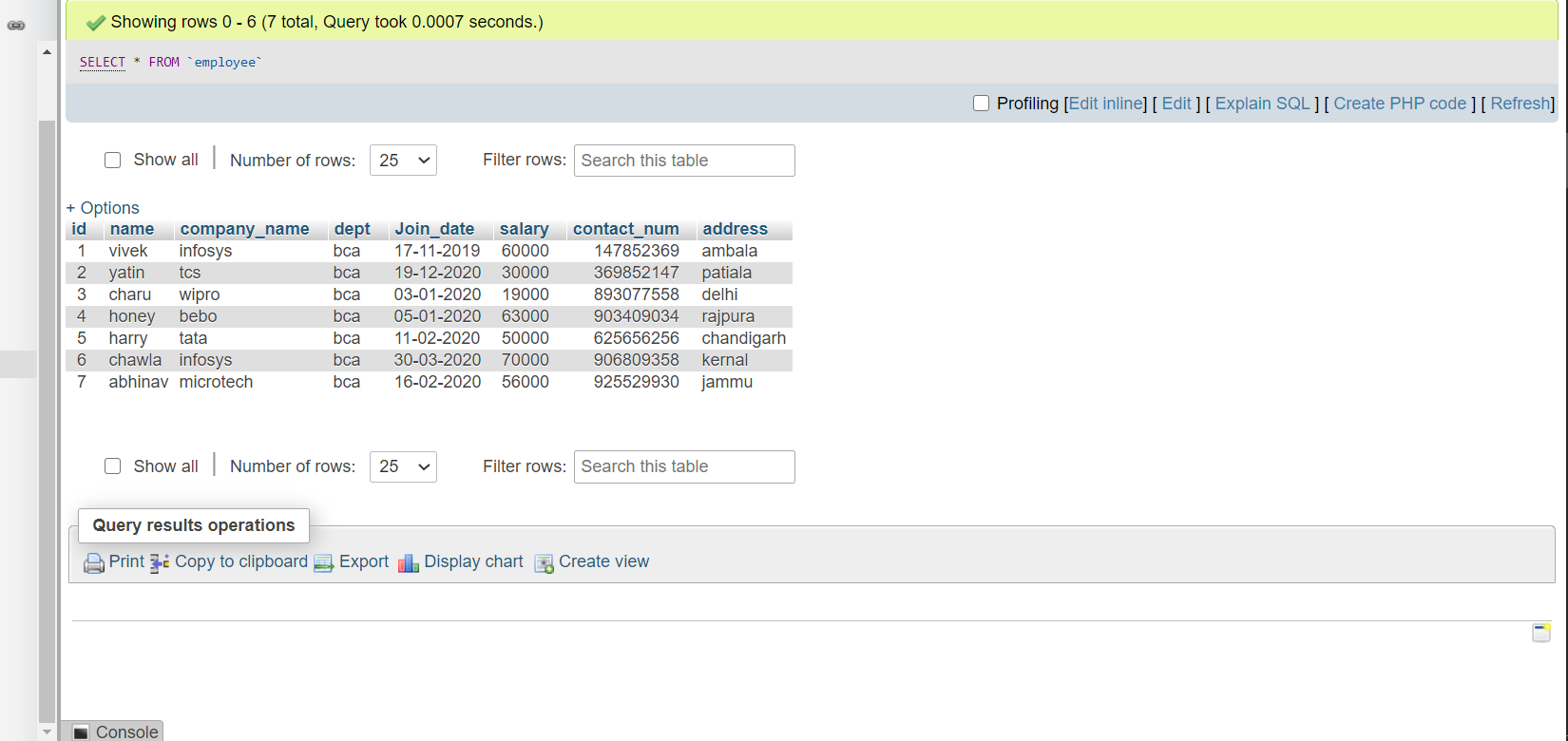
From Access to Excel file.

**Solution)**

a.

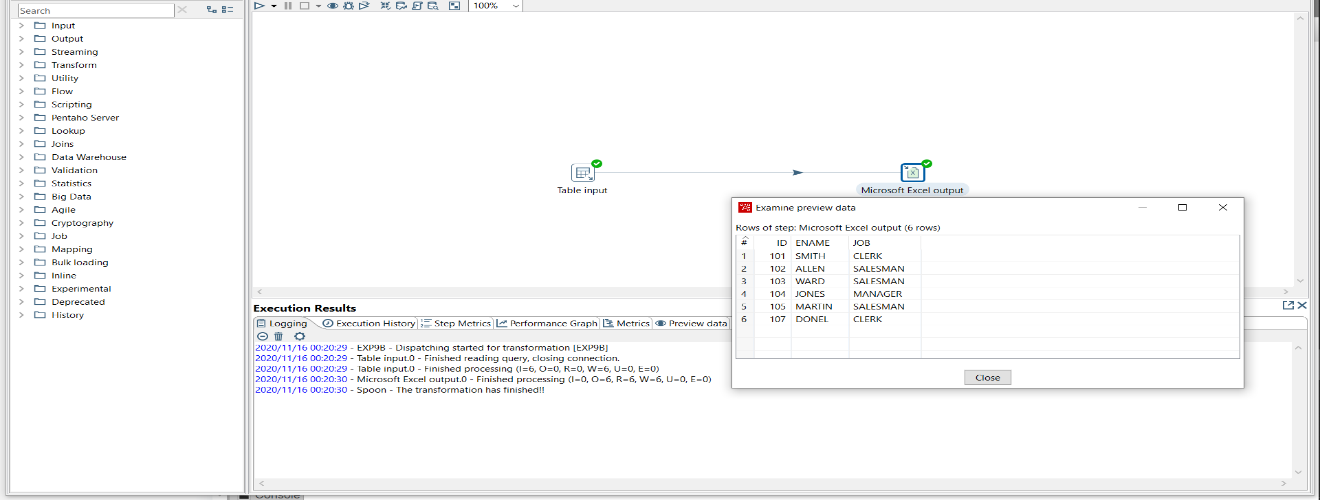
1. Drag and drop the Database table input and Output text file.
2. Double click on table input then, click on connection name (new).
3. Select database type , enter appropriate database name and provide additional info.
4. Test weather the connection is established or not.
5. Now , click on setup wizard and select the table that needs to be converted into Text file.
6. Draw a hop from Database table to Text File, get desired fields by double clicking on text file output and get fields , save location.
7. Run the transformation and preview the data, if any error occurs check the connection /drivers.





b.

1. Drag and Drop the Database table input and Output excel file.
2. Double click on table input then, click on connection name (new).
3. Select database type , enter appropriate database name and provide additional info.
4. Test weather the connection is established or not.
5. Now , click on setup wizard and select the table that needs to be converted into excel file.
6. Draw a hop from Database table to excel File, get desired fields by double clicking on excel file output and get fields, save location.
7. Run the transformation and preview the data, if any error occurs check the connection /drivers.



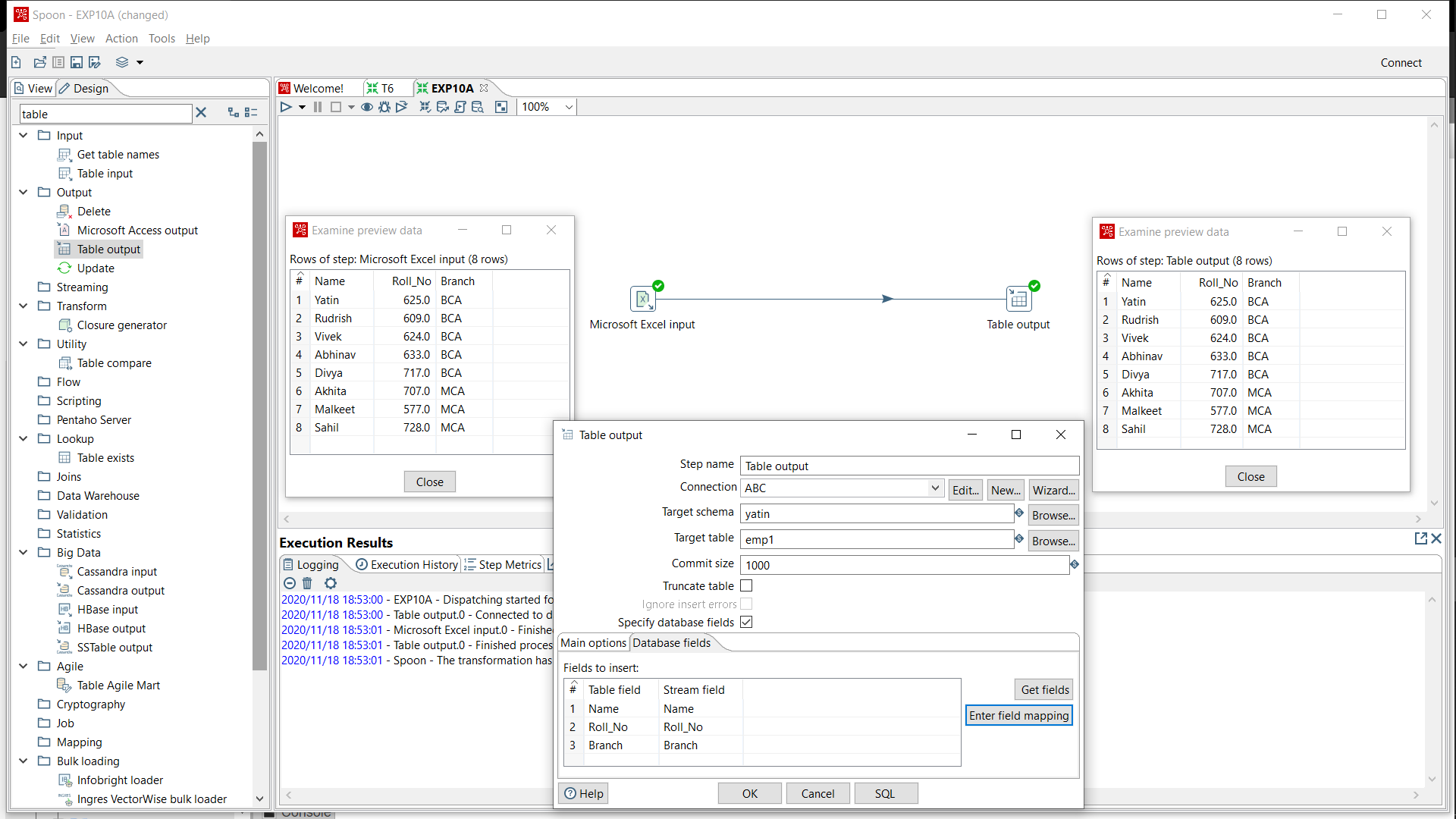
**Experiment 10:** Create a Simple PDI Transformation for transforming data from:

1. From Excel file to Database table.
2. From Text file to Database table.

**Solution)**

a.

1. Drag and Drop the excel file input and Output database table.
2. Double click on excel input, and add the data source file in them and get sheets fields and then preview rows .
3. Draw a hop from excel input to table output .
4. Select database type , enter appropriate database name and provide additional info.
5. Test weather the connection is established or not.
6. Now , select the table and get fields and perform mapping between excel and table fields.
7. Run the transformation and preview the data, if any error occurs check the connection /drivers.



b.

1. Drag and Drop the text file input and Output database table.
2. Double click on text input, and add the data source file in them and get fields and then preview rows .
3. Draw a hop from text input to table output .
4. Select database type , enter appropriate database name and provide additional info.
5. Test weather the connection is established or not.
6. Now , select the table and get fields and perform mapping between text and table fields.
7. Run the transformation and preview the data, if any error occurs check the connection /drivers.

