KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20

T.Y. B.Sc. I.T. Semester VI

INDEX

Sr. No.	Topic	Date	Signature
Assignment			
1	Assignment 1	17|11|22	
2	Assignment 2	01|12|22	
Tutorial			
1	Test 1	19|12|22	
SLE & Project Report			
1	Business Intelligence Applications in Clickstream Mining	06|01|23	
2	Clickstream Data Mining using Google Analytics	06|01|23	
Practical			
1A	Installation of Power BI Desktop	31|01|23	
1B	Import the legacy data from different sources such as Excel, SQL Server and OData Field and load into the target system	31|01|23	
2A	Perform the ETL process to construct the database in SQL Server	31|01|23	
2B	Perform the ETL process in Power BI	31|01|23	
3A	Perform data visualization in Power BI and create the data staging area for the selected database	08|02|23	
3B	Create the data staging area for the selected database using Star Schema	08|02|23	
4A	Create the cube with corresponding dimension and fact table based on OLAP	08|02|23	
4B	Perform Multi-Dimensional Expressions (MDX) queries for OLAP database	08|02|23	

KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20

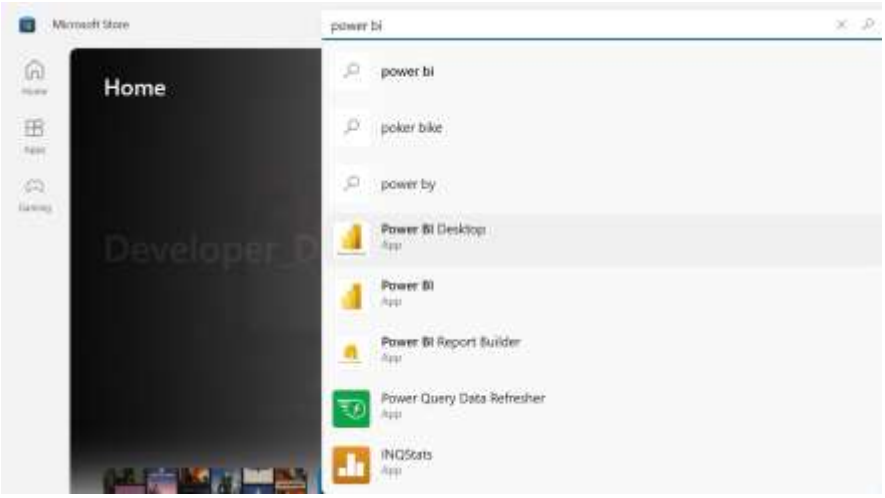
T.Y. B.Sc. I.T. Semester VI

Sr. No.	Topic	Date	Signature
5A	Import the data warehouse in Microsoft Excel and create Pivot Table to perform Data Analytics	13|02|23	
5B	Import the cube in Microsoft Excel and create the Pivot Chart to perform Data Modelling	13|02|23	
6	Perform the What-If Analysis for Data Visualization	13|02|23	
7	Data Analysis using Time Series Analysis	13|02|23	
8	Implementation of Decision Tree using R Tool	13|02|23	
9	K – Means Clustering using R Programming	13|02|23	
10	Prediction using Linear Regression	13|02|23	

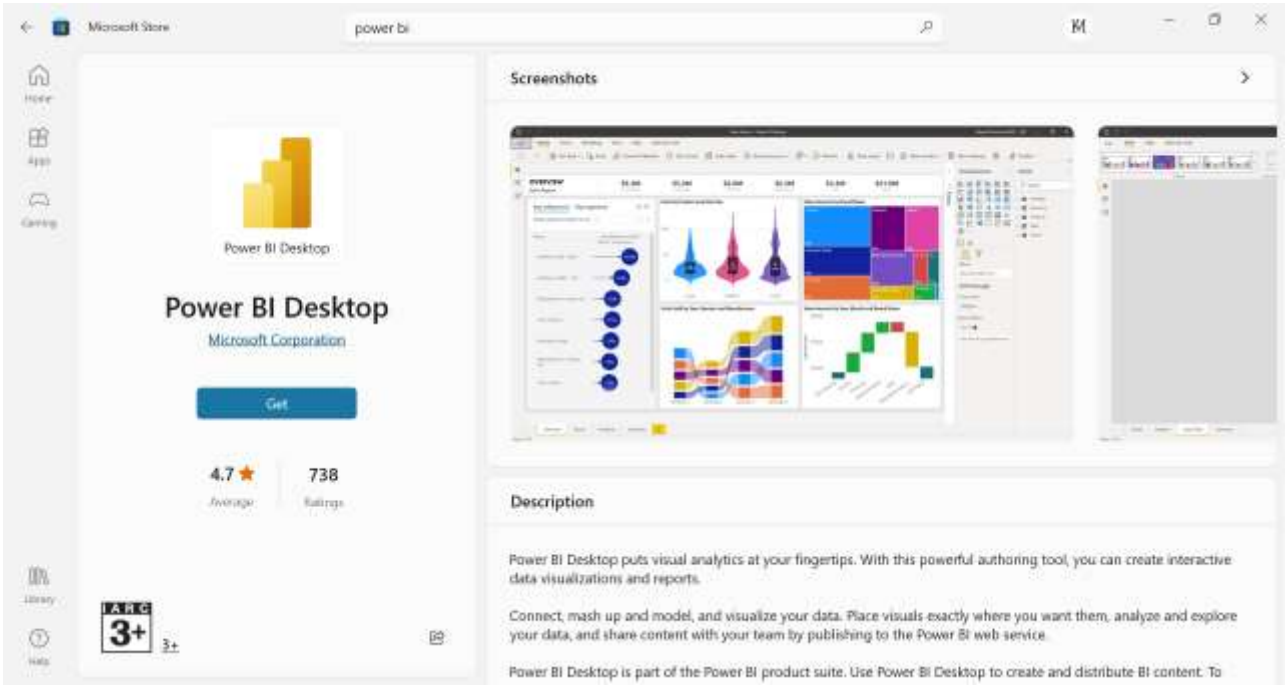
KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

Practical 1A – Installation of Power BI Desktop

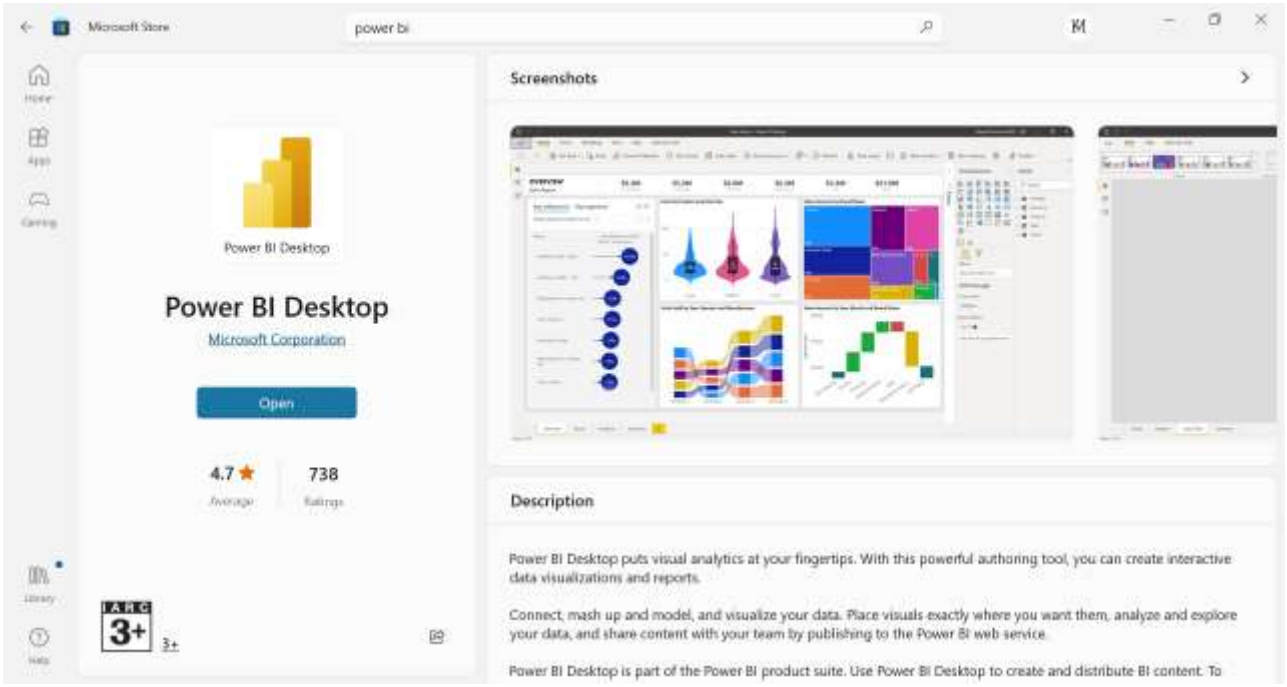
Go to Microsoft Store and search for power bi, click on Power BI Desktop.



Click on Get button.



The installation will start. Click on Open once the installation is complete and start working.

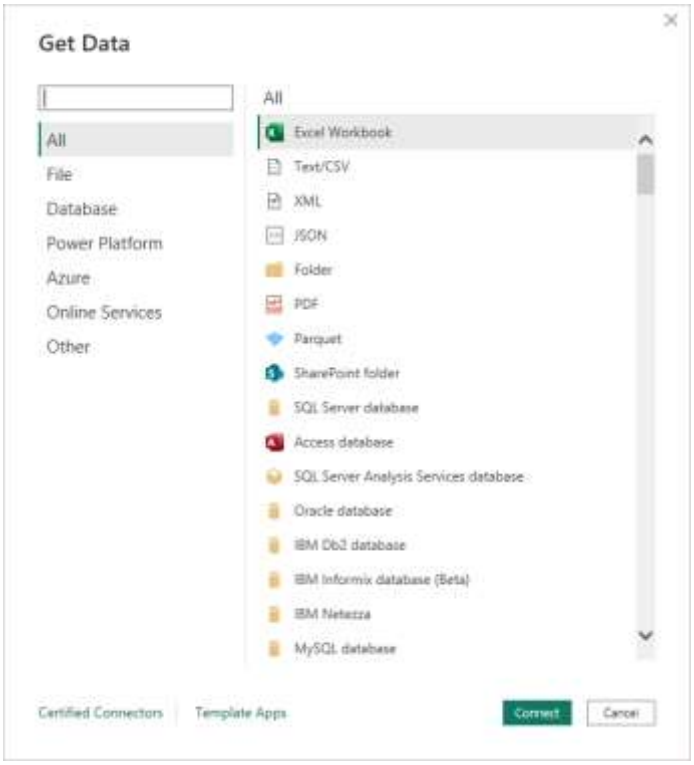


KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

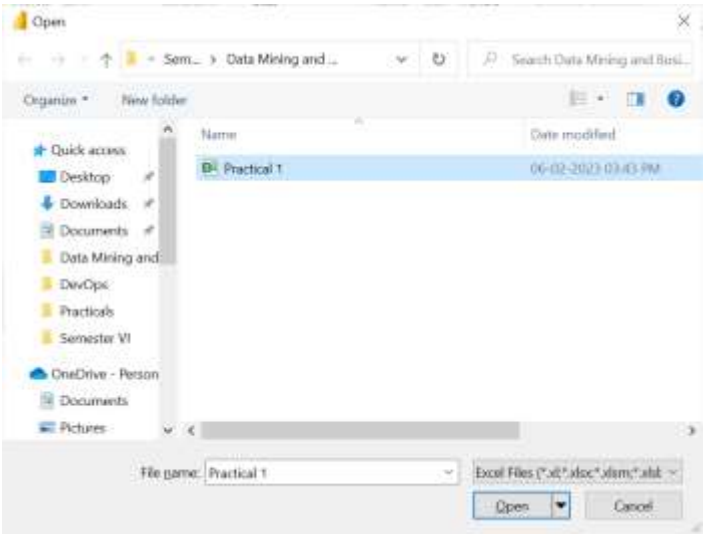
Practical 1B – Import the legacy data from different sources such as Excel, SQL Server and OData Field and load into the target system

Open Microsoft Power BI.

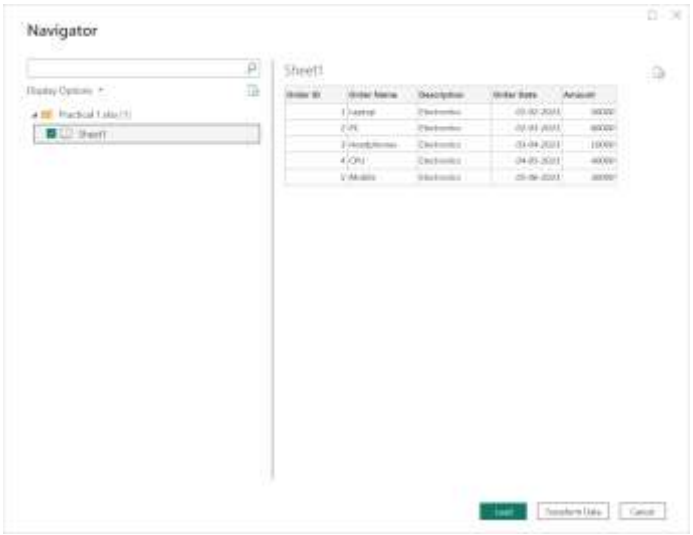
Select Excel Workbook under the Get Data option. Click on Connect.



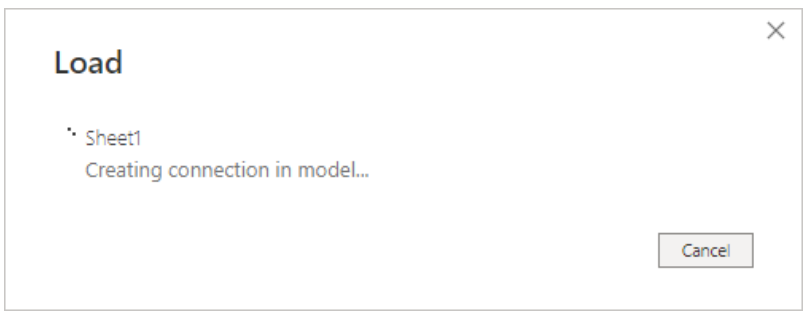
Browse your file and select it.



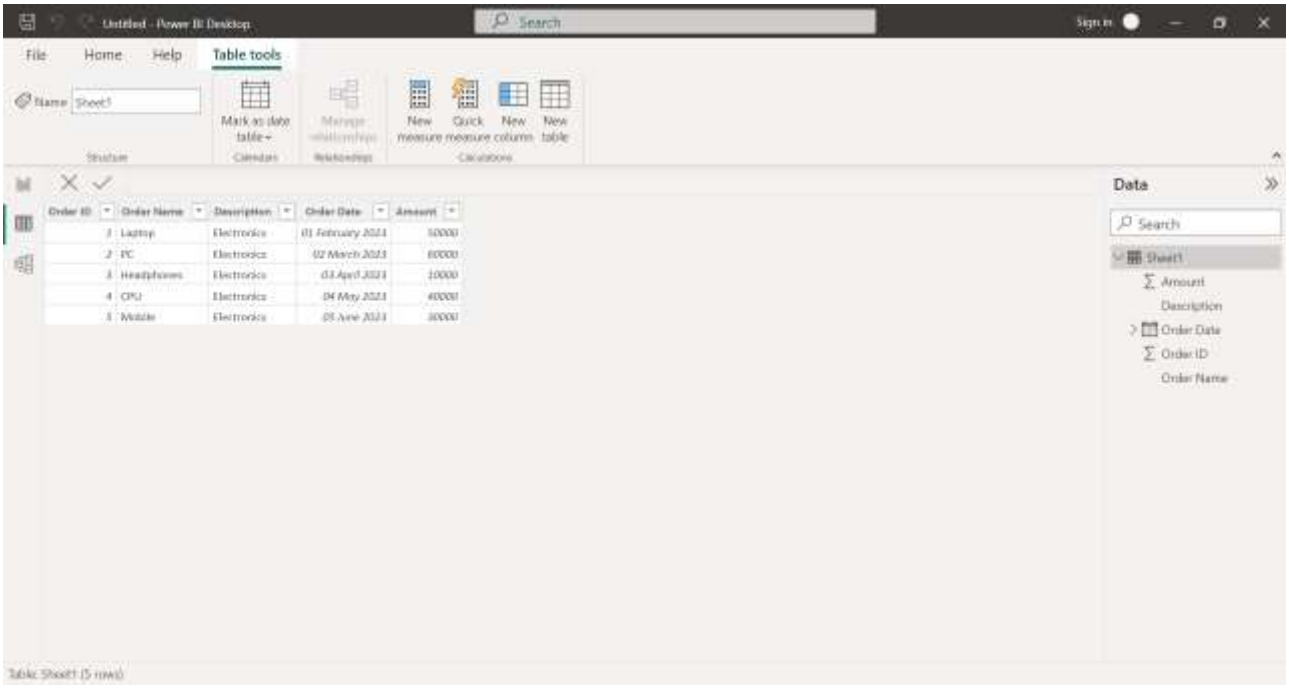
Select the sheet and click on Load once the table is displayed.



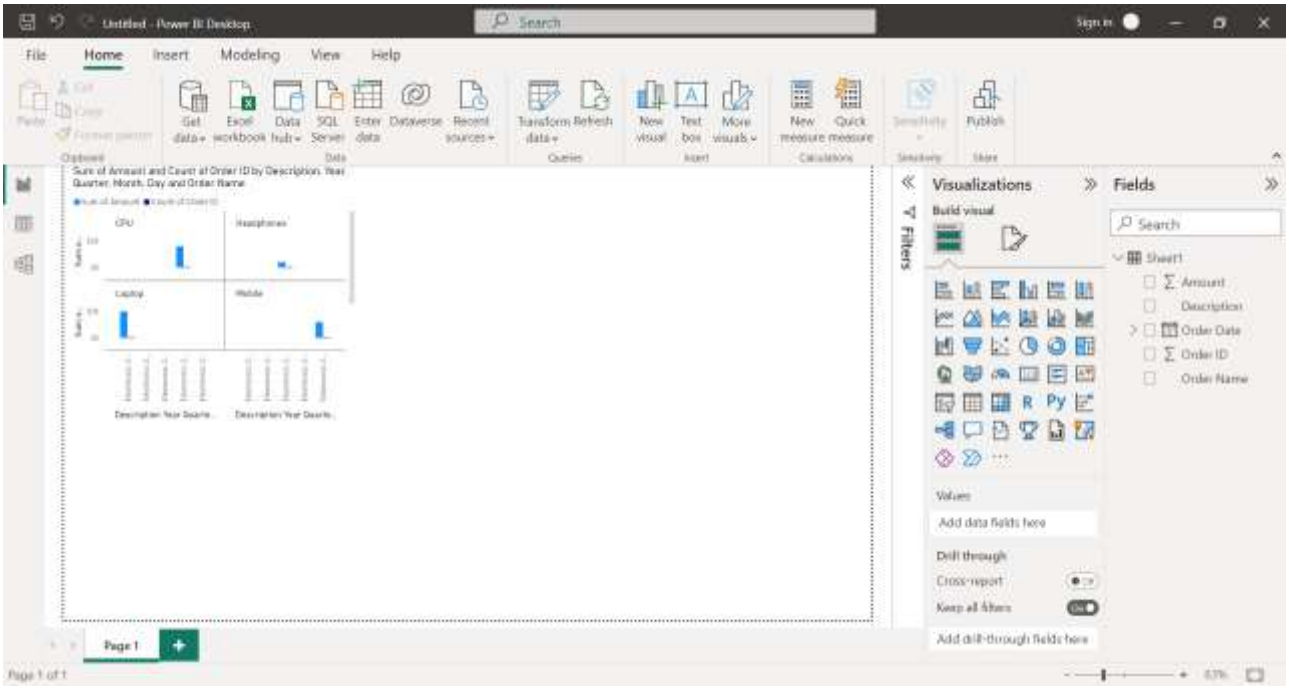
Wait for the loading to complete.



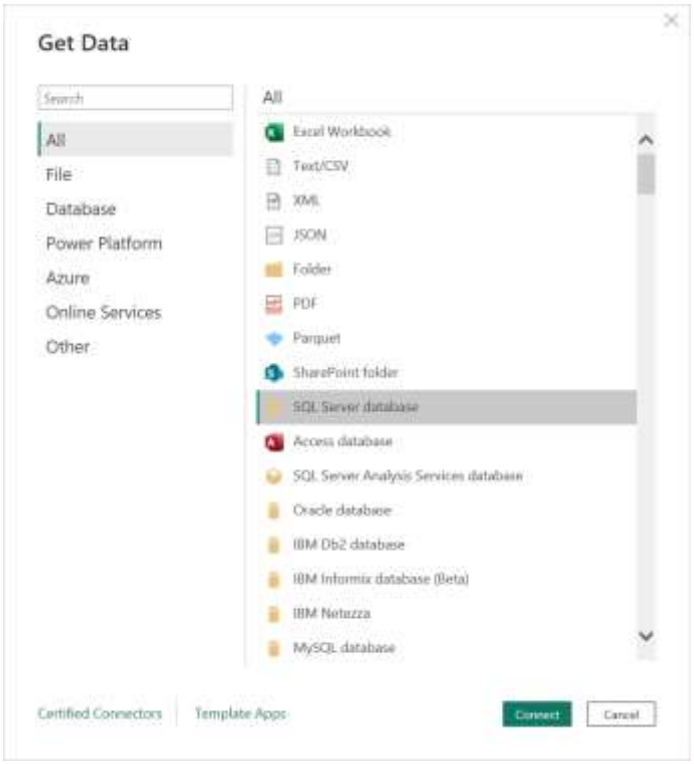
The data will be shown on the screen once the loading is complete.



Click on the Charts icon to display the graphical representation of the data.

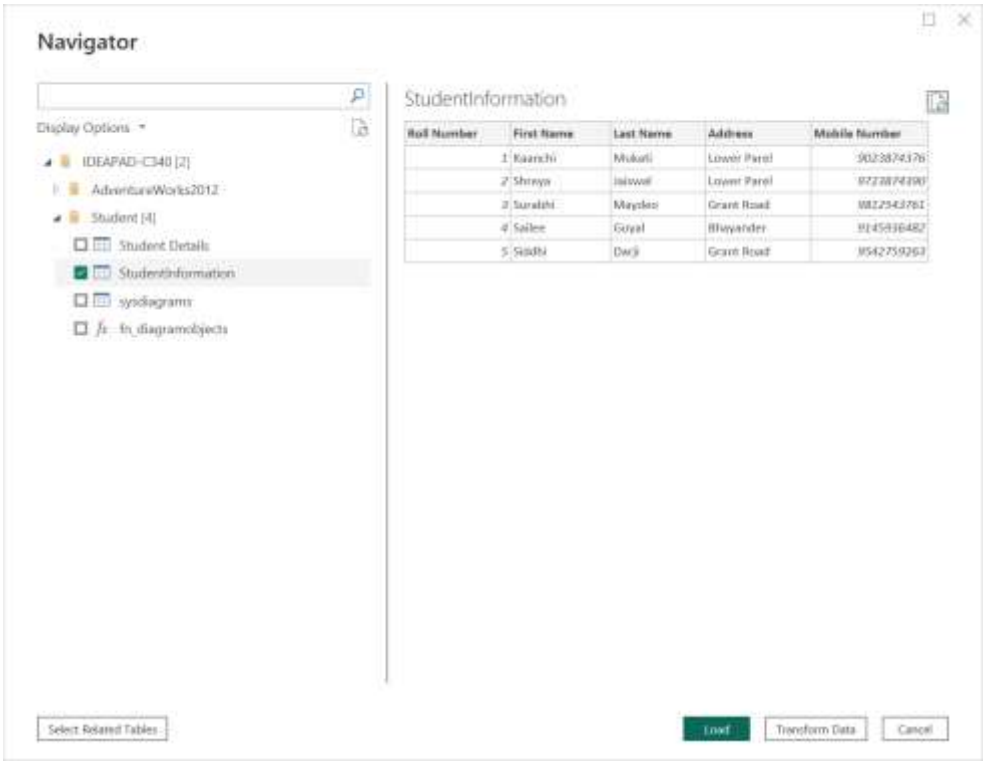


Click on Get Data and select SQL Server Database.



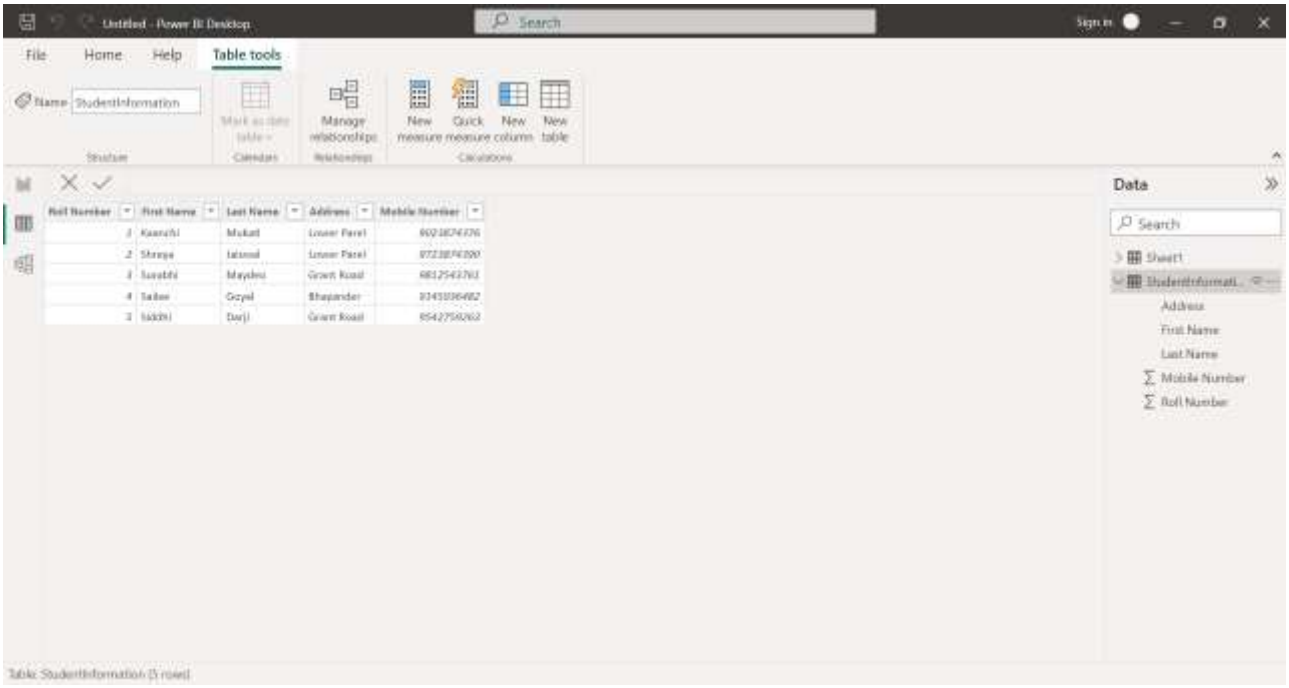
Copy the server name from Microsoft SQL Server Management Studio and paste it where the server is asked for connection between the two apps. Select Import and click on OK.

Select the table you wish to go ahead with and load it.

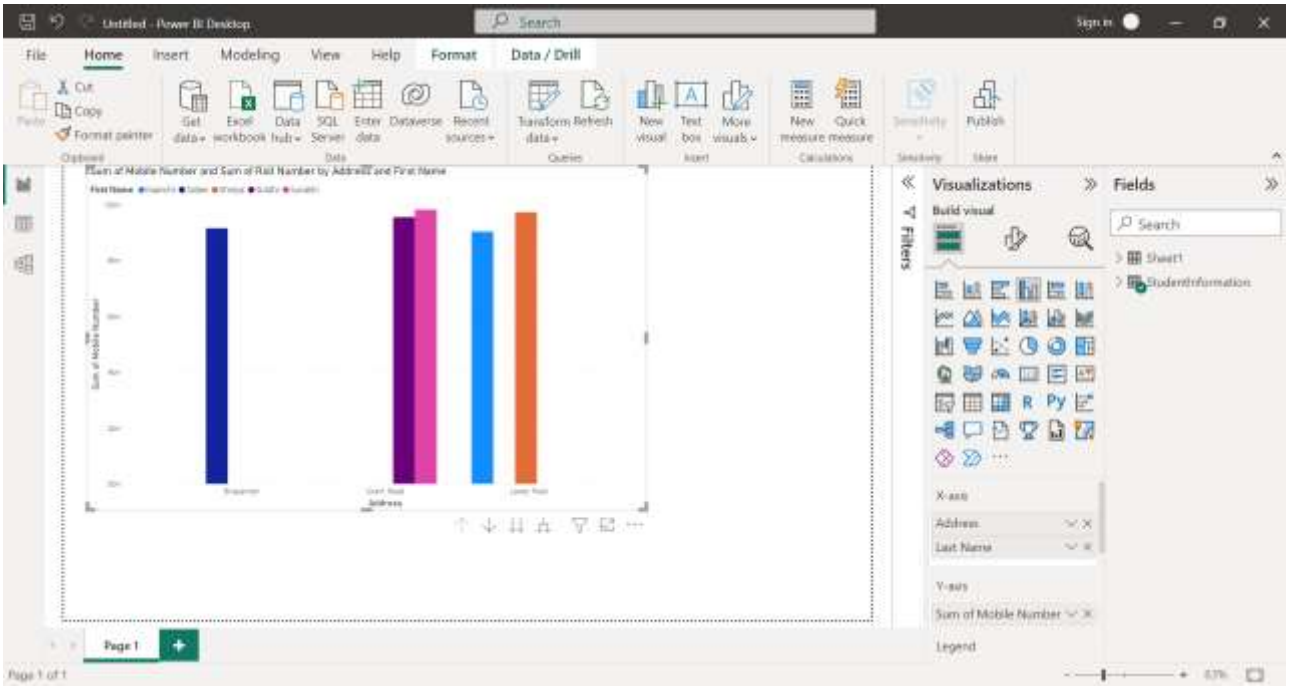


Once the data is loaded the following dialogue box will be displayed.

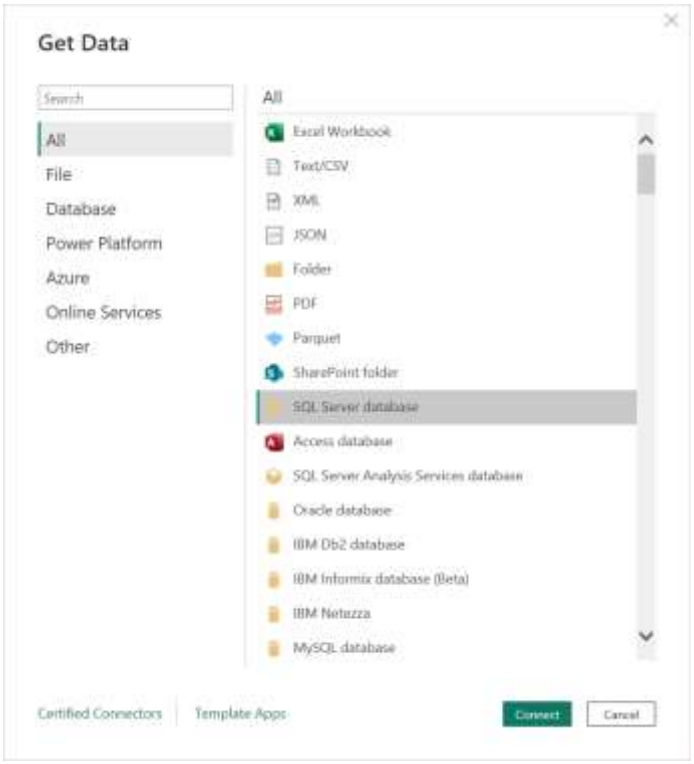
The data will be displayed on the screen.



On clicking the charts icon the graphical representation of the data will be displayed.

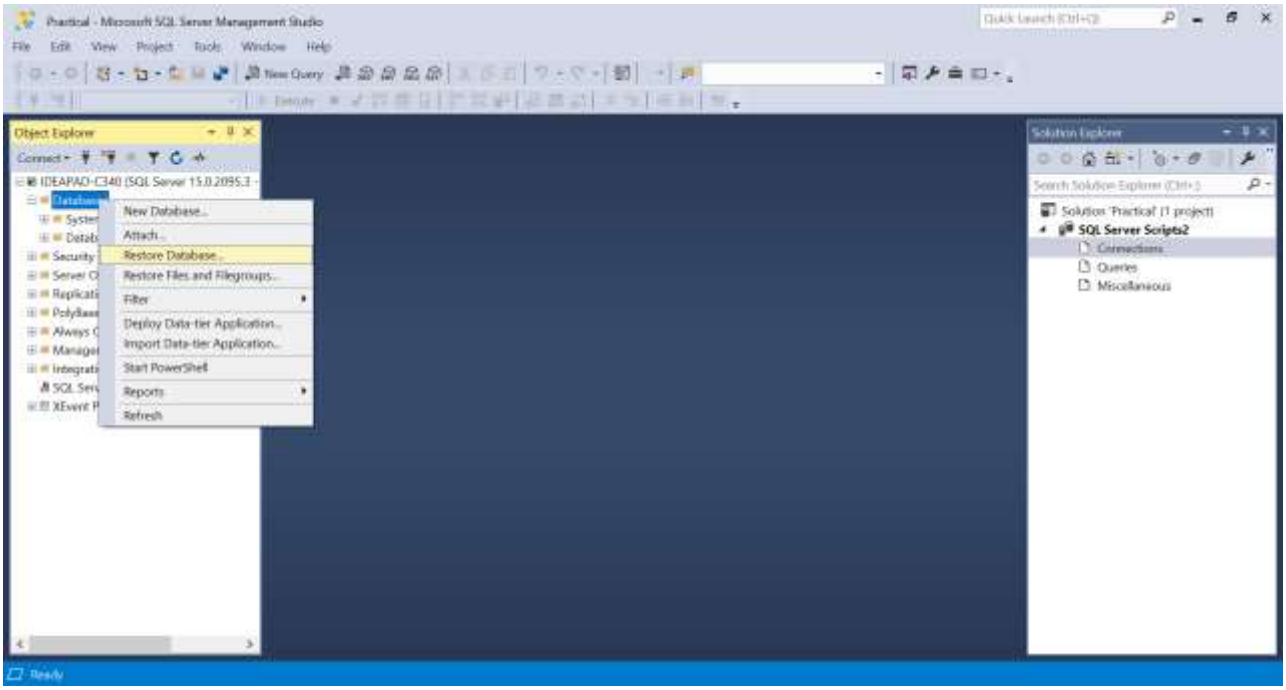


Select SQL Server Database under Get Data and connect.

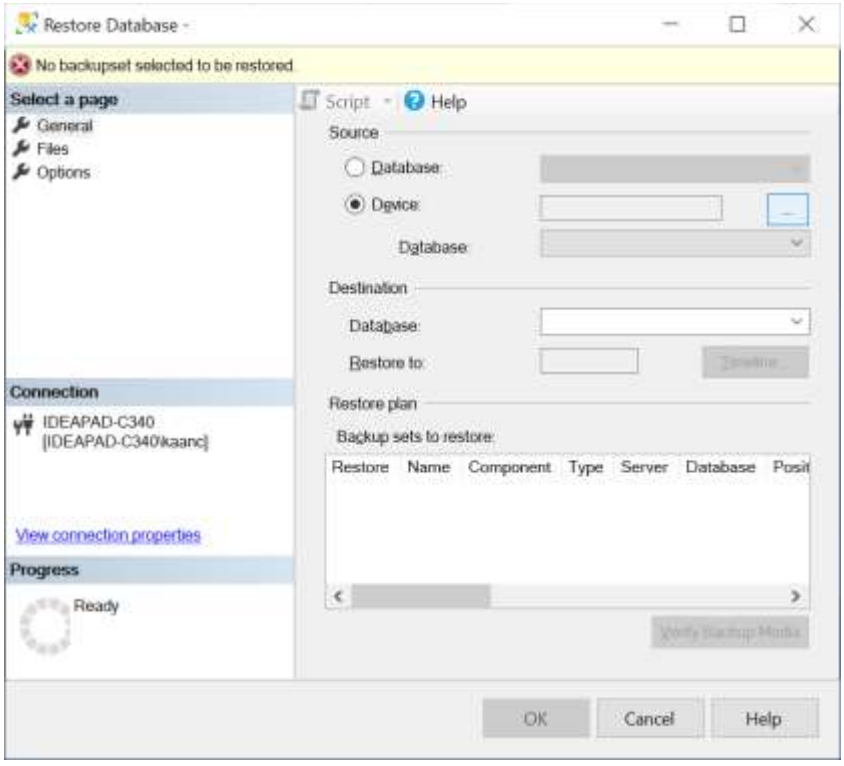


Enter the server name and the file name with .bak extension. Click on Import and OK.

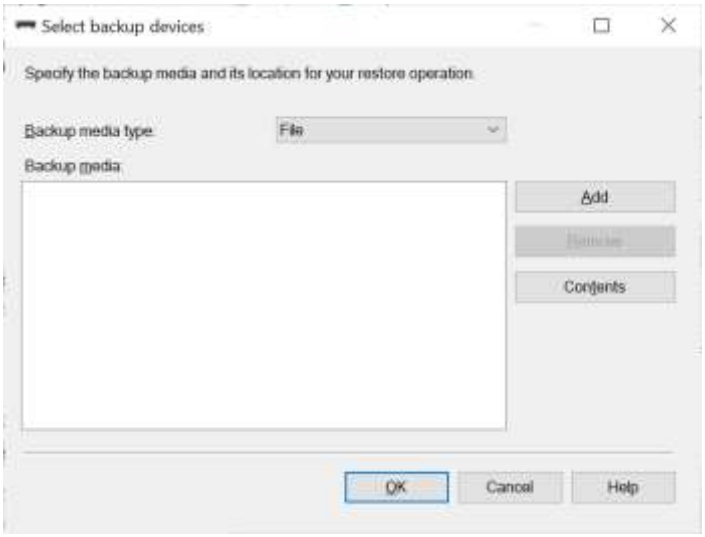
Open Microsoft SQL Server Management Studio. Right click on Databases and click on Restore Database…



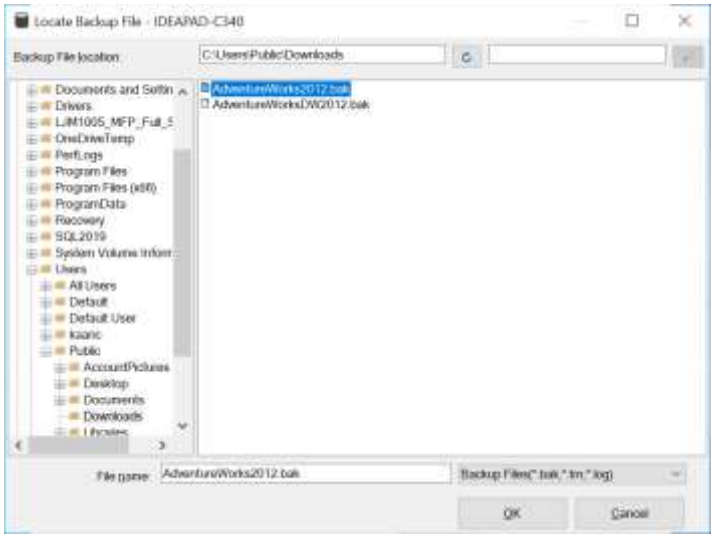
Select Device and click on the button next to it for additional steps.



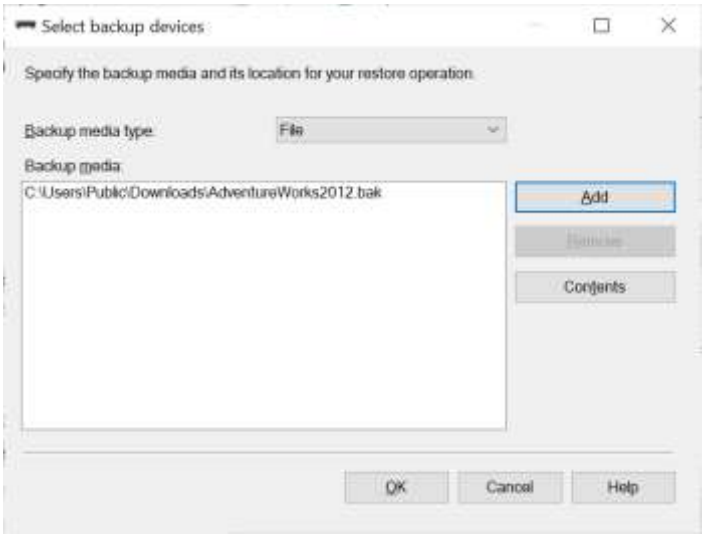
Click on Add.



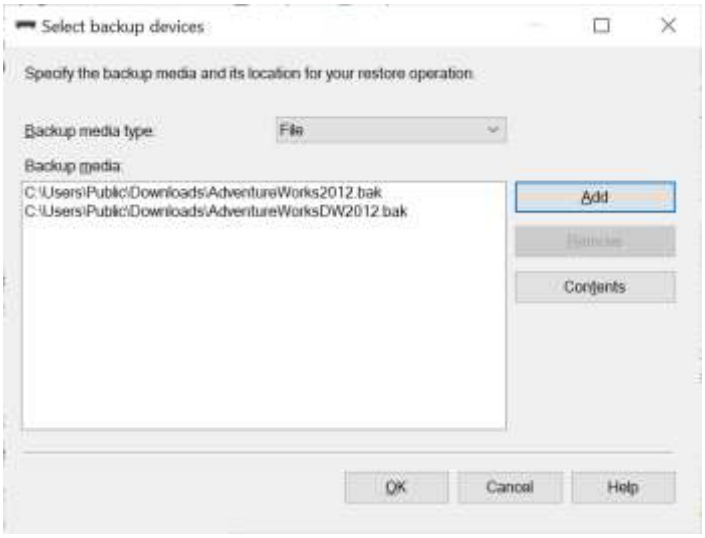
Search for your file with .bak extension and select it.



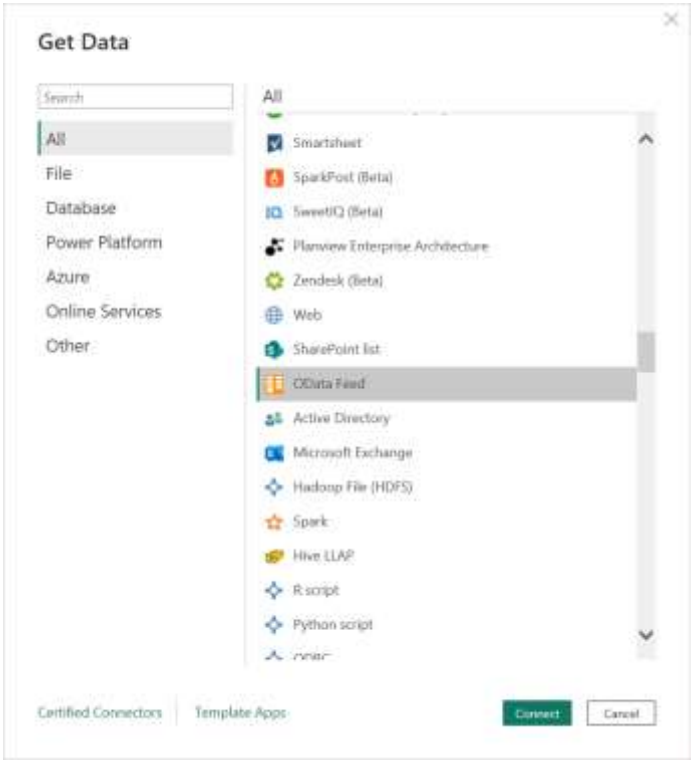
Click on Add.



Search for another file and add it as well.



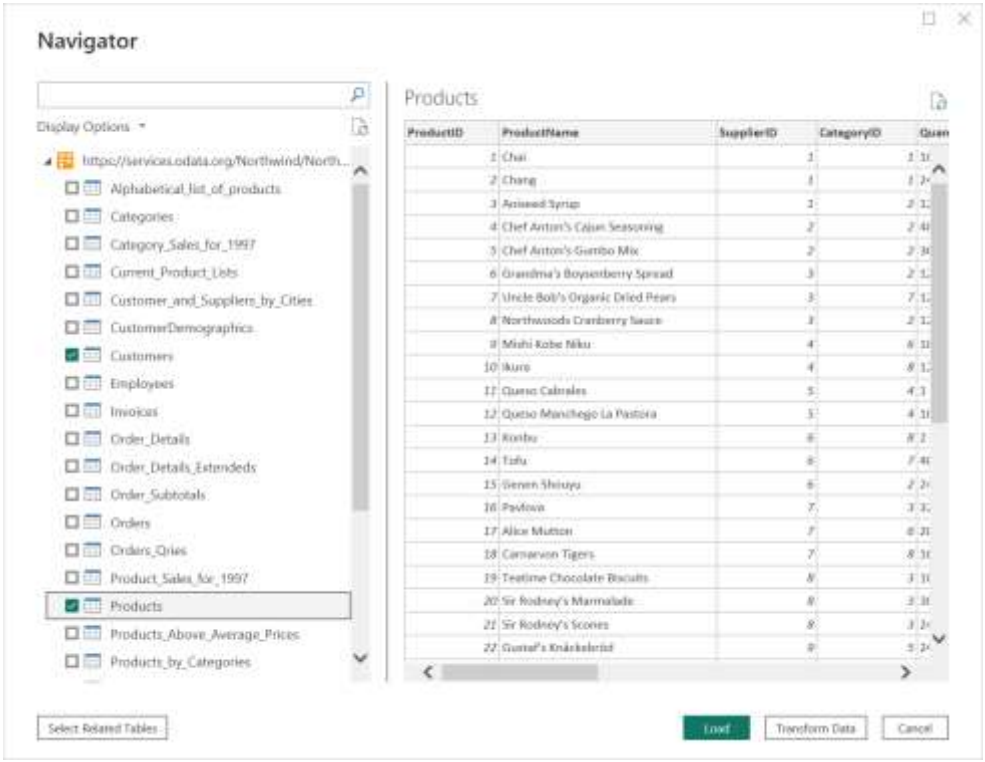
Get Data – O Data Feed



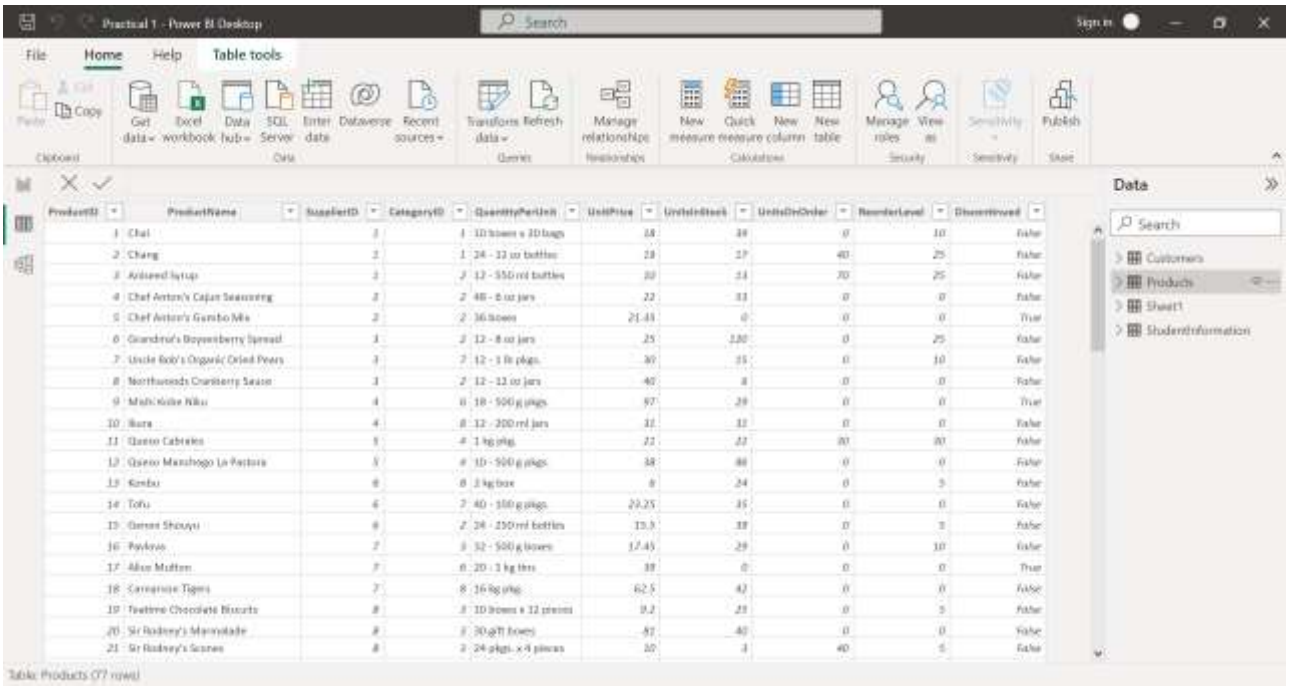
Paste the URL.

Select the services level to apply the settings to and connect.

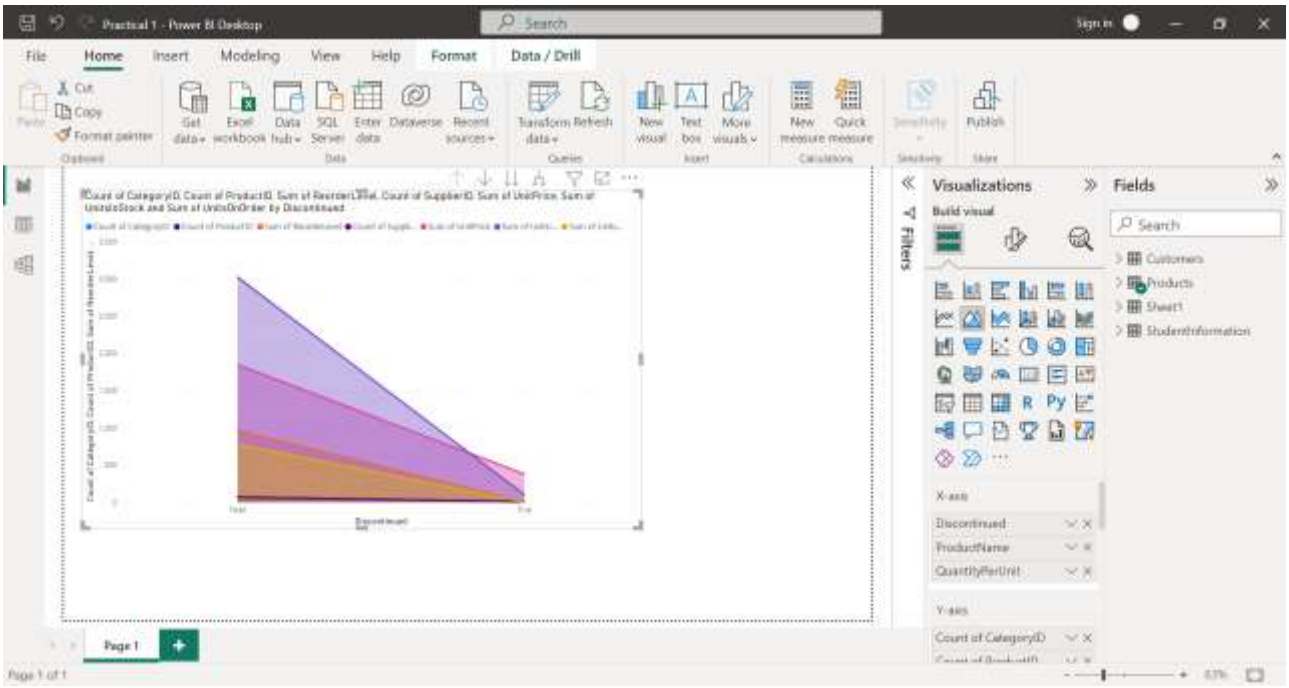
Select two tables that are Customers and Products table and load them.



Select the Products table.



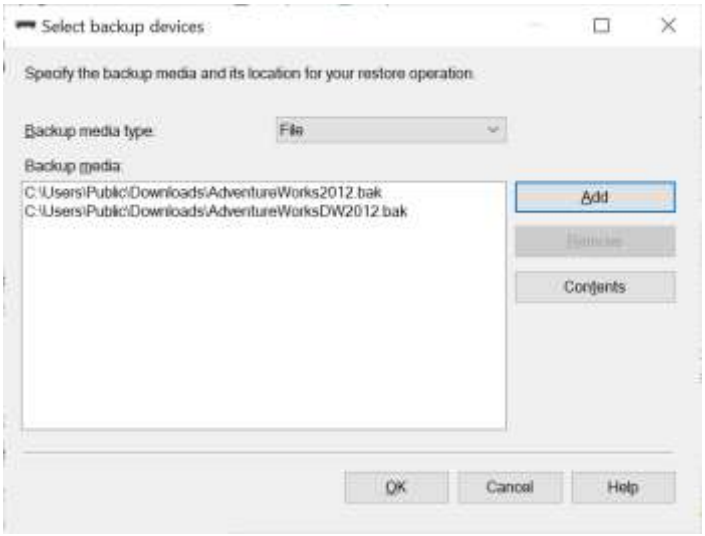
Click on the charts icon and select the type of graph to be displayed.



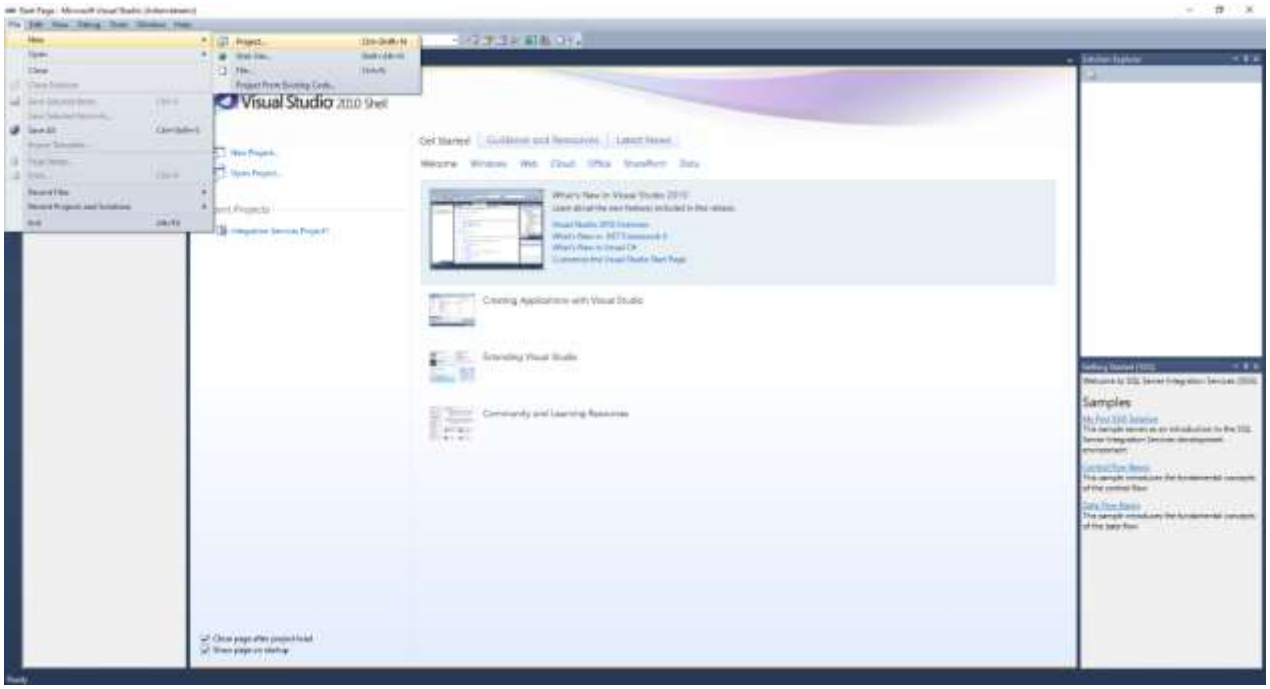
KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

Practical 2A - Perform the ETL process to construct the database in SQL Server

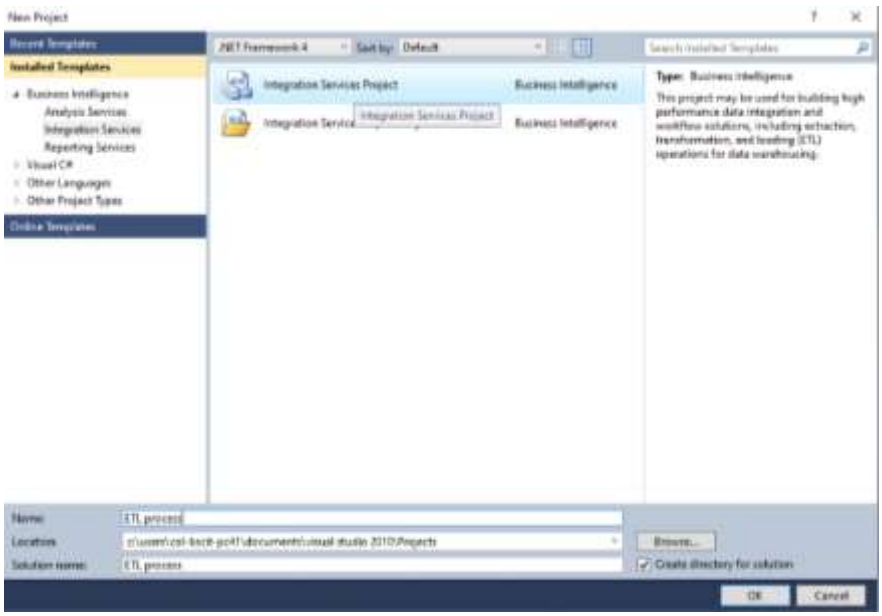
Add the AdventureWorks file in SQL Server Management Database.



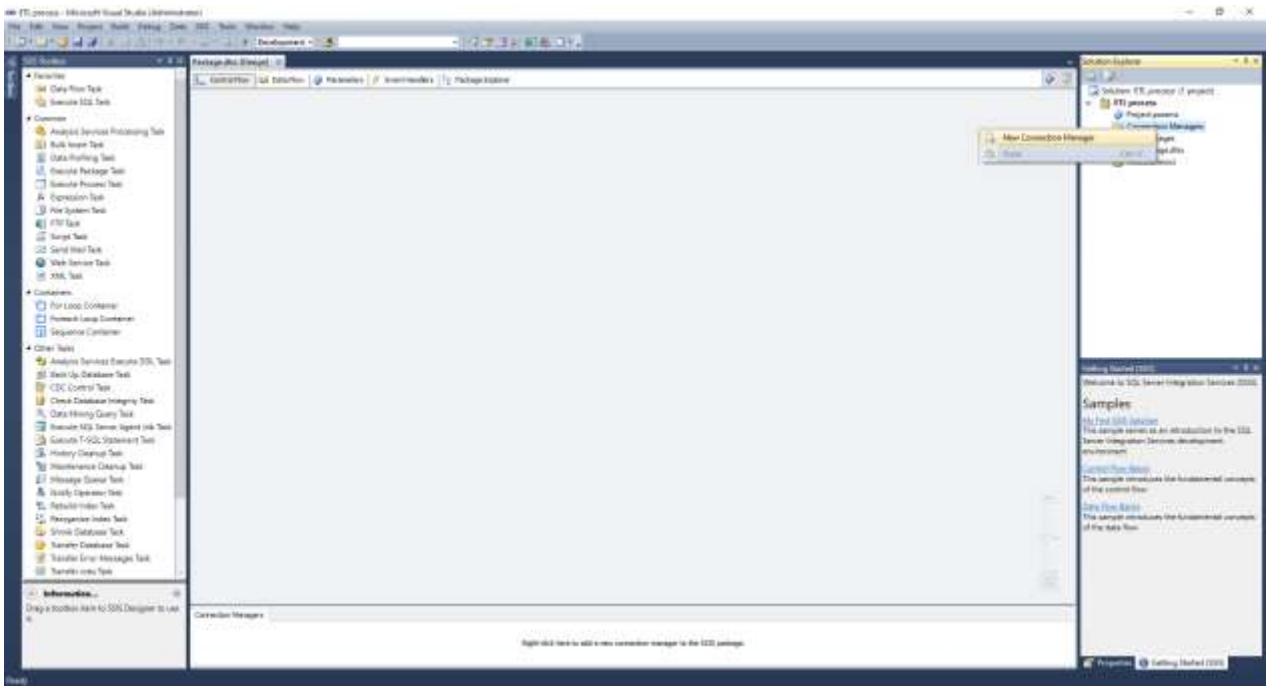
Open SQL Server Data Tools. Go to File then New and then Project



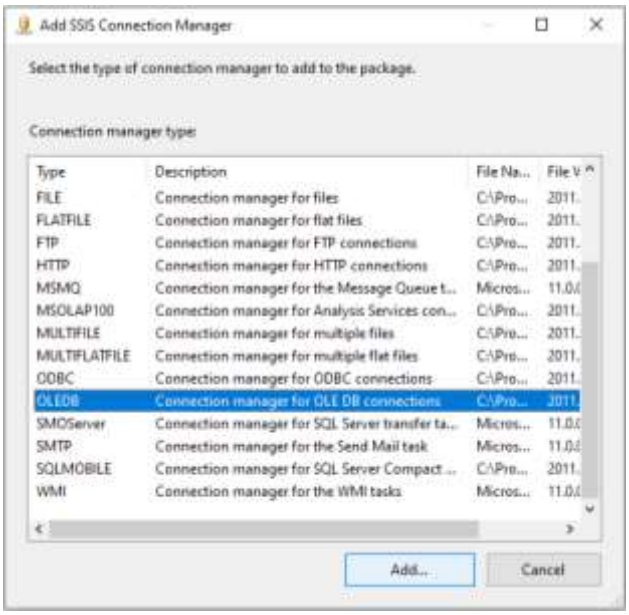
Select Integration Services and select Integration Services Project, rename it as ETL Process.



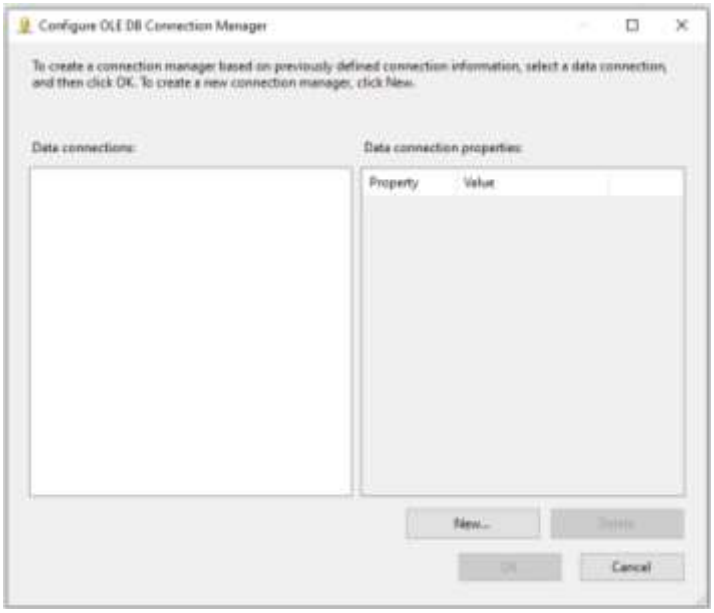
In the Solution Explorer right click on Connection Managers under ETL Process the file name and select the New Connection Manager.



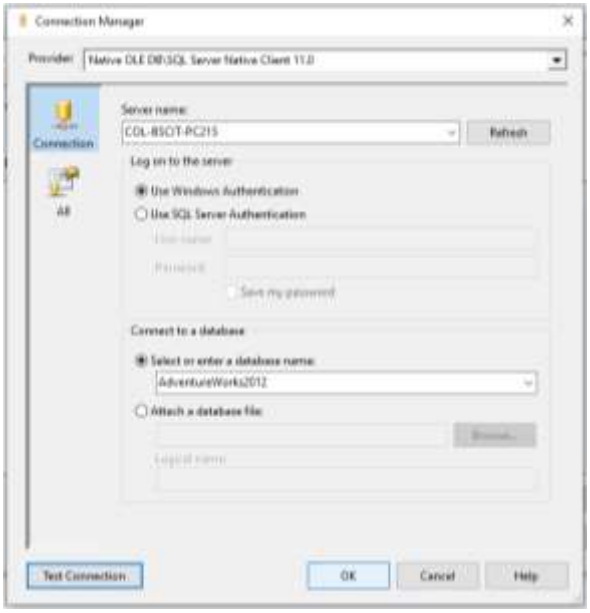
Add OLEDB as the Connection Manager.



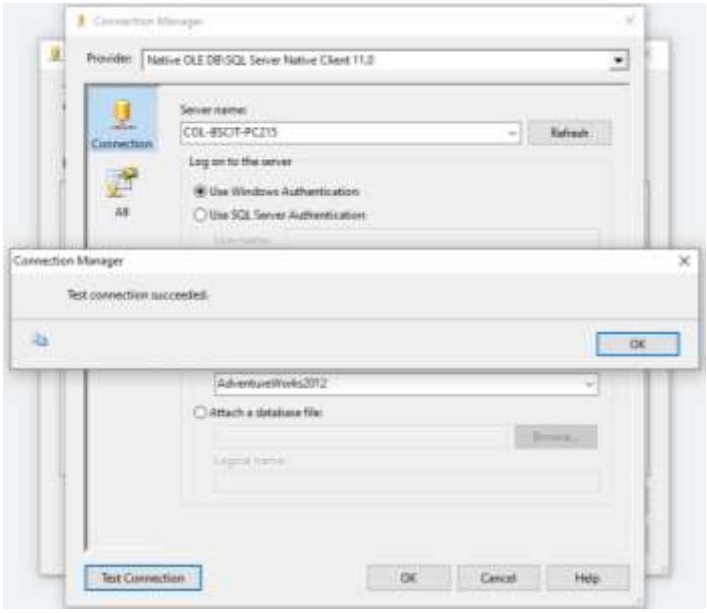
After adding the Connection Manager the following dialogue box will open. Click on New.



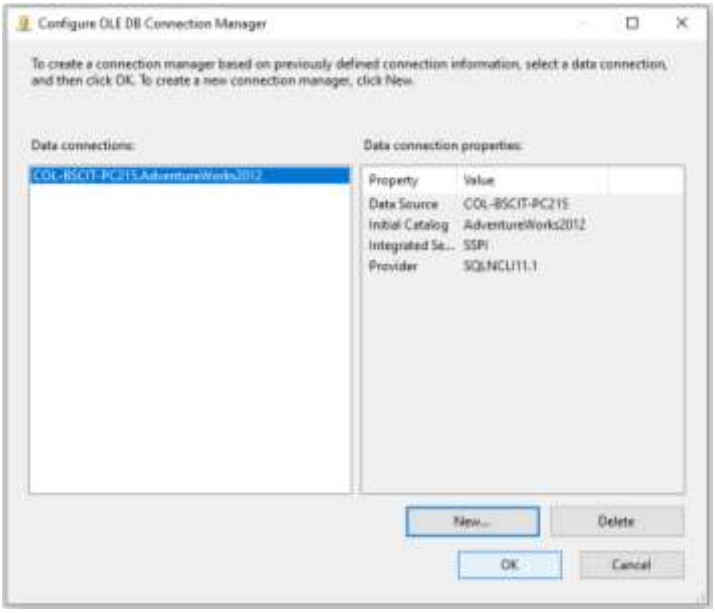
Enter the Server Name and select the database that is AdventureWorks file.



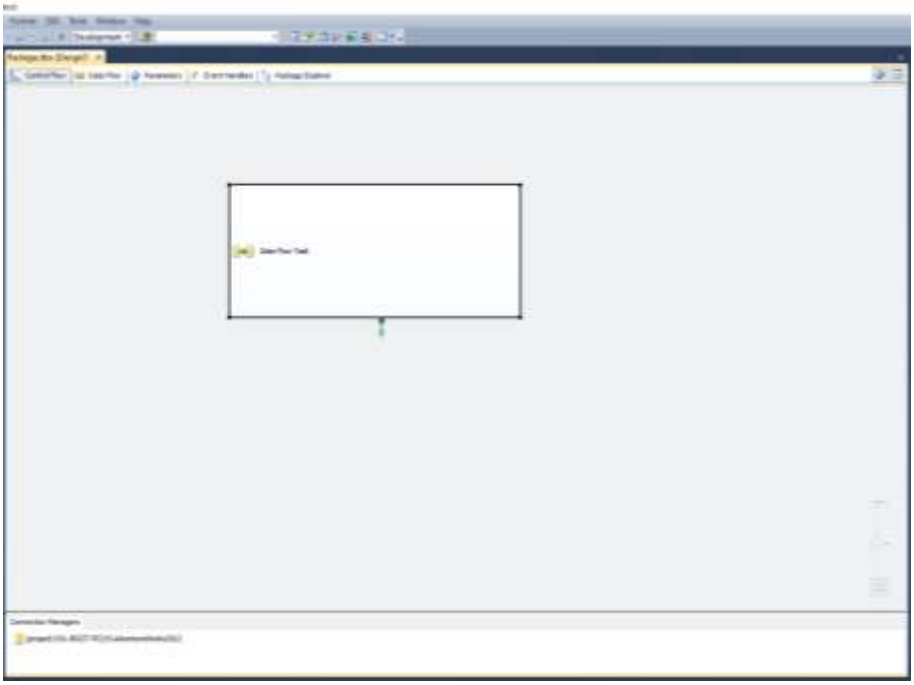
Test the connection, it should succeed.



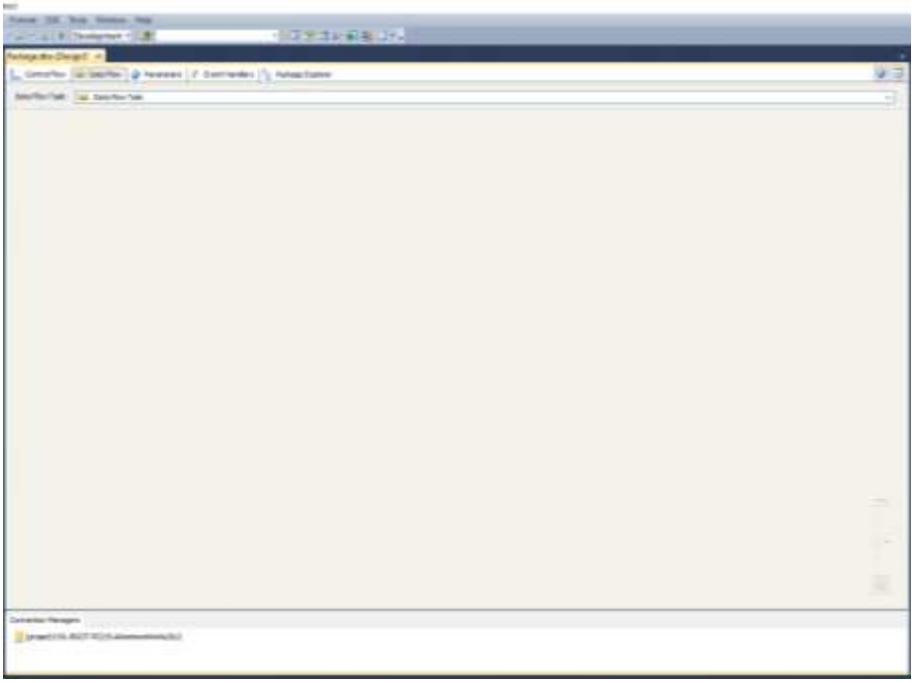
Click on OK.



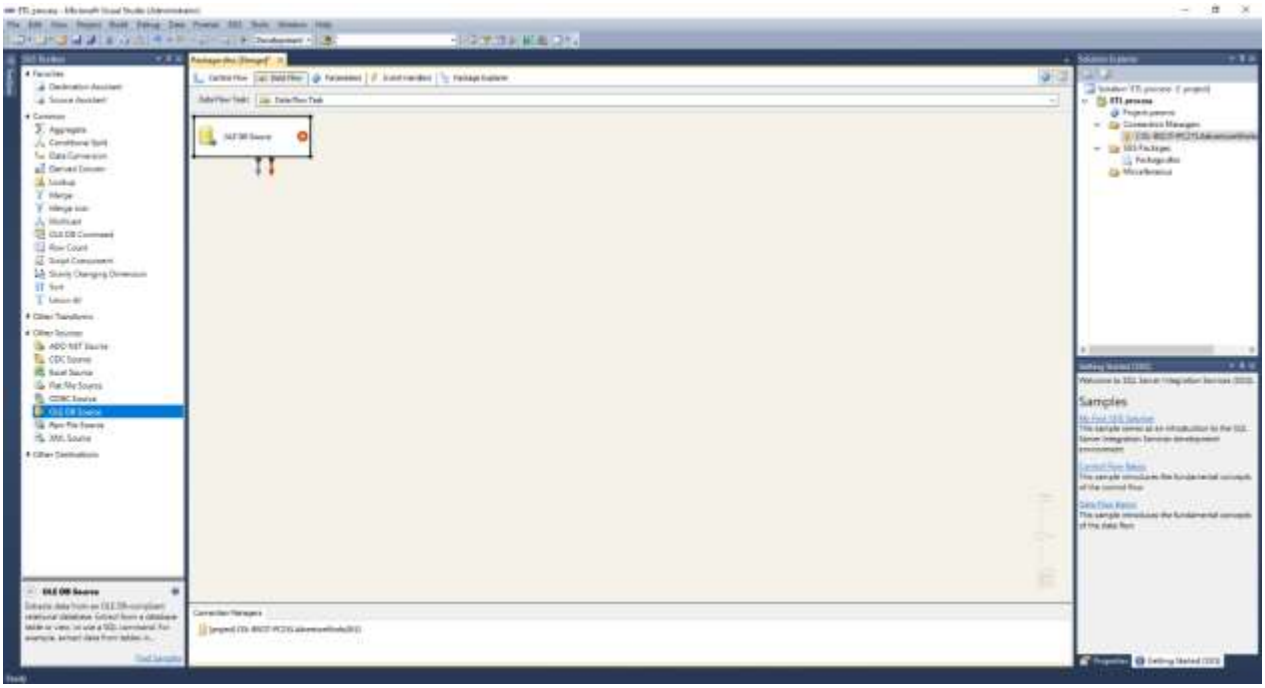
The following will be displayed on the main screen in the Control Flow view.



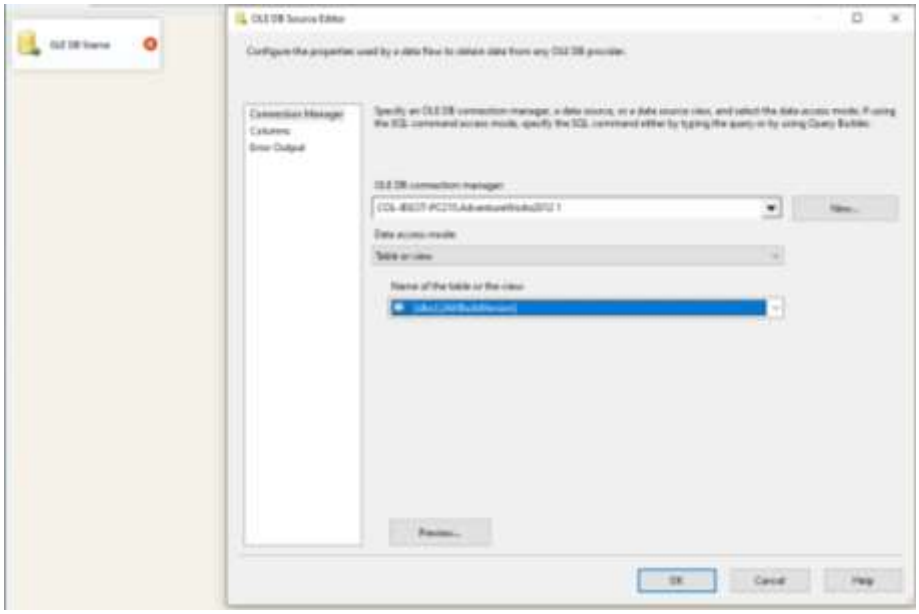
In the Data Flow task, the following will be displayed.

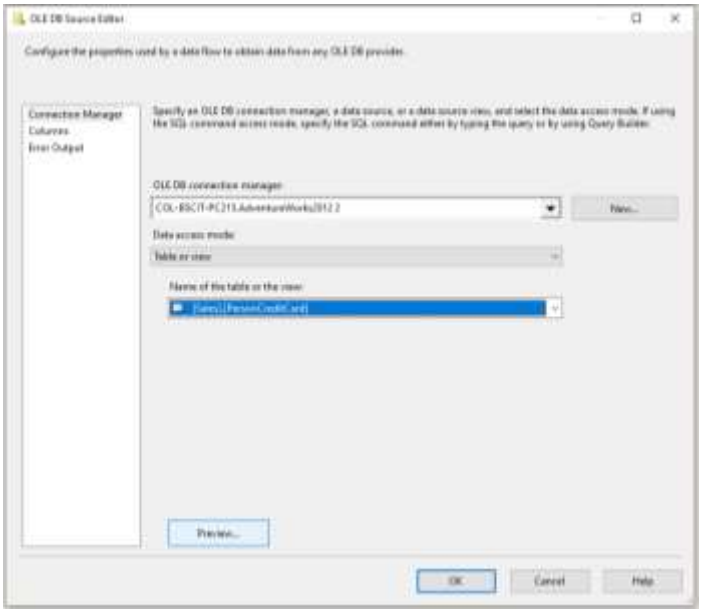


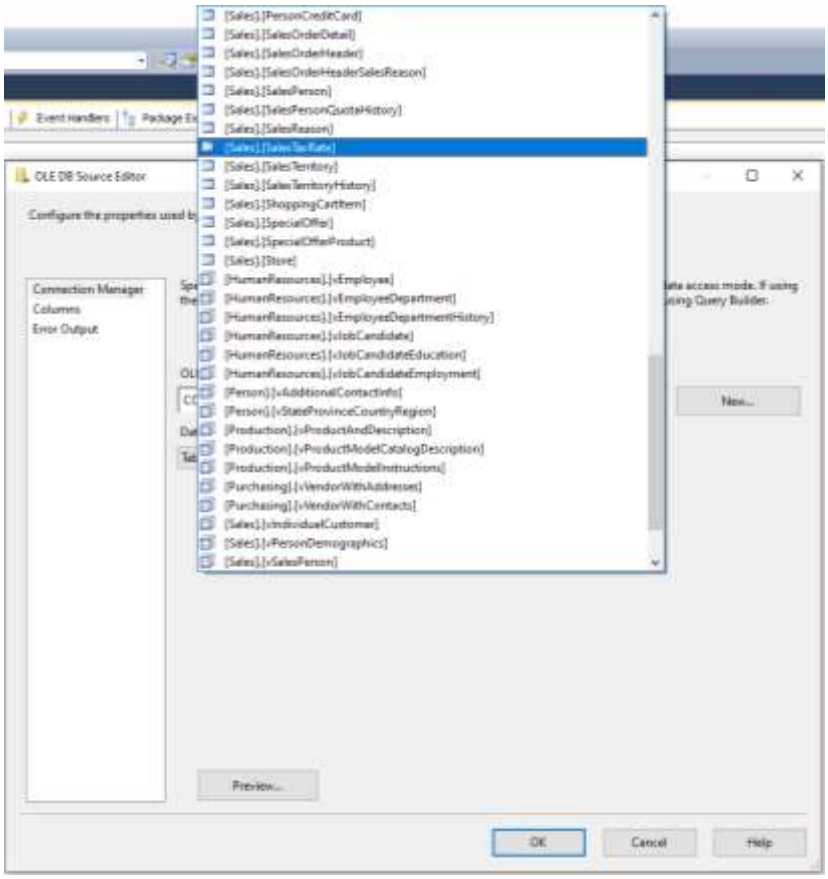
Under Other Sources in the SSIS Toolbox drag and drop OLE DB Source to the screen.



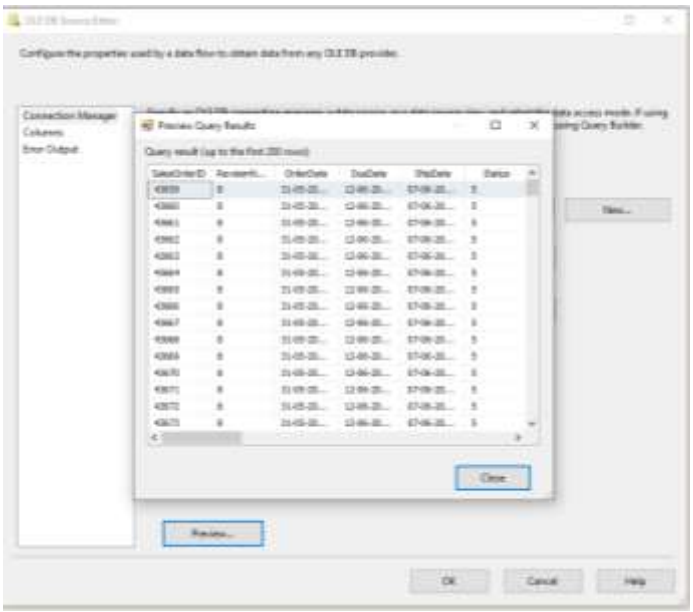
Right click on OLE DB Source. Under Connection Manager enter the OLE DB Connection Manager and the name of the tables by clicking on New.



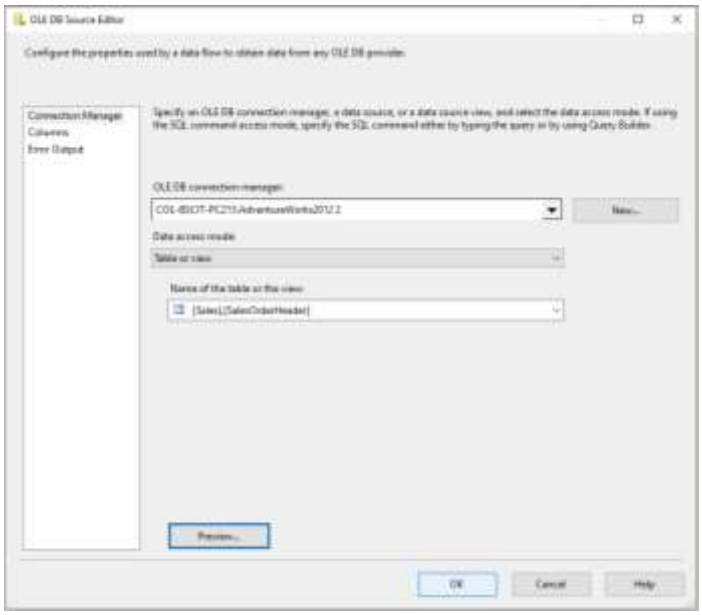




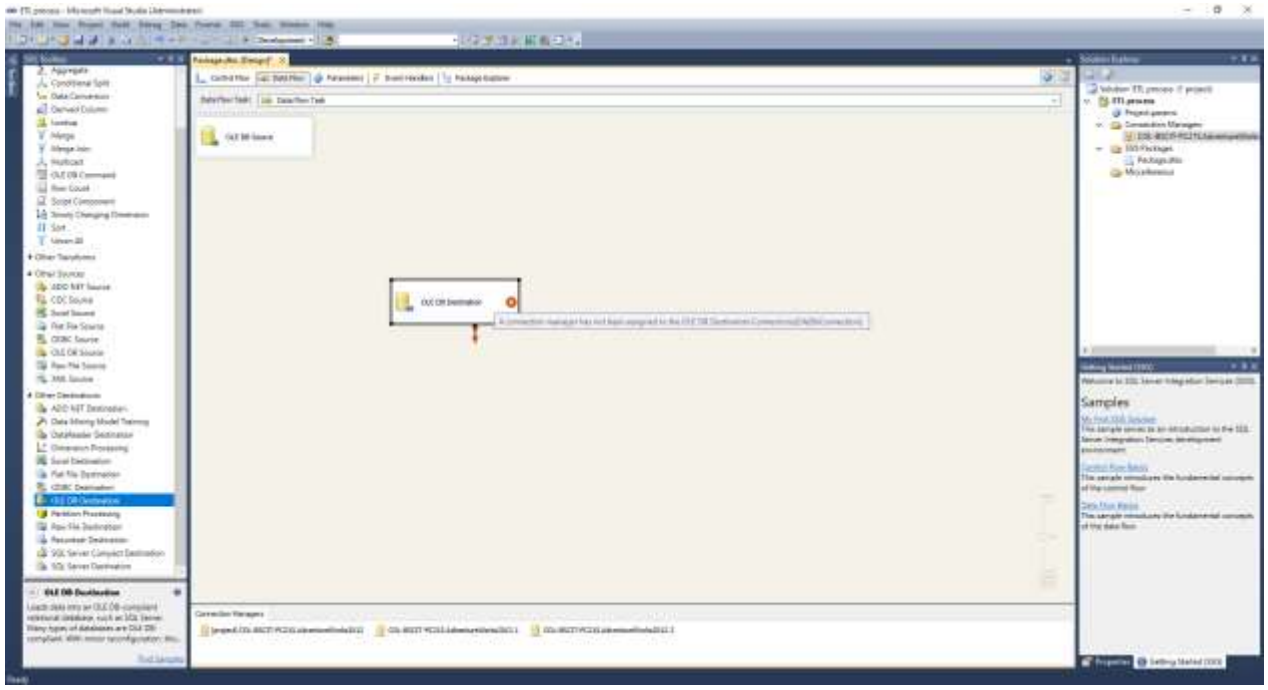
Select Preview to see the query results.

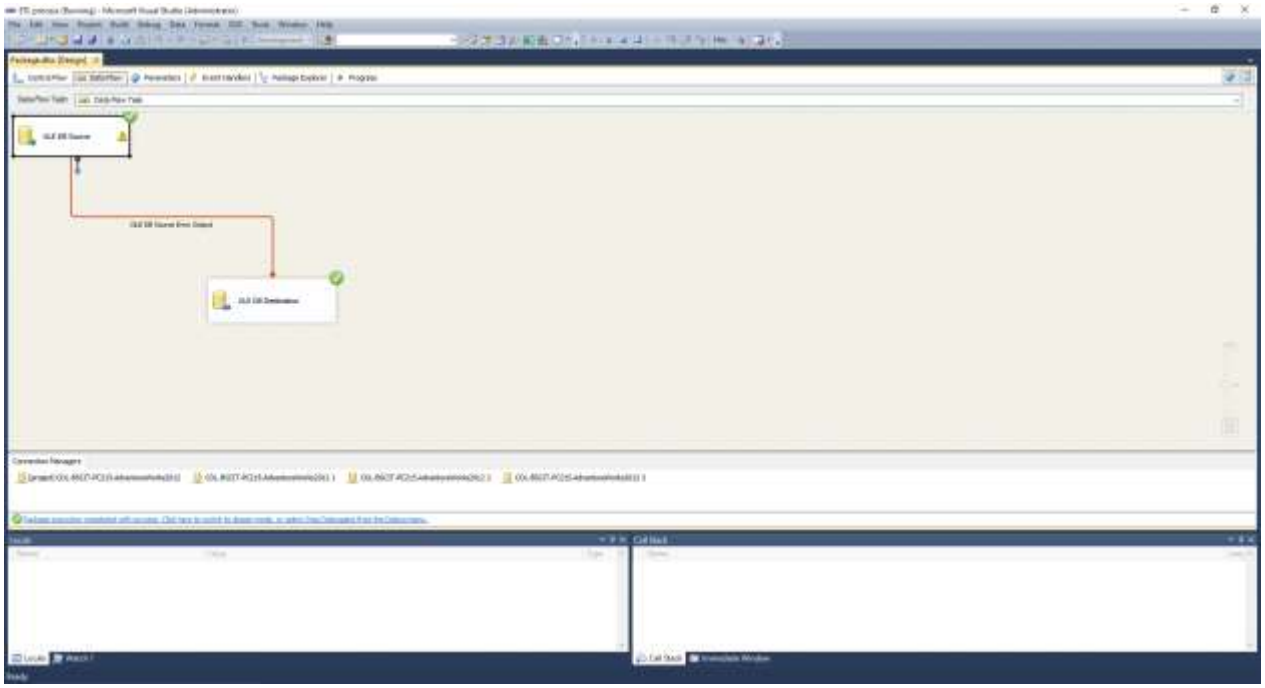


Click on OK.

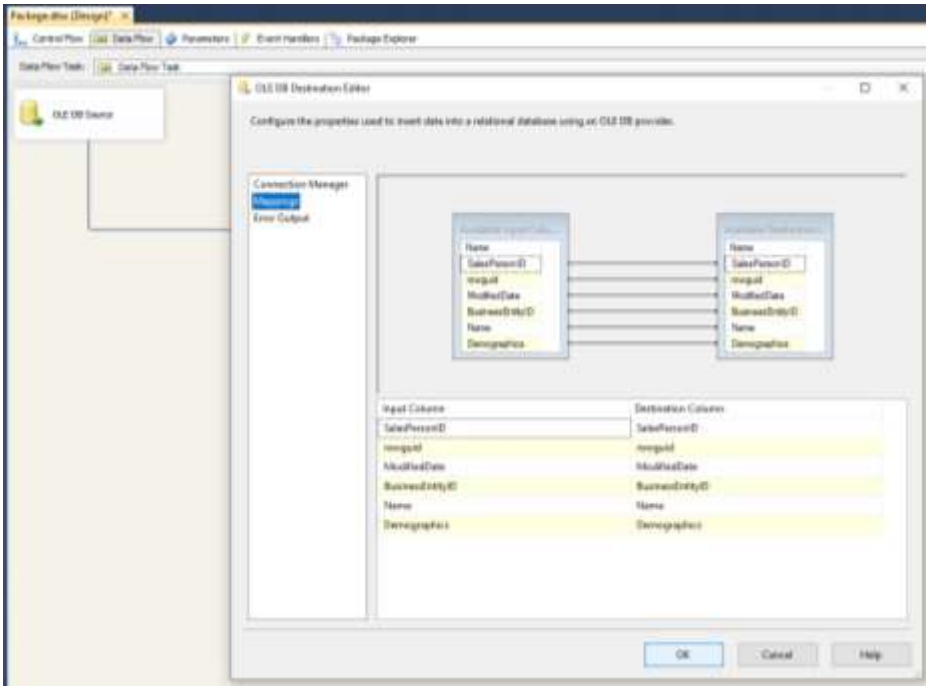


Under Other Destinations, drag and drop OLE DB Destination to the Data Flow screen.

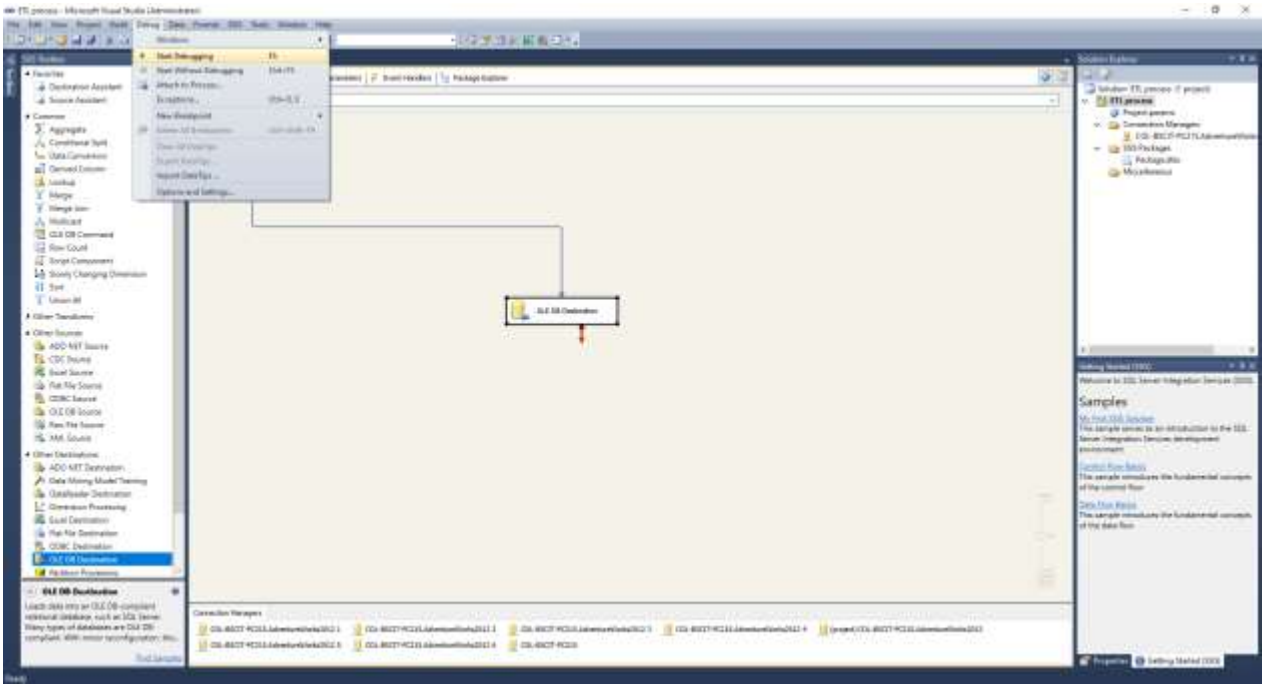




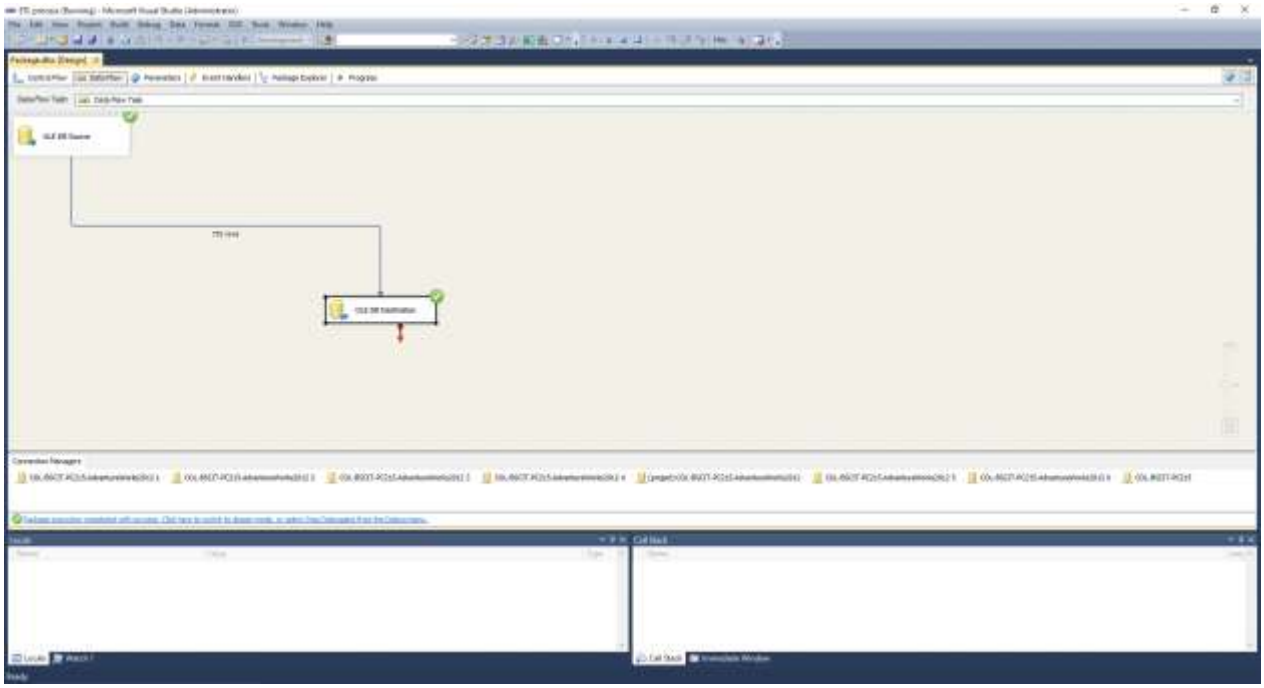
In OLE DB Destination Editor check the Mappings.



Start the debugging.

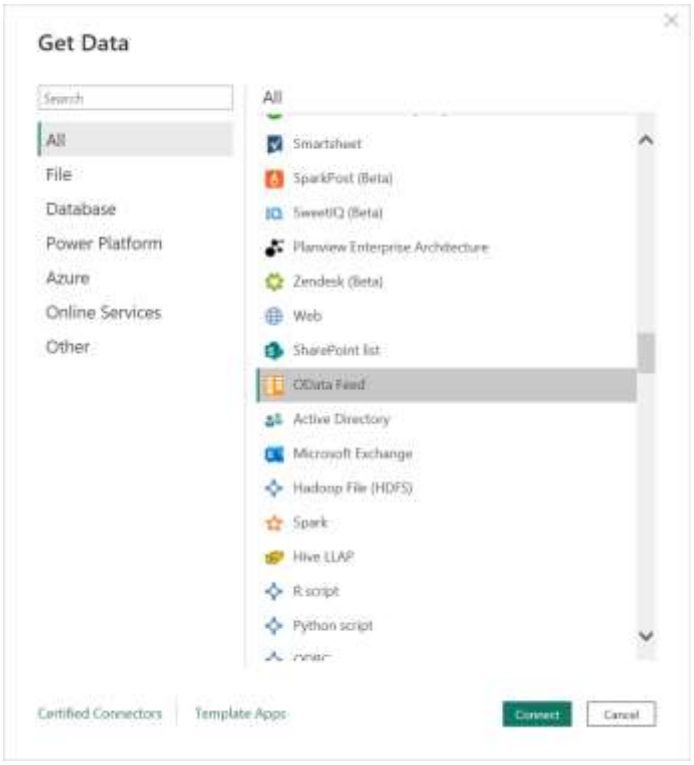


It should be successful.



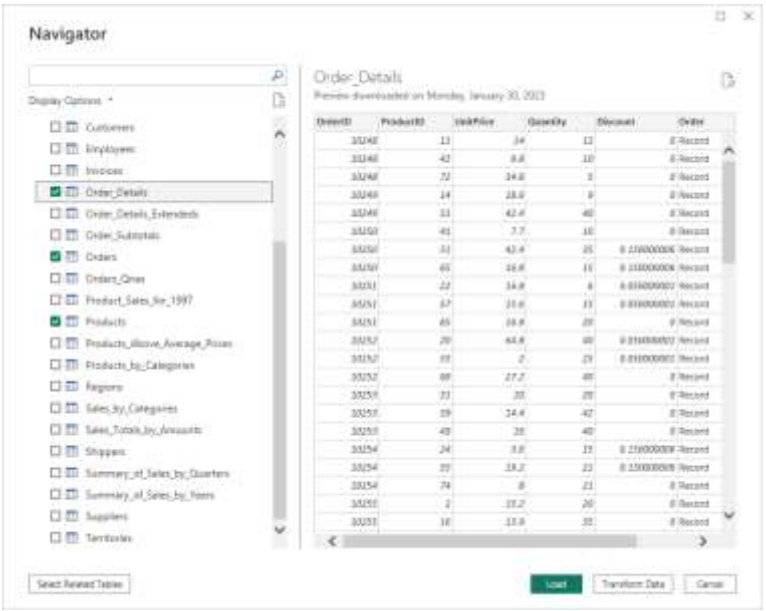
Practical 2B – Perform the ETL process in Power BI

Select OData Feed from Get Data and click on Connect.



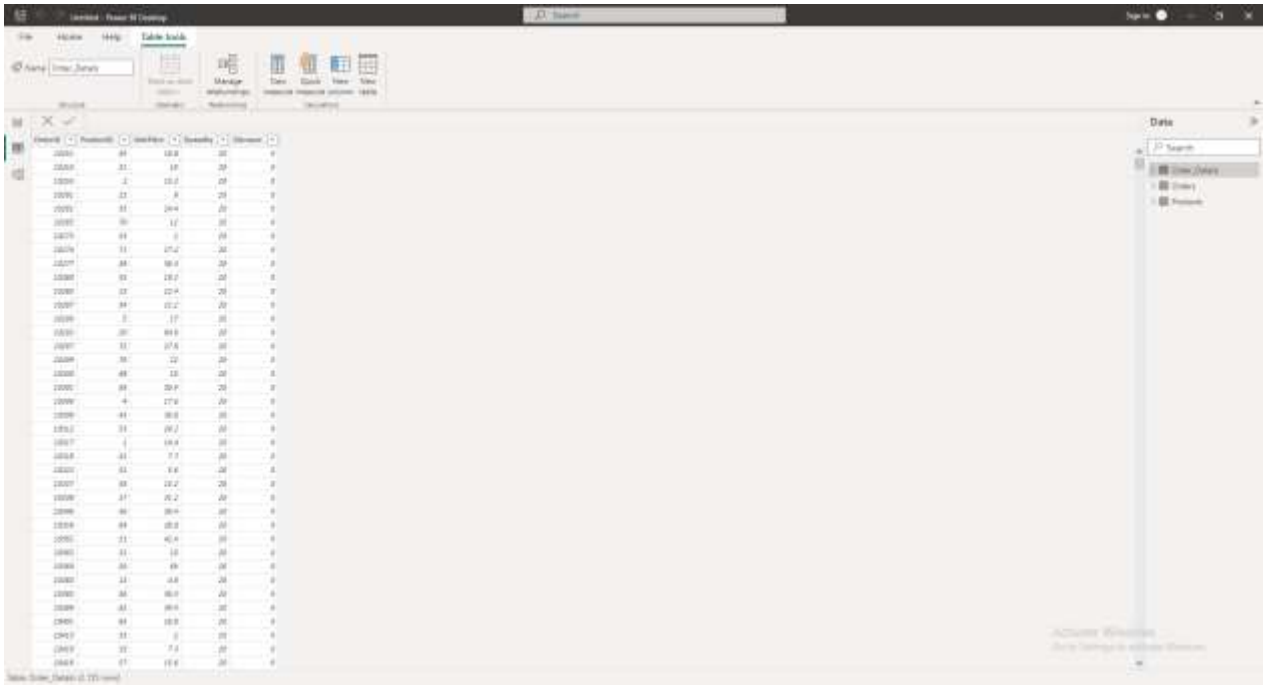
Enter the URL provided that is https://services.odata.org/Northwind/Northwind.svc and click on OK.

Select the tables – Orders, Order_Details and Products and load them.

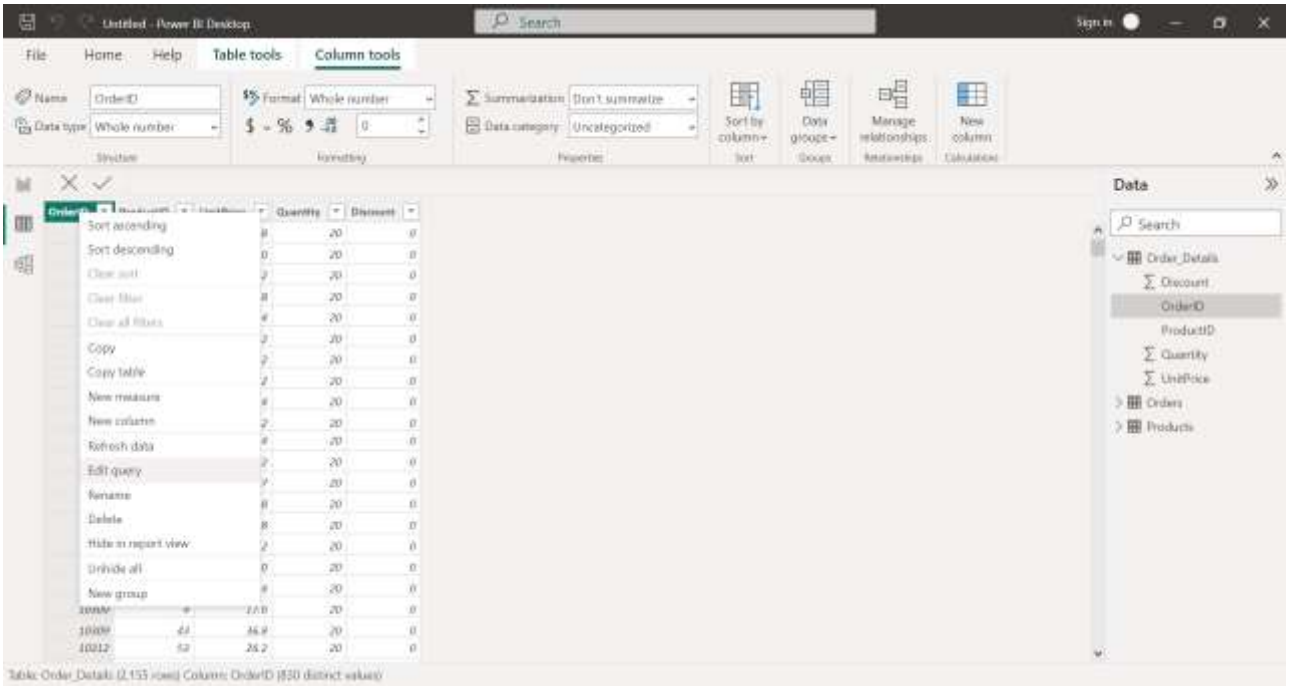


The loading process takes a bit of time.

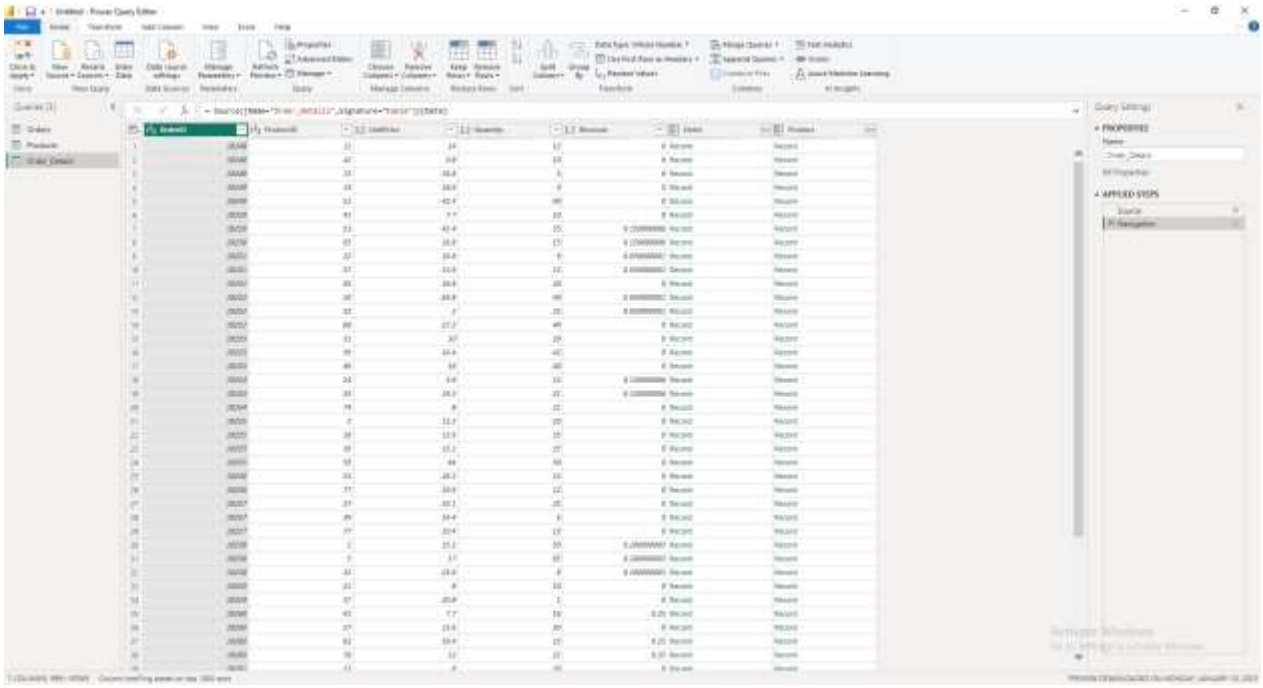
The data will be displayed of the tables selected once the loading is complete.



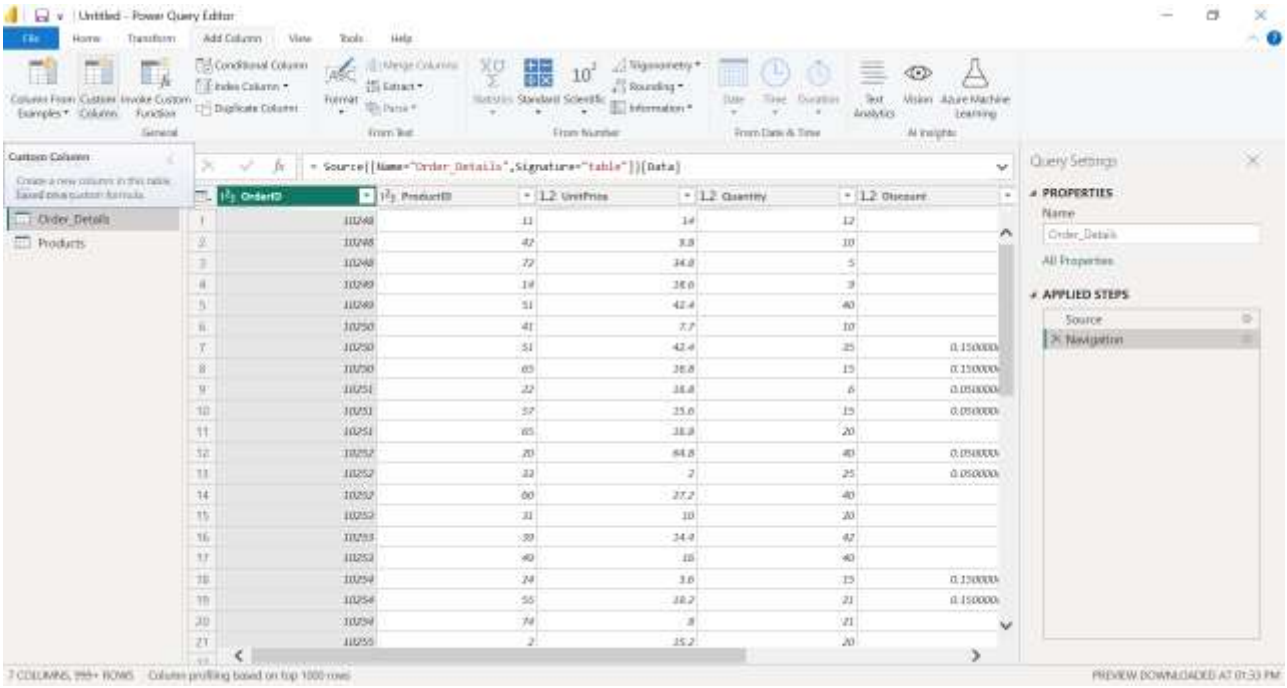
Right click on any of the column and select Edit Query.



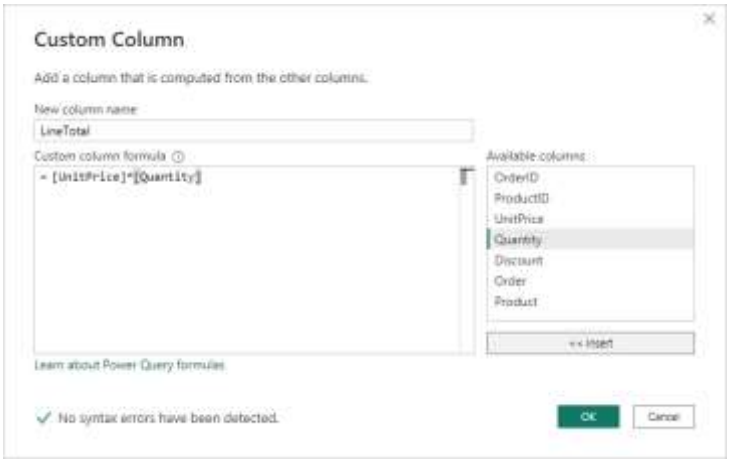
A new window will be opened.



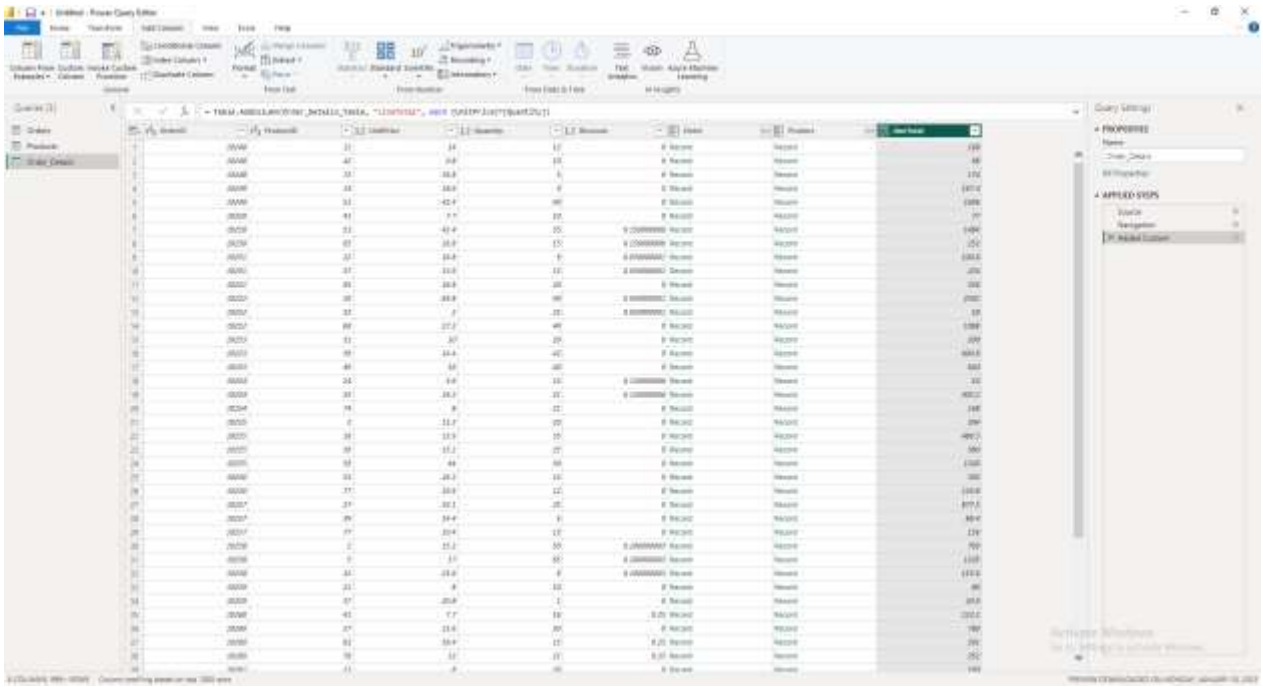
Click on Custom Column.



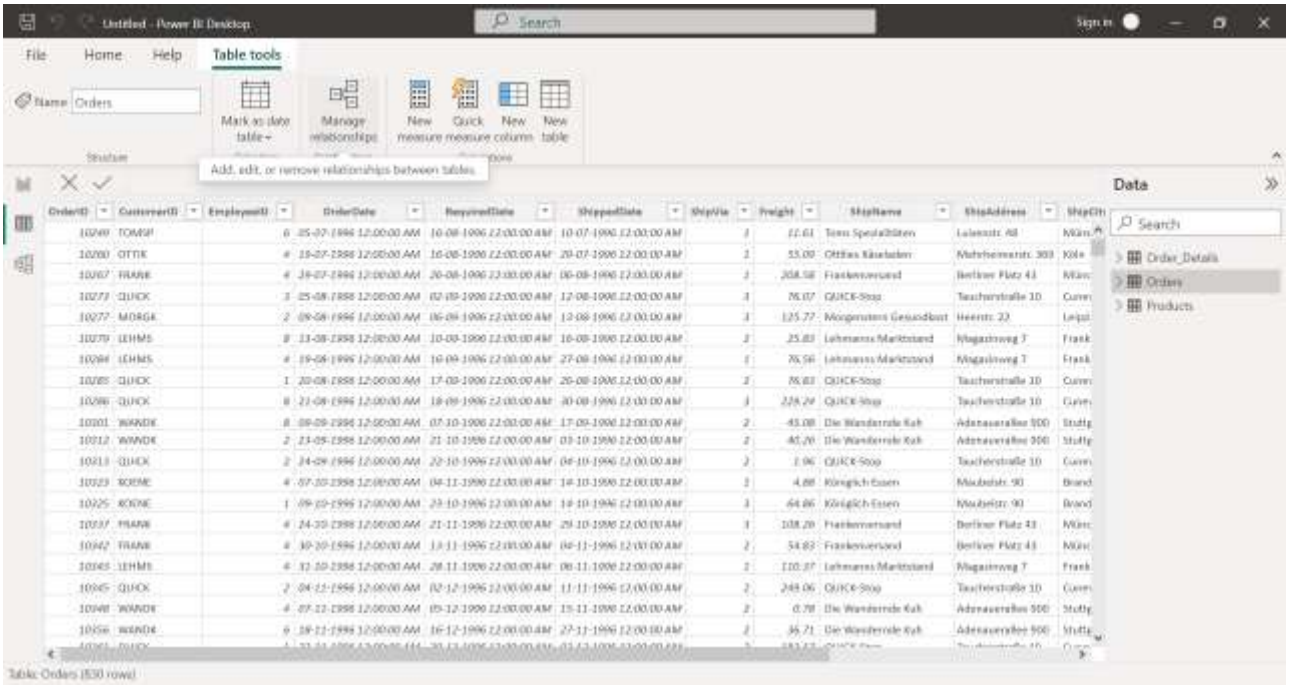
Name the column name as LineTotal and enter the formula by inserting the columns as the variables and then click on OK.



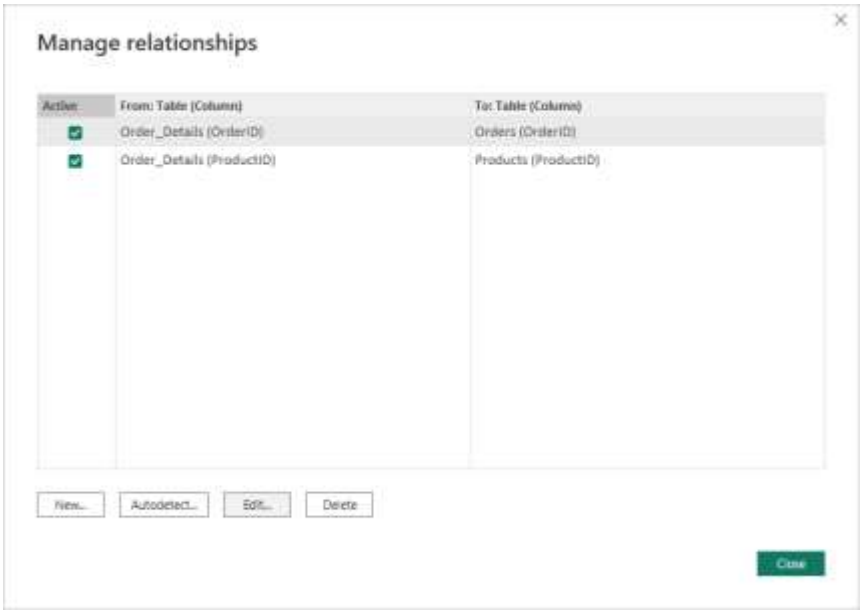
The LineTotal column will be added.



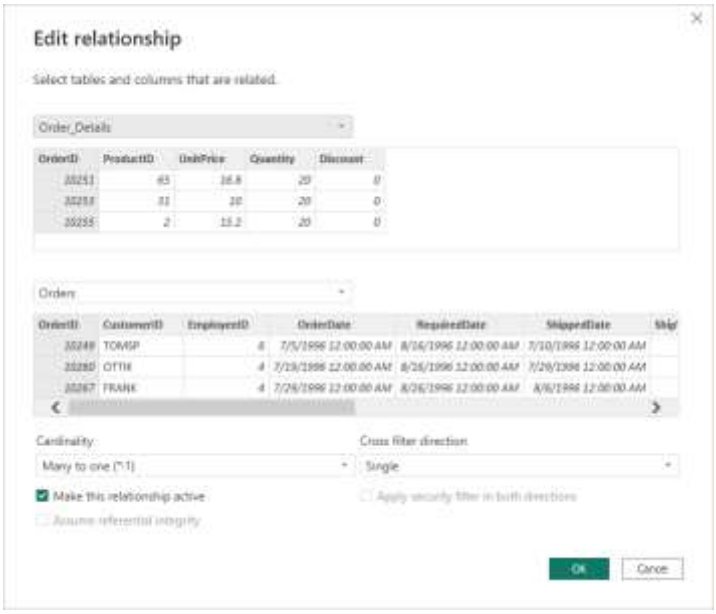
Select the Orders table and under Table tools select Manage Relationships.



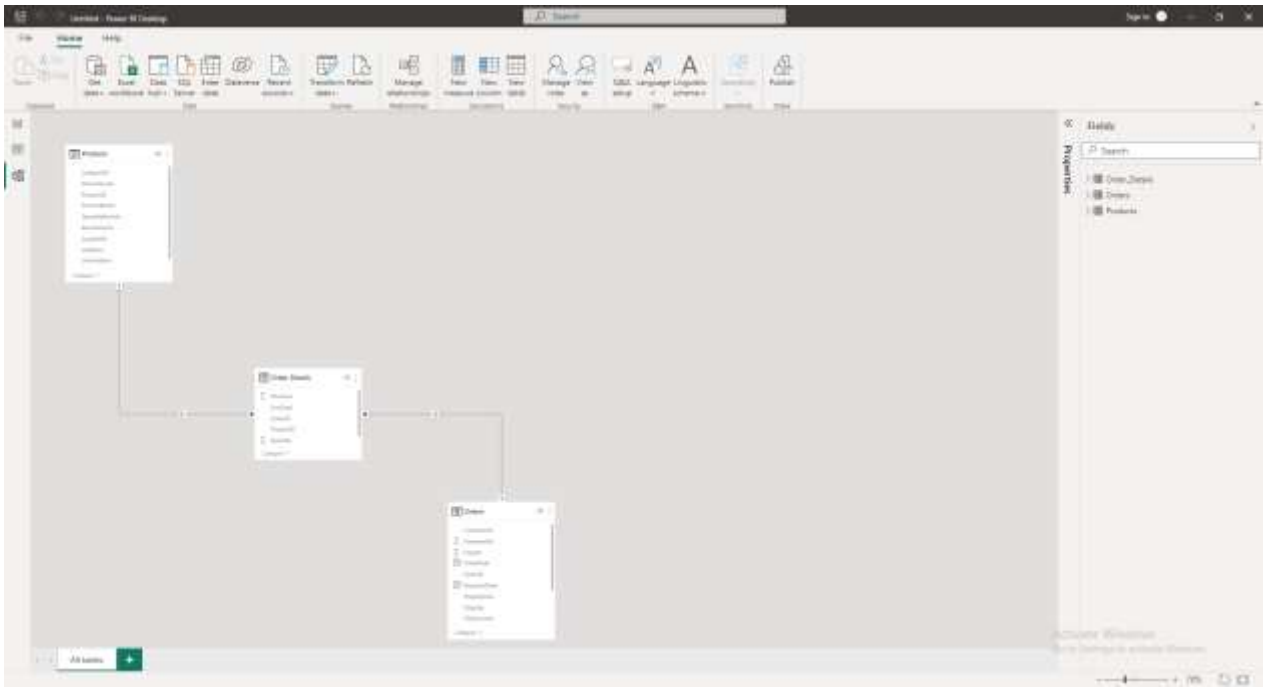
The columns will come pre-selected.



Edit the details according to the requirement.



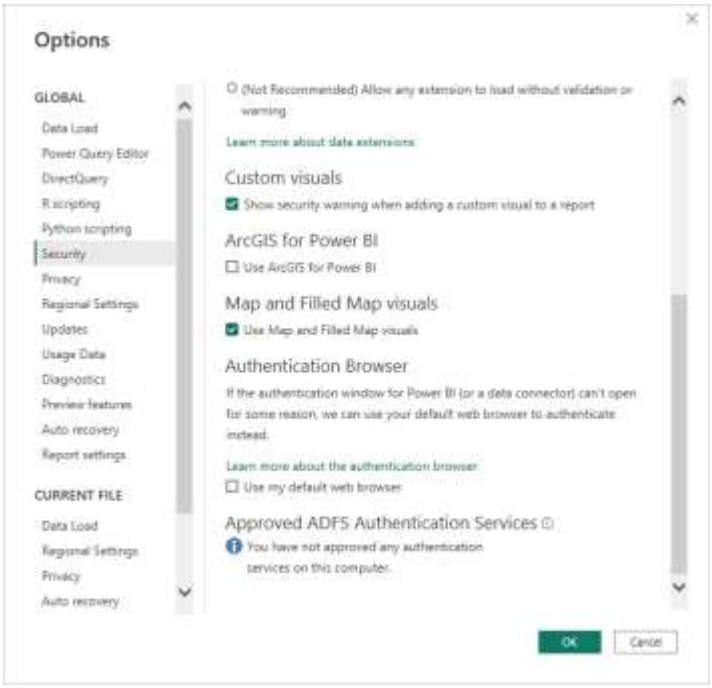
Click on the relationships icon.



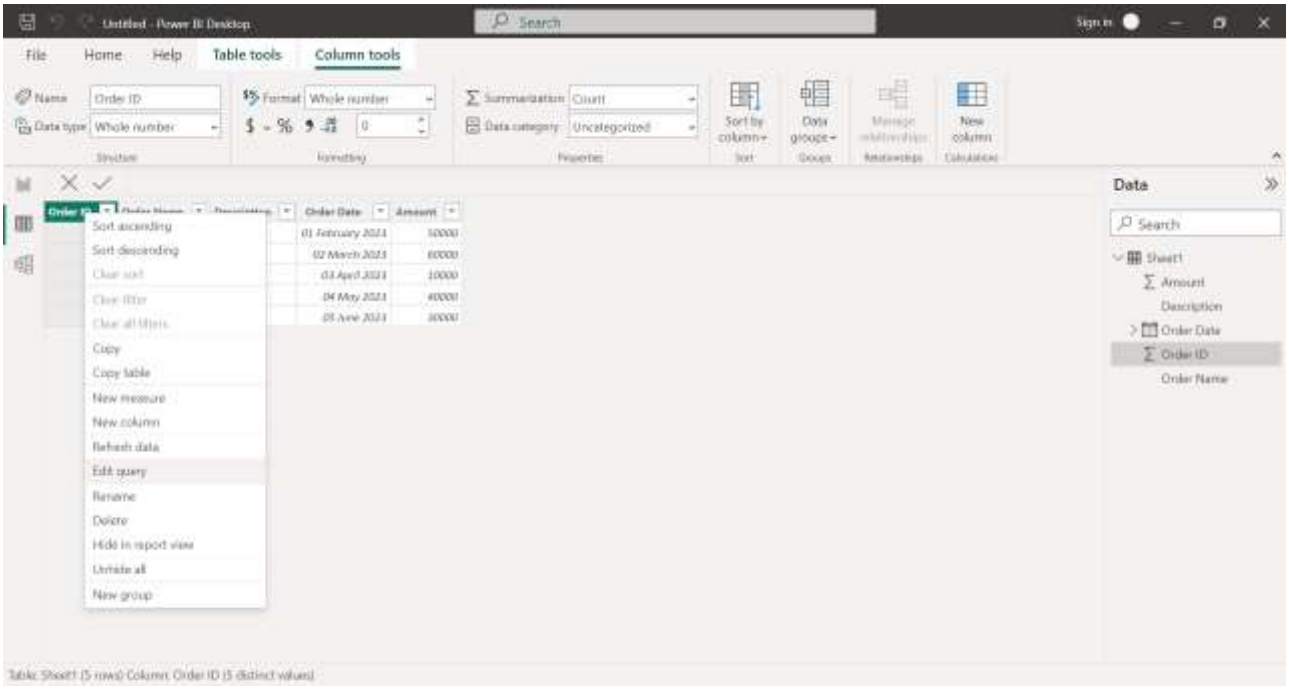
KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

Practical 3A – Perform data visualization in Power BI and create the data staging area for the selected database

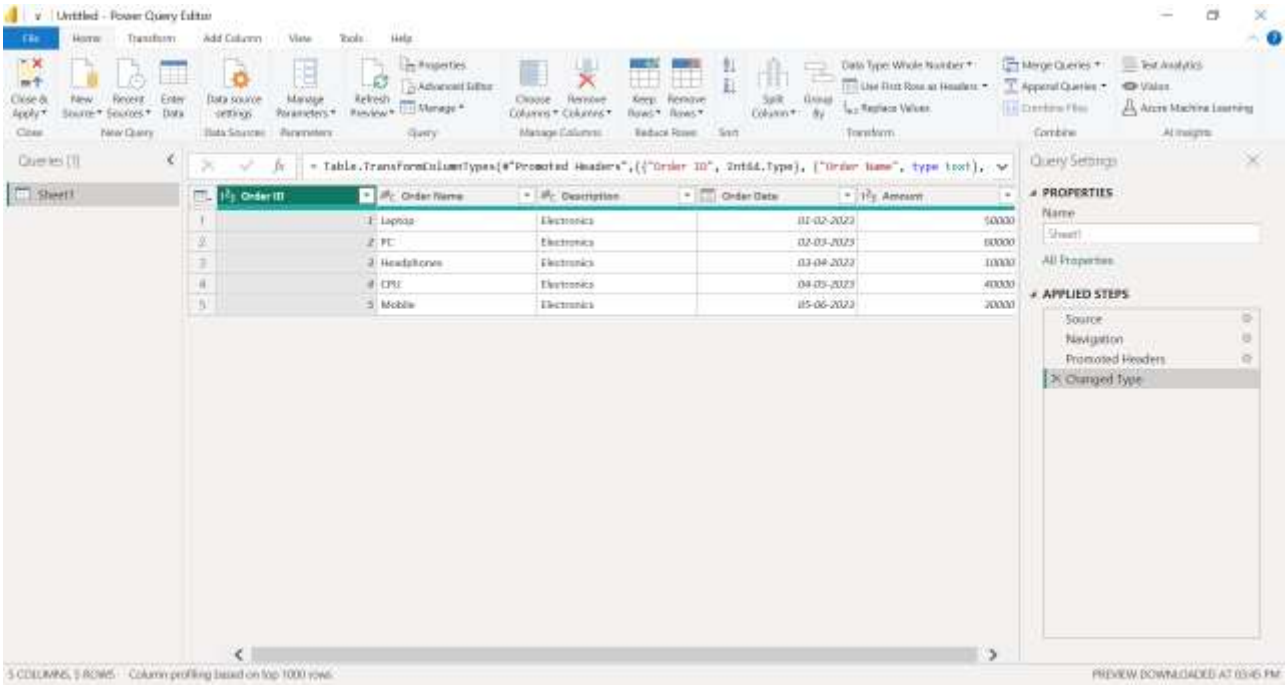
In Power BI click on File then go to Options and Settings click on Options and check Map and Filled Map Visuals. Click on OK.



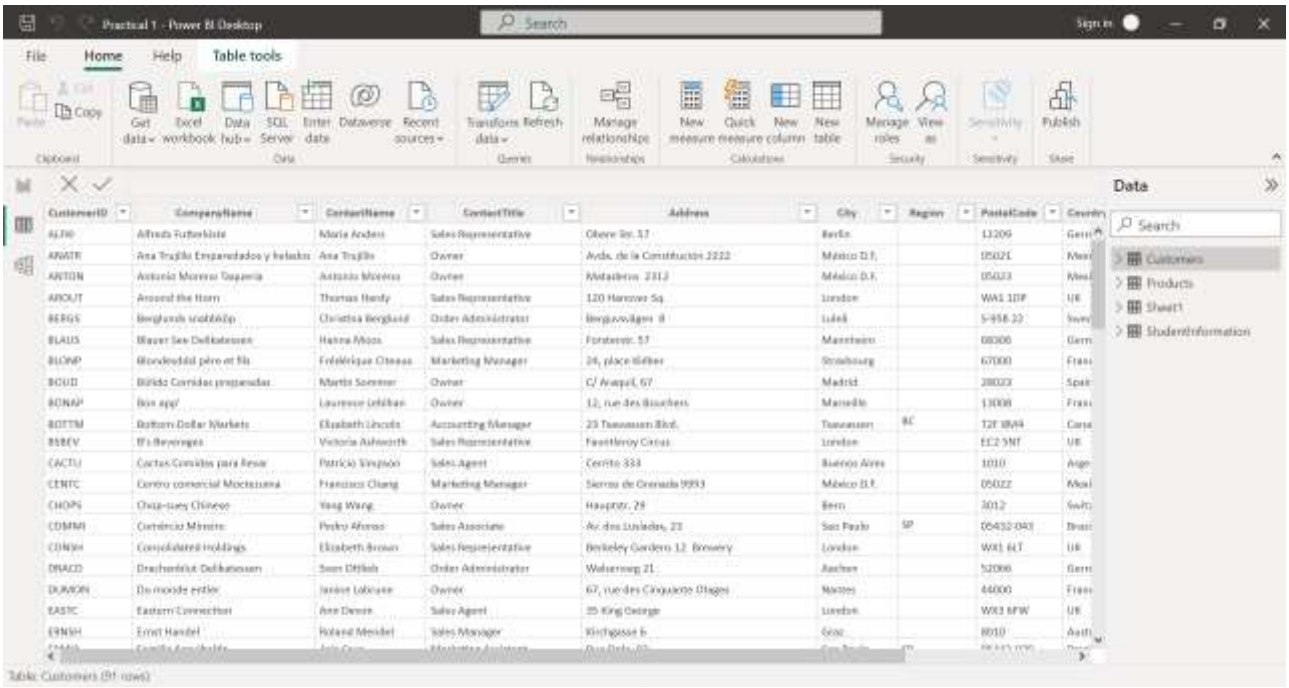
Open a table. Right click on any column and select Edit query.



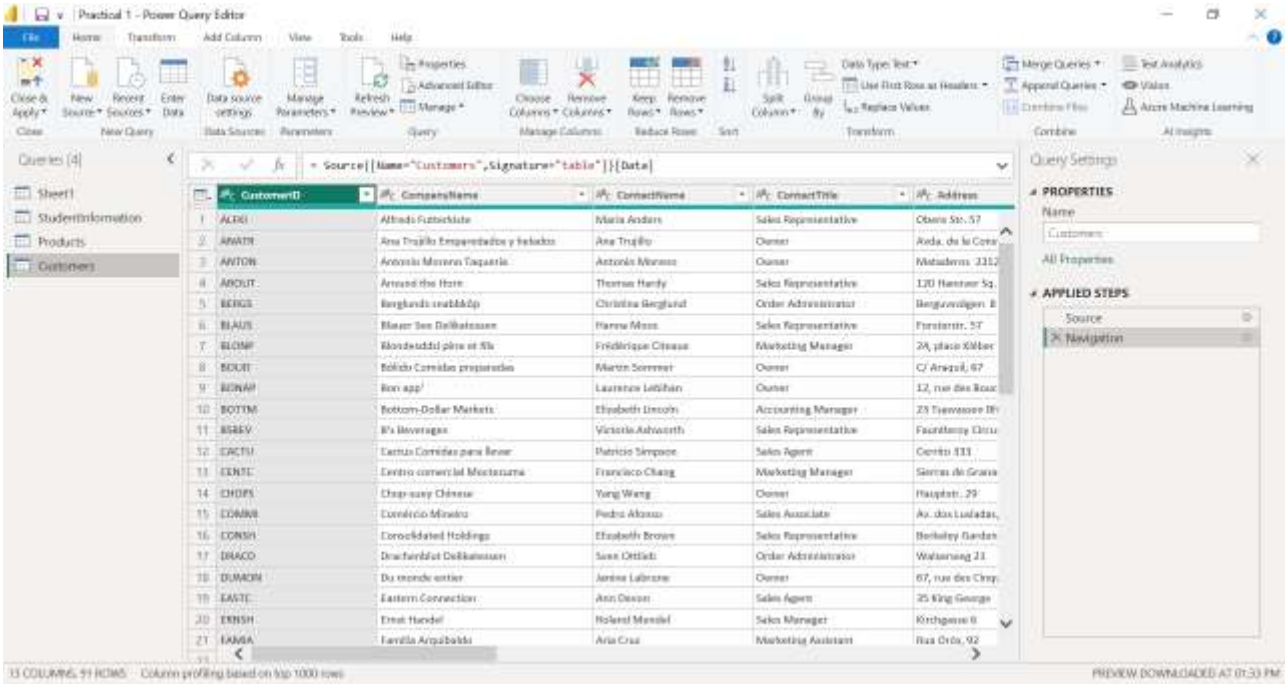
Another window will be opened which look like below.



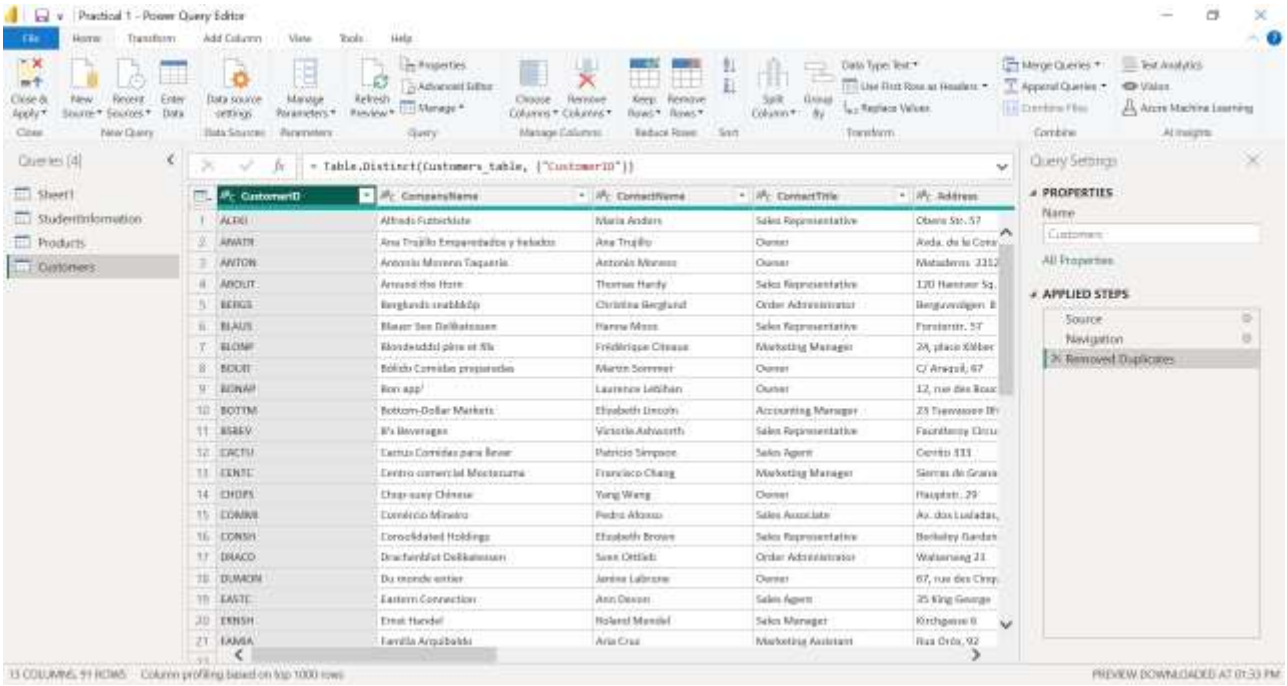
The tables will be now loaded. Select the Customers table.



Select any column and right click to select Edit Query. New window will open.



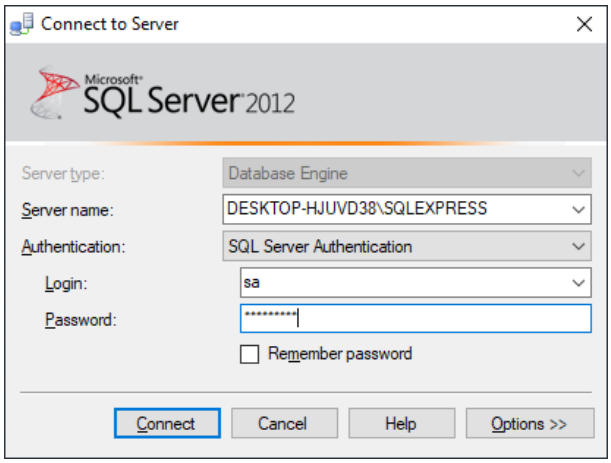
Remove duplicate entries by right clicking on the column and selecting Remove Duplicates.



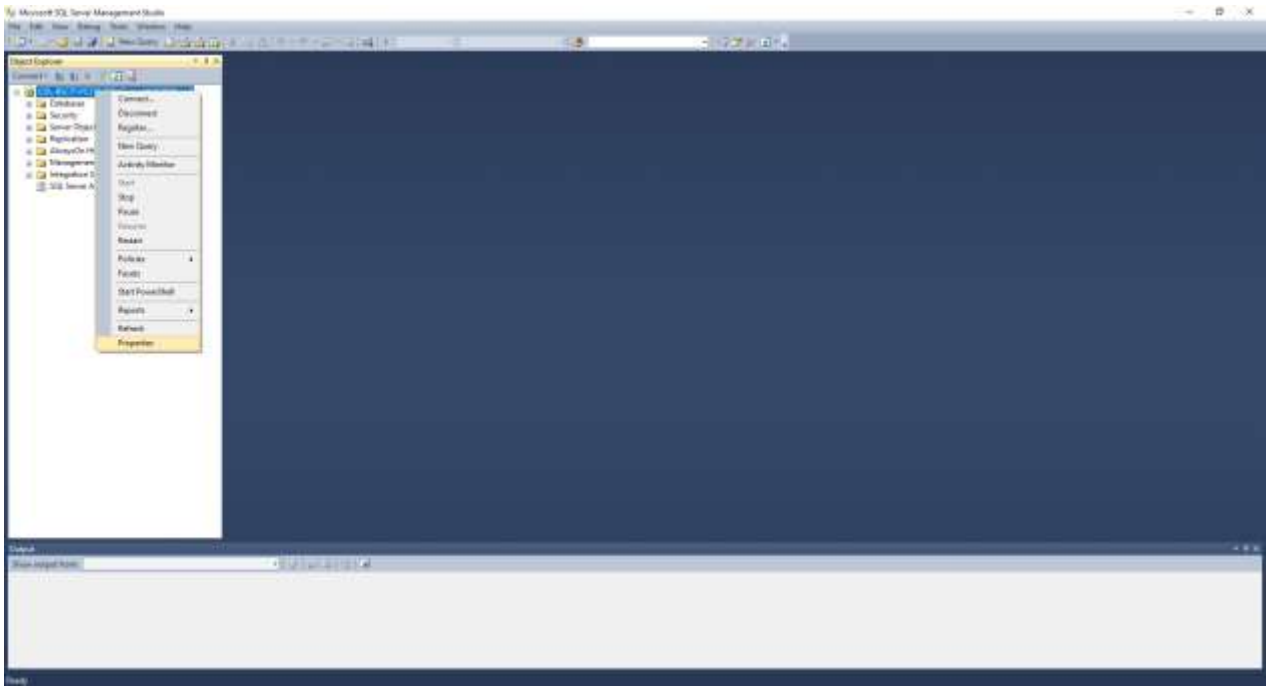
KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

Practical 3B – Create the data staging area for the selected database using Star Schema

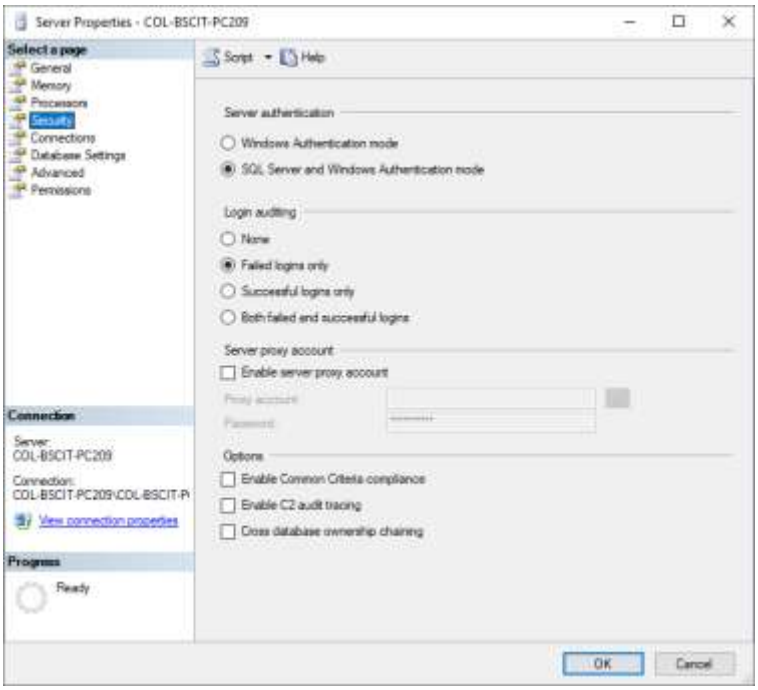
Open SQL Server Management Studio. Authentication will be SQL Server Authentication. Login username is sa and password is bscit@123.



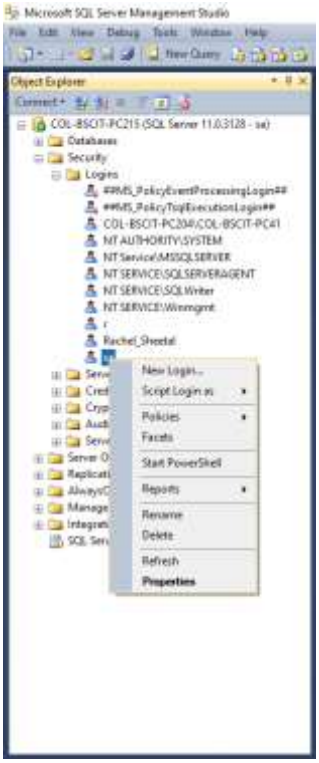
In the Object Explorer, right click on the server and click on Properties.



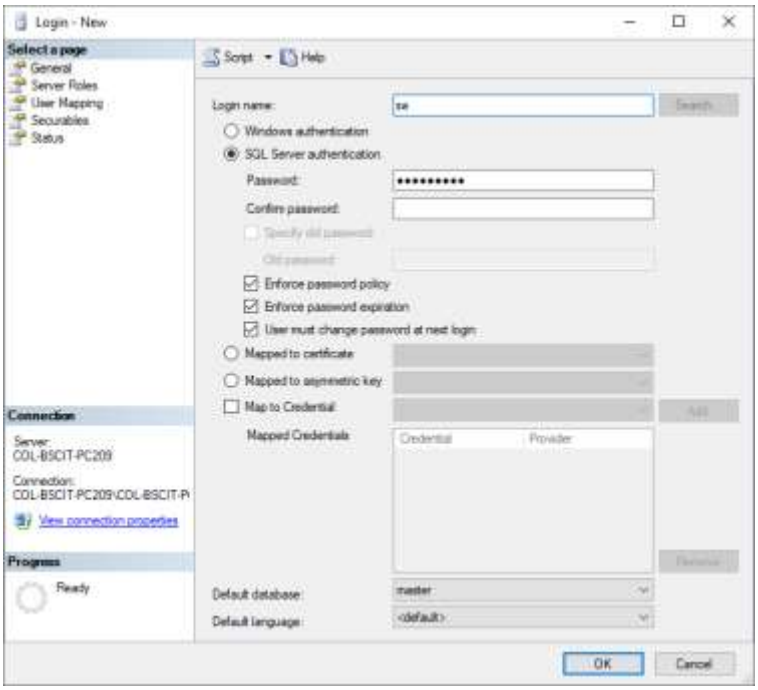
Open Security under Server Properties and check if everything is proper.



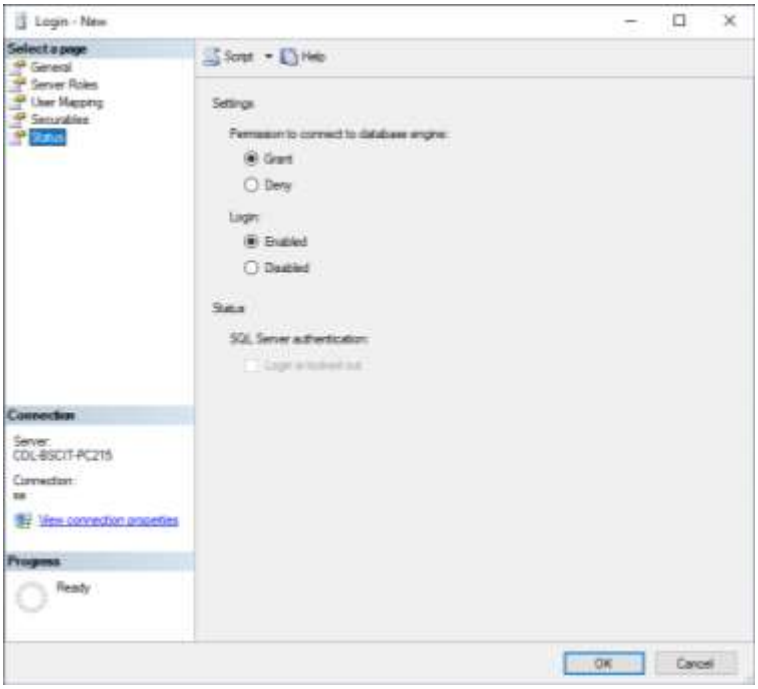
Under Security and then under Logins there is a folder named sa that is the login username, right click on it and select New Login.



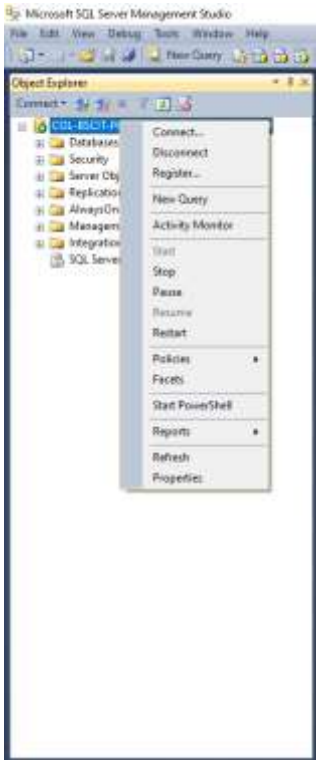
Check the enforcement of passwords.



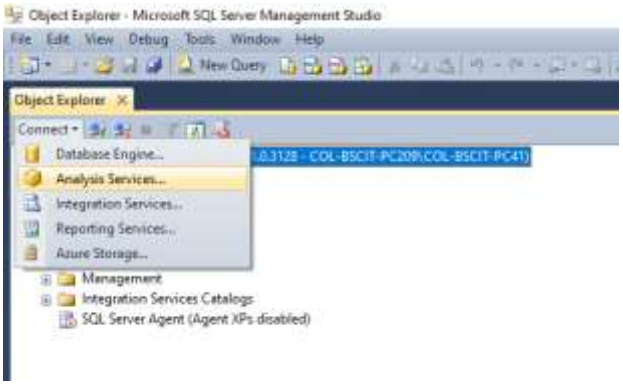
Enable the login in Status.



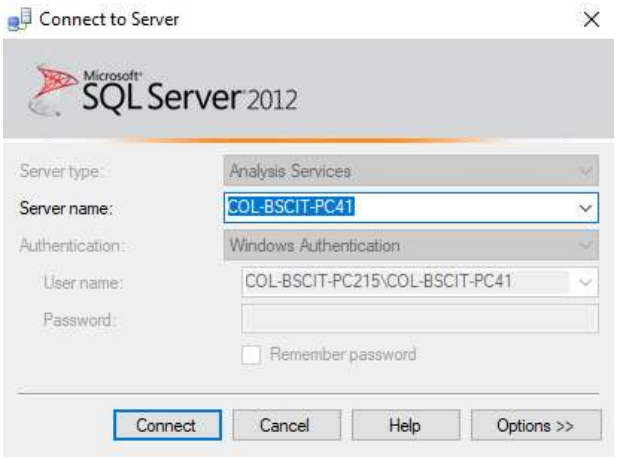
Right click on the server name and click on Restart.



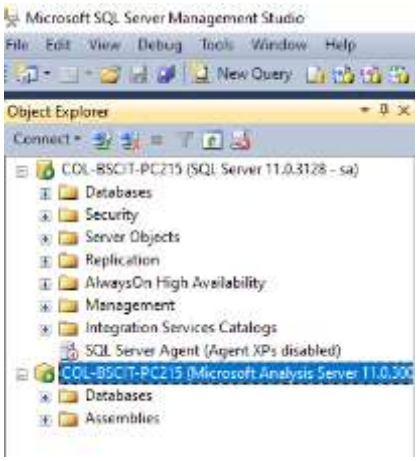
Click on Connect and select Analysis Services.



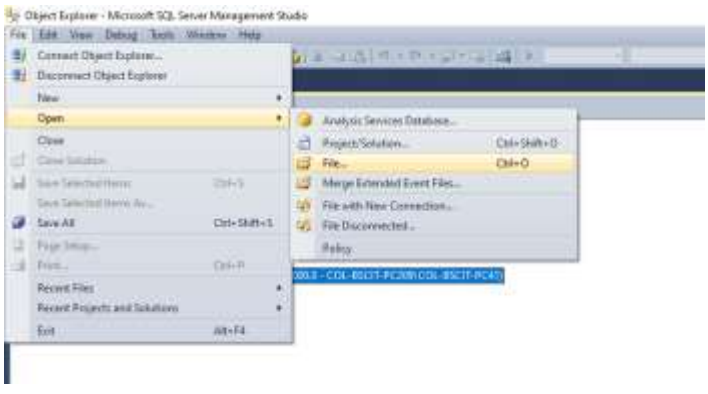
Enter the credentials.



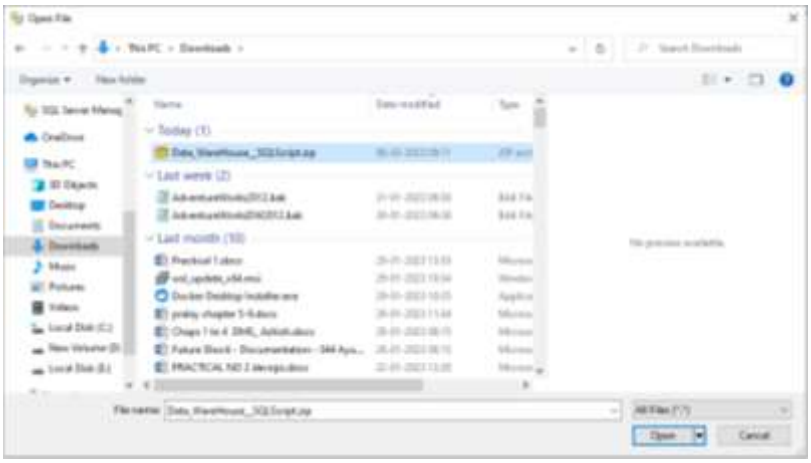
The Analysis Server will now be added in the Object Explorer.



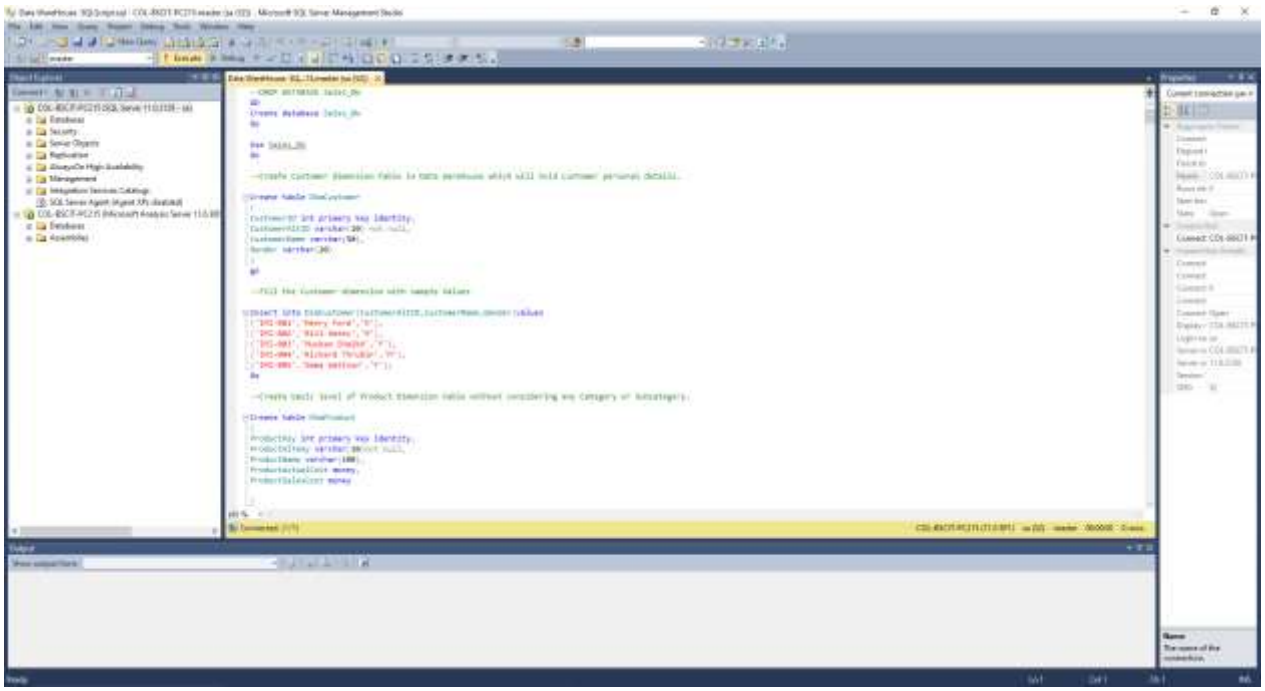
In Microsoft SQL Server Management, click on File then Open and then File.



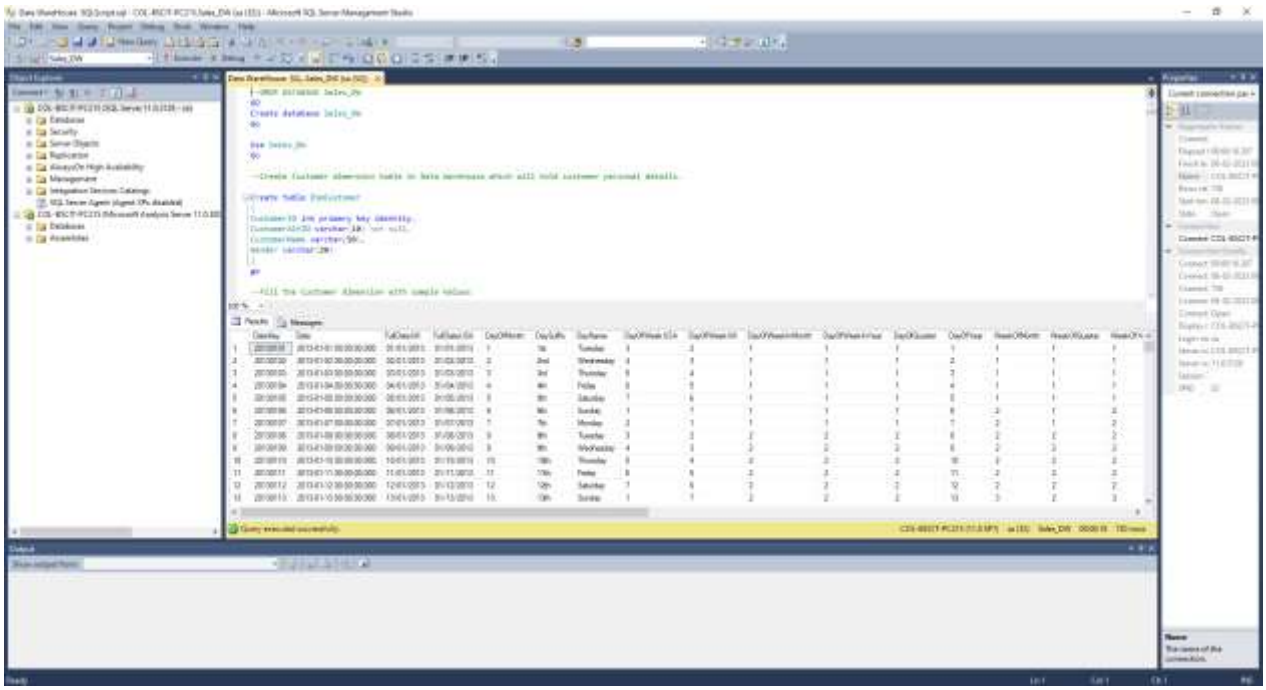
Search for the file Data_WareHouse_SQLScript.zip.



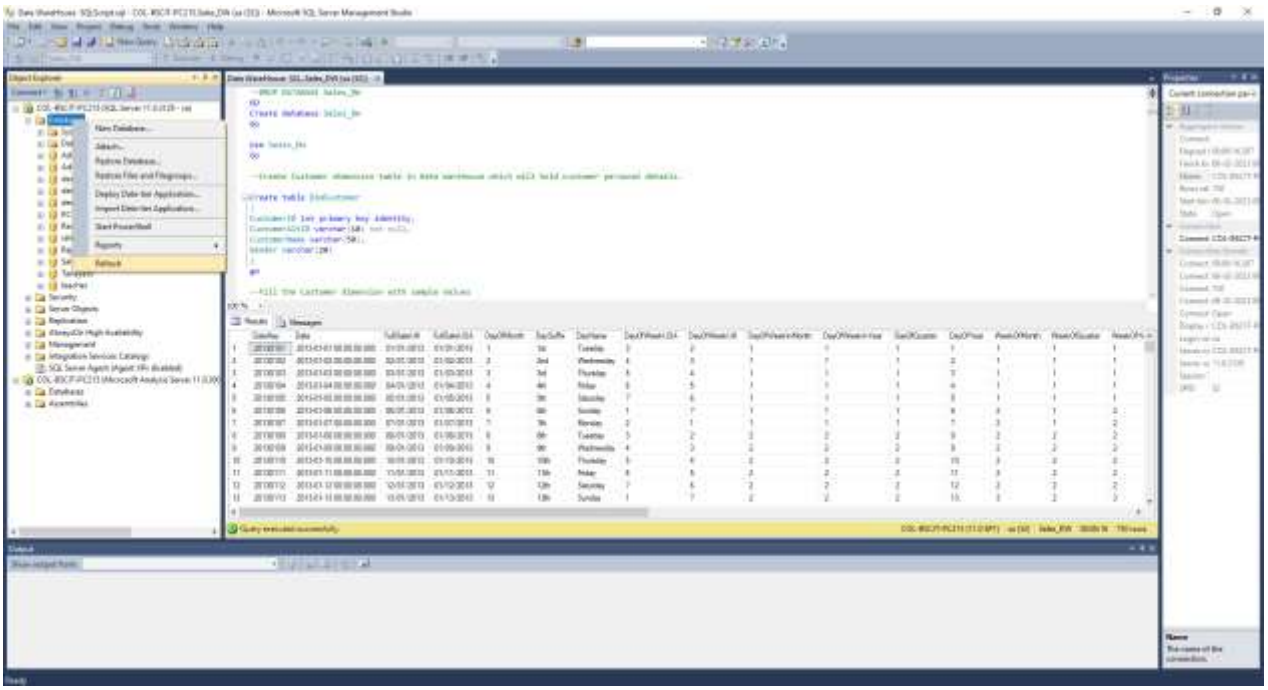
Click on Execute after importing the file. The following screen is displayed.



Query executed successfully message will be displayed at the bottom of the screen.

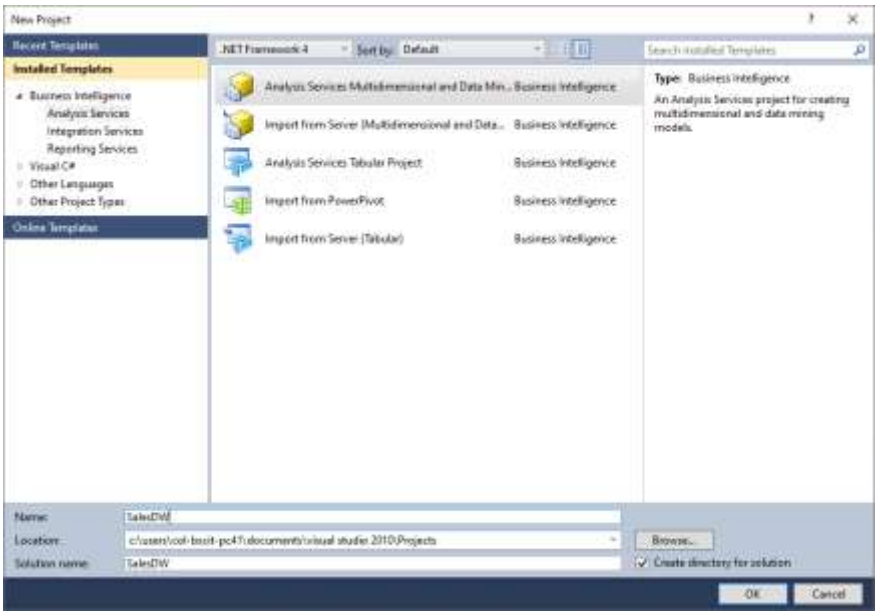


Right click on Databases and click on Refresh.

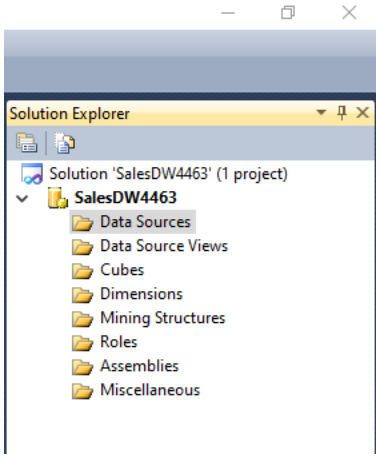


Open Microsoft Server Data Tools, click on File then New and Project.

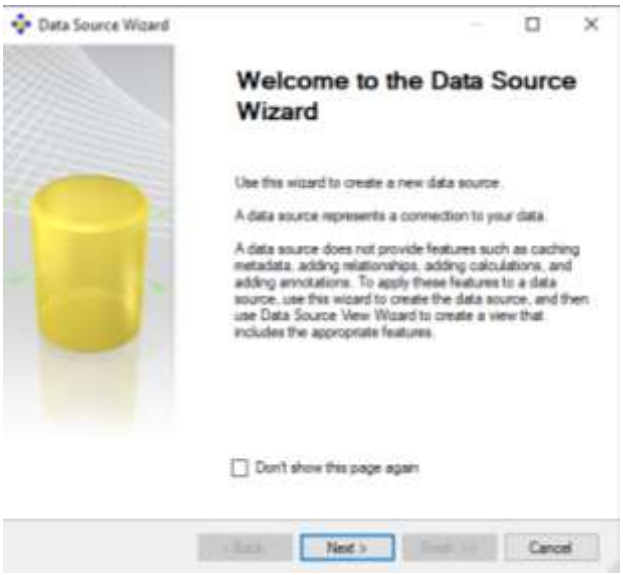
Under Business Intelligence and Analysis Services, select Multidimensional and Data Mining, name the file as SalesDW.



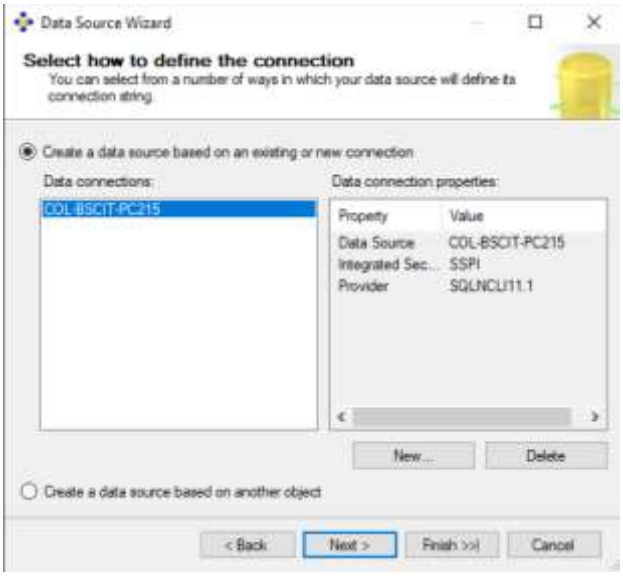
In the Solution Explorer, right click on Data Sources and select Data Source View Wizard.



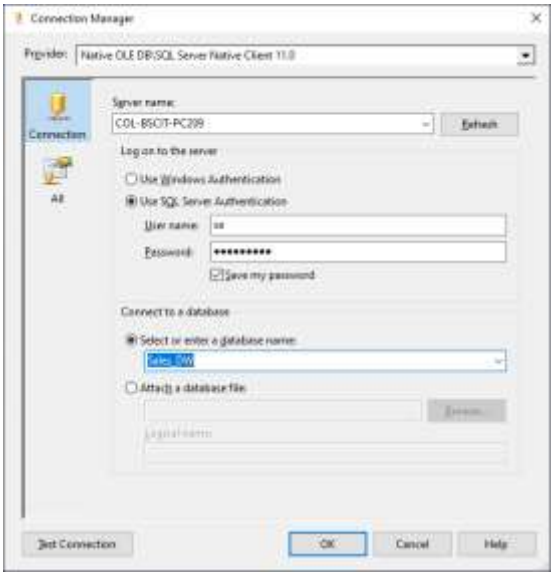
Click on Next.



Delete the current if any and add a new one and click on Next.

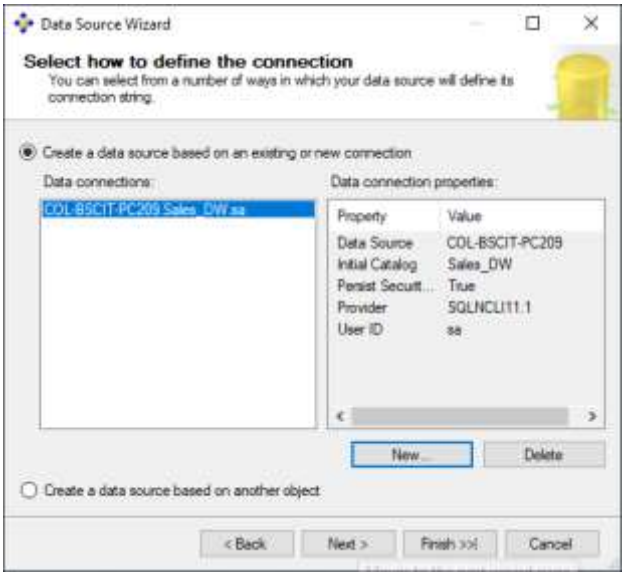


Enter the username and password and enter the database that is SalesDW.

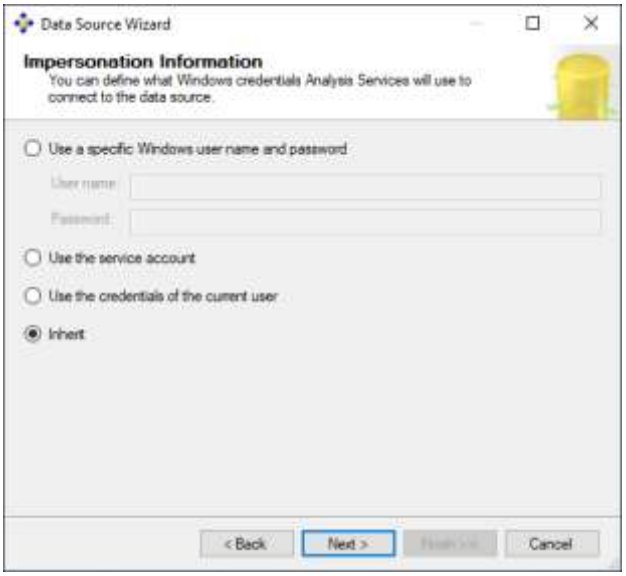


Test the connection it should succeed.

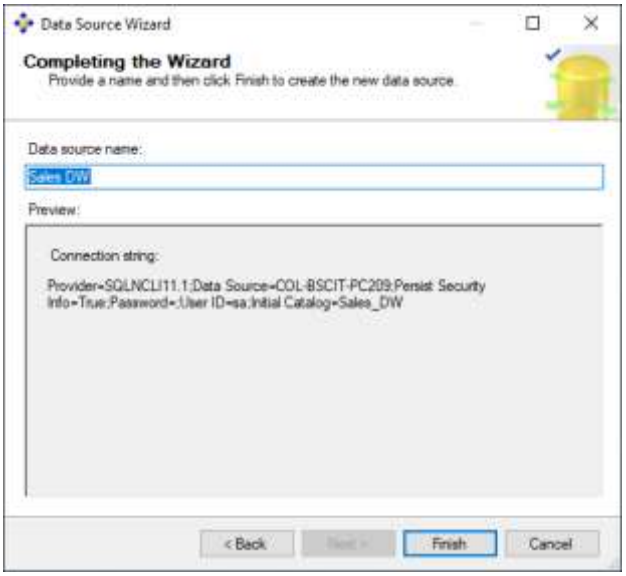
Click on Next.



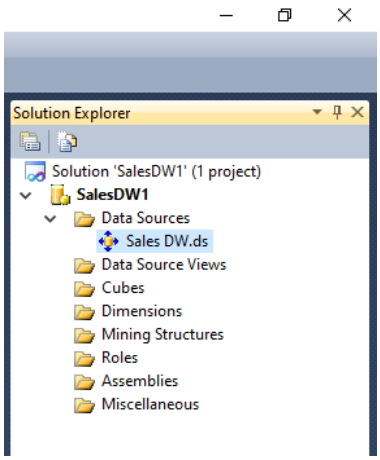
Check Inherit.



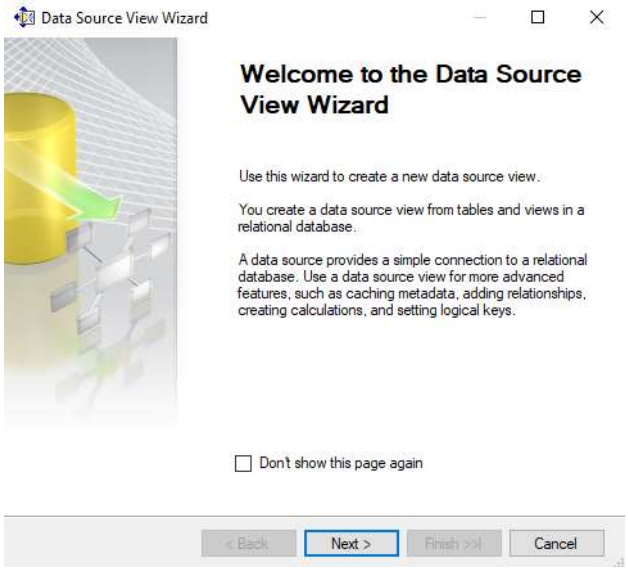
Click on Finish.



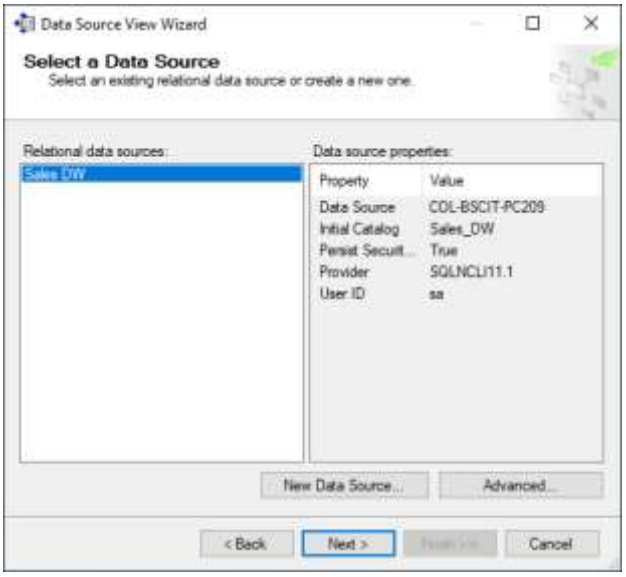
The data source is added in the Solution Explorer.



Again, open Data Source View Wizard click on Next.

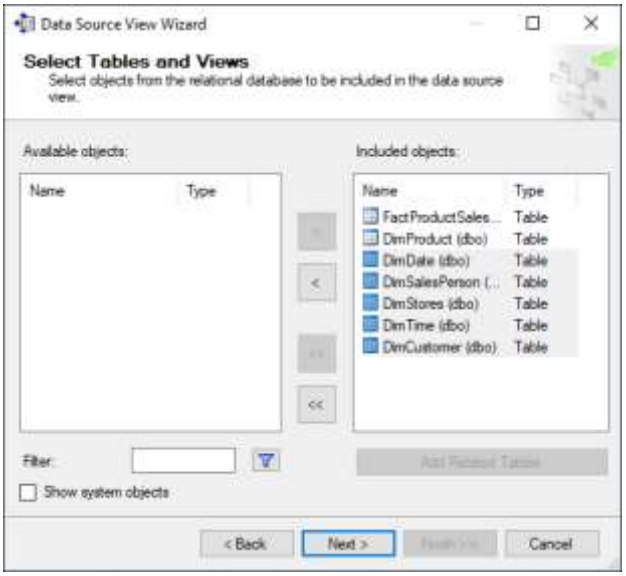


Click on Next.

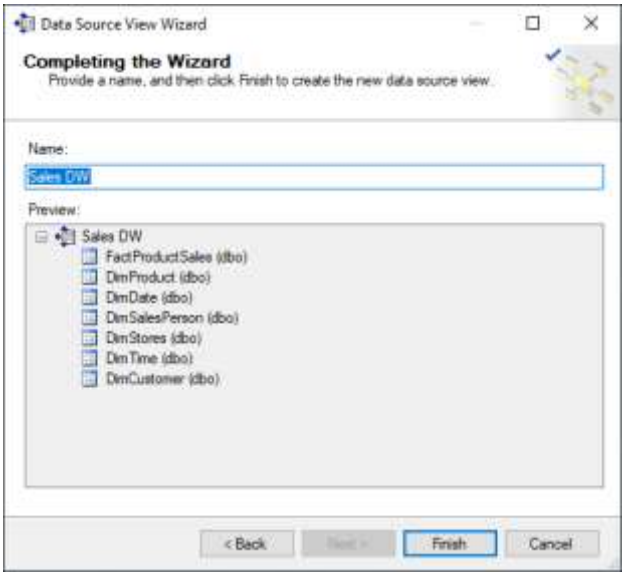


Select Add Related Tables.

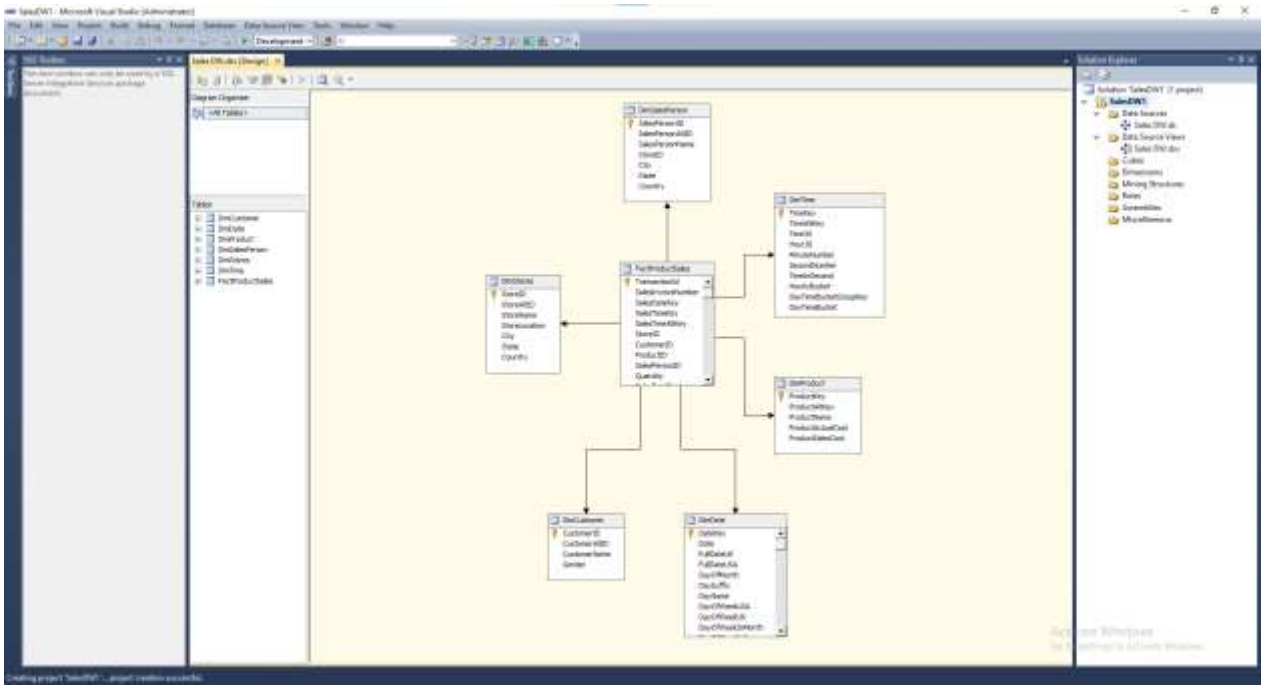
Click on Next.



Click on Finish.



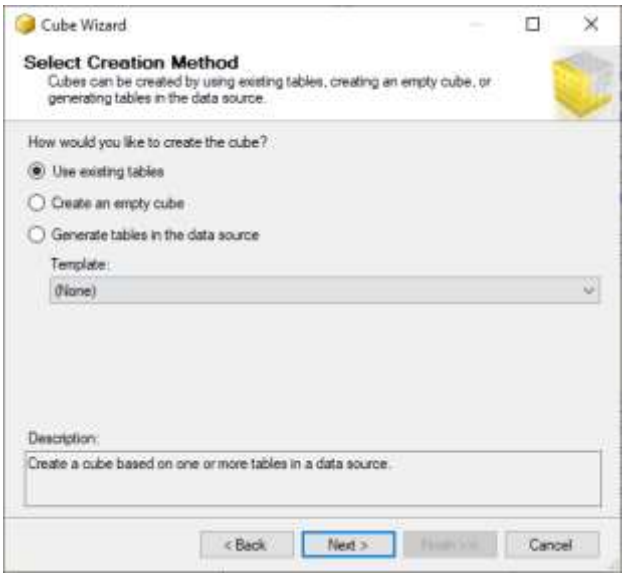
The relationship chart is now shown.



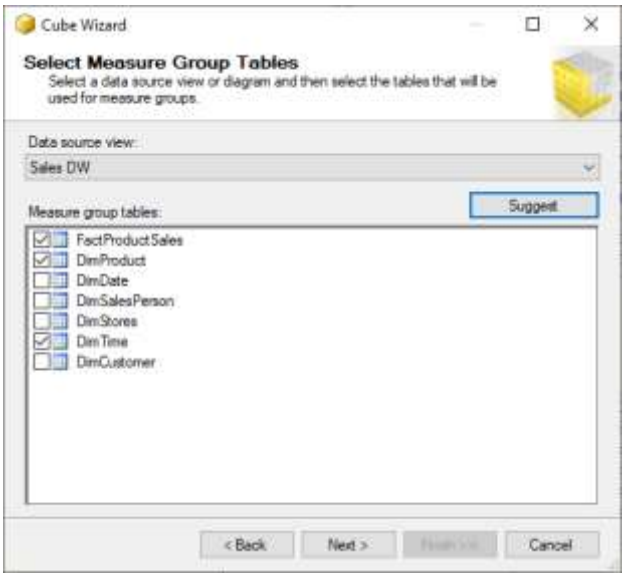
Practical 4A – Create the cube with corresponding dimension and fact table based on OLAP

Continue from the above practical.

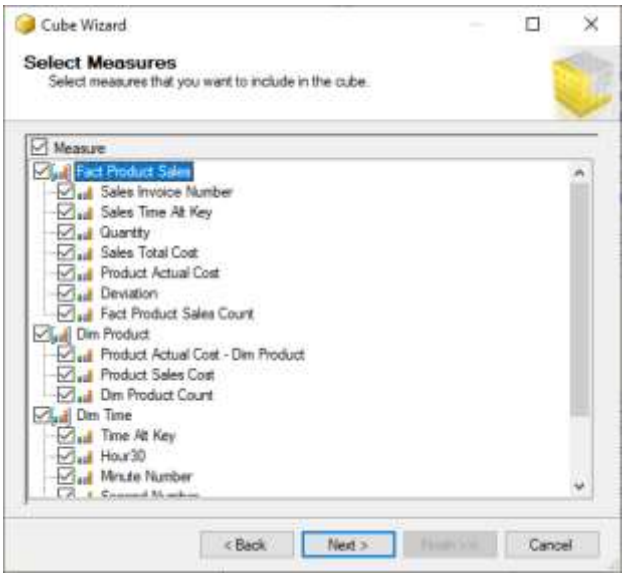
In Solution Explorer select Cube which opens a dialogue box.



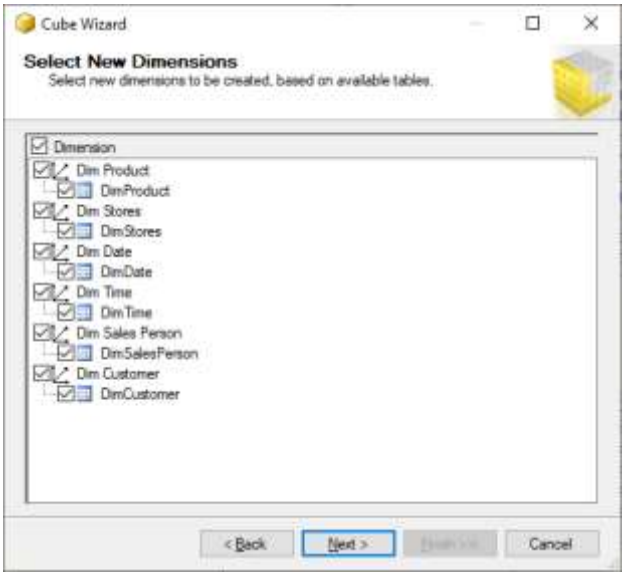
Select the tables – FactProductSales, DimProduct, DimTime and click on Next.



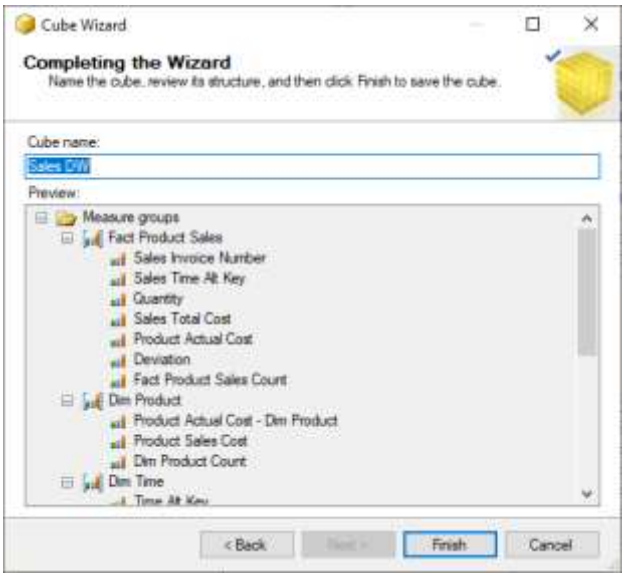
Select all the measures to include in the cube and click on Next.



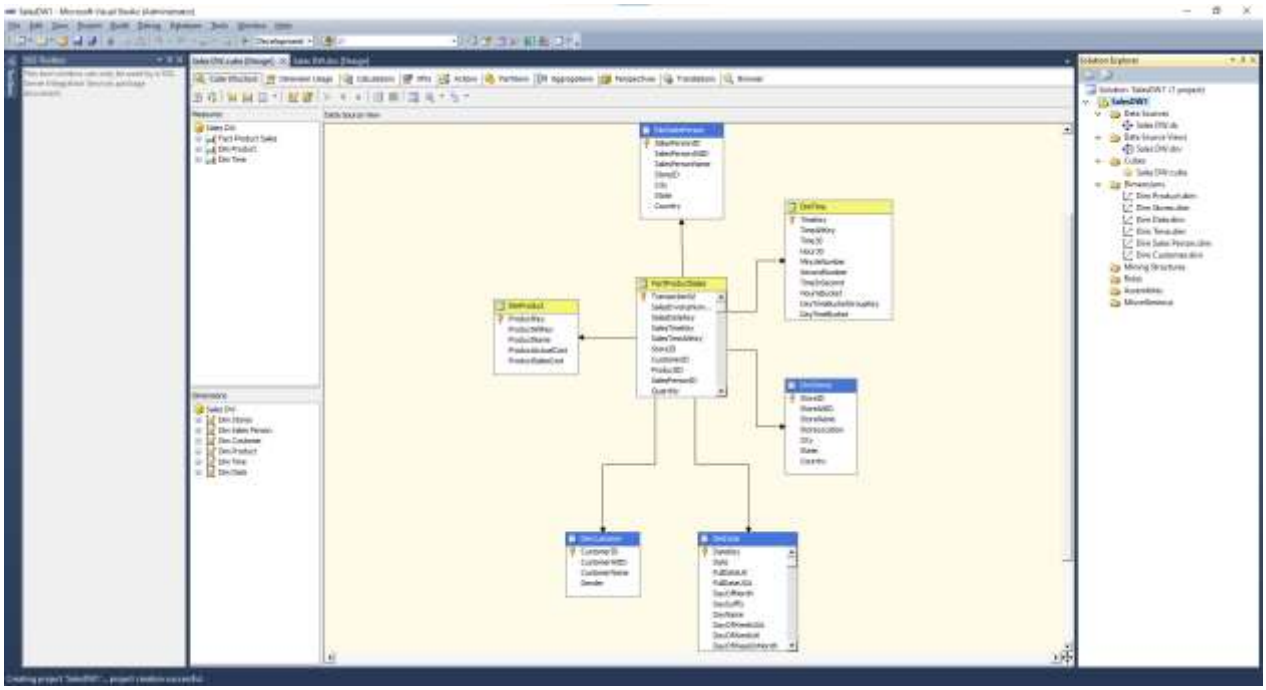
Select all the new dimensions and click on Next.



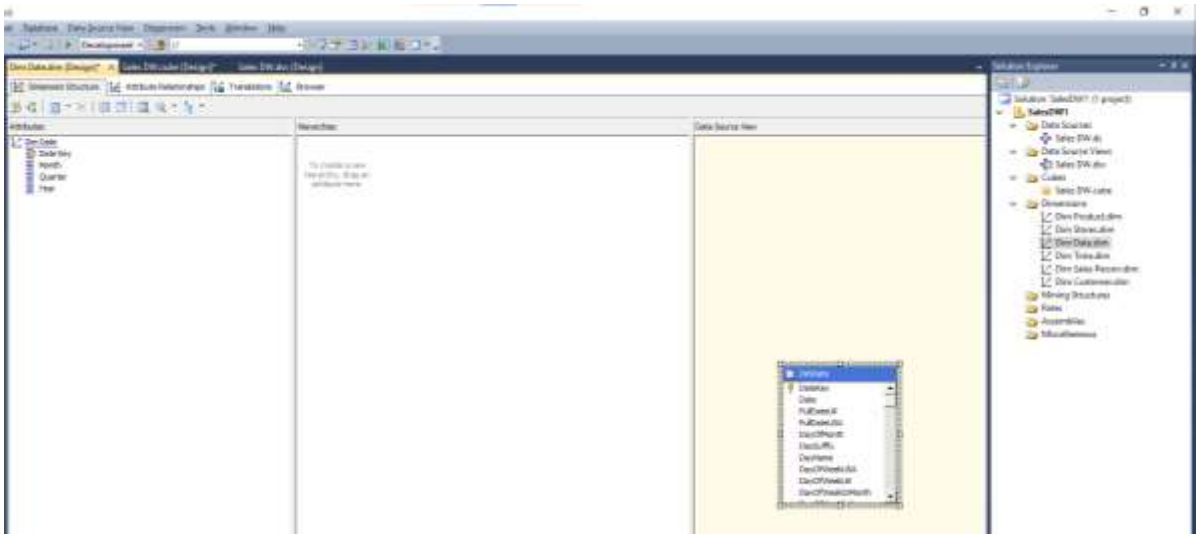
Name the cube as Sales DW and click on Finish.



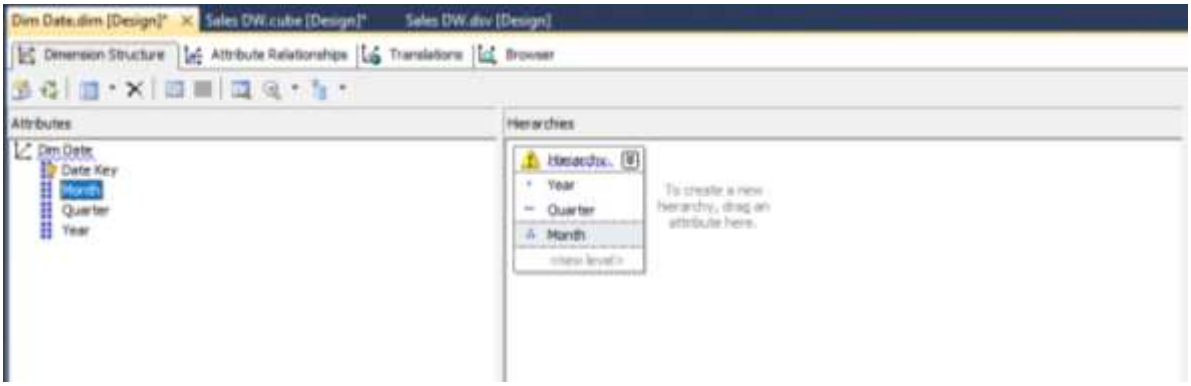
The relationship will be displayed on the screen.



Drag and drop Year, Quarter and Month to the Attributes.

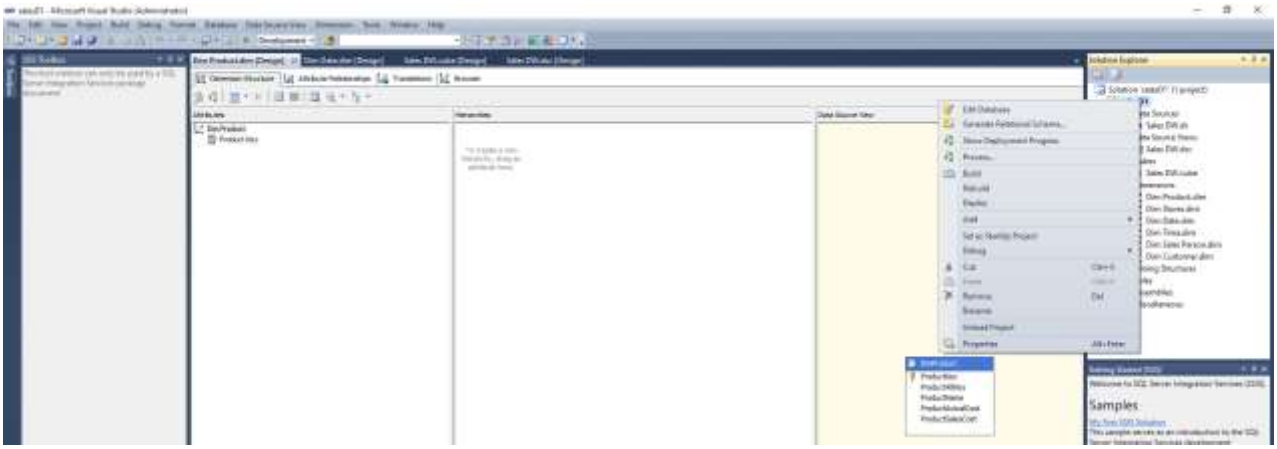


Drag and drop the attributes into the same Hierarchy.

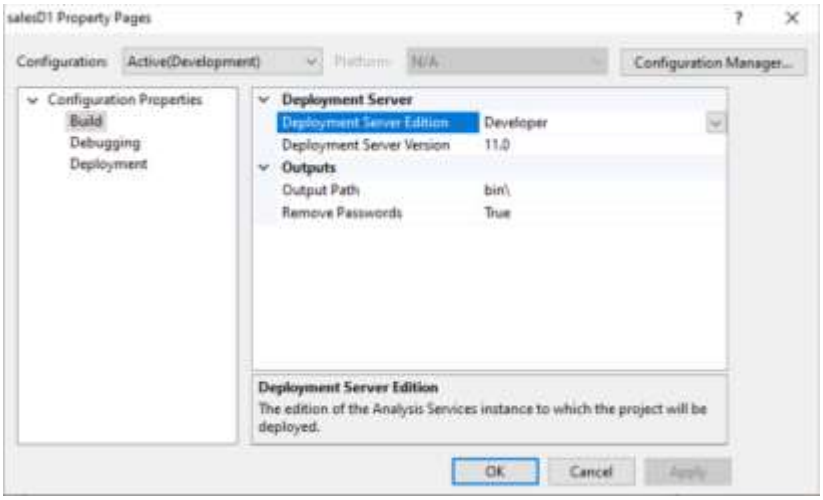


Save the progress and click on OK.

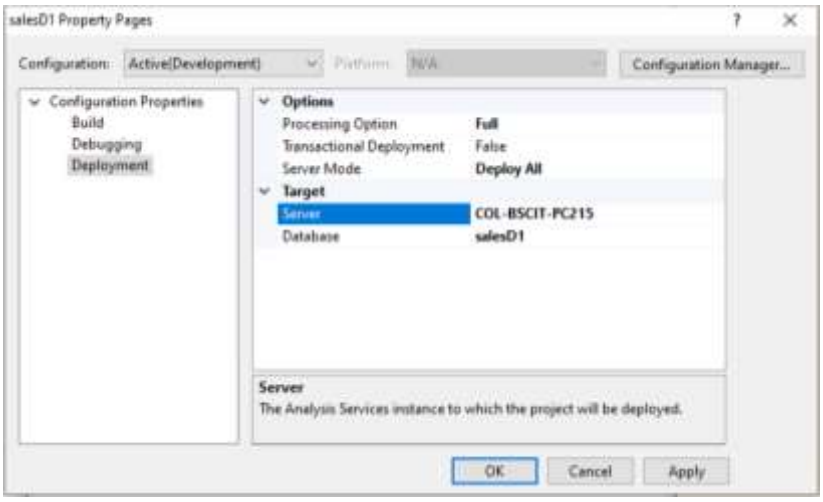
Right click on the file name and select Properties.



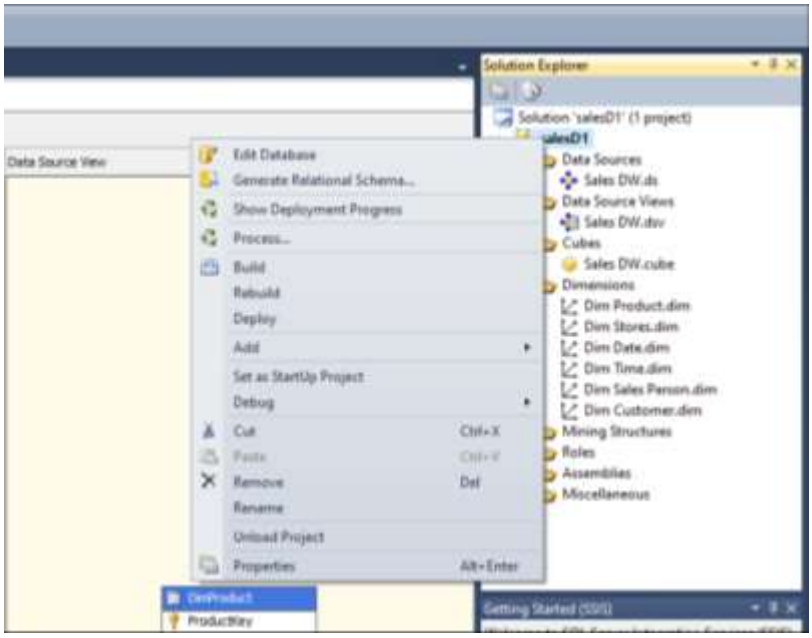
In Build make sure everything is correct.



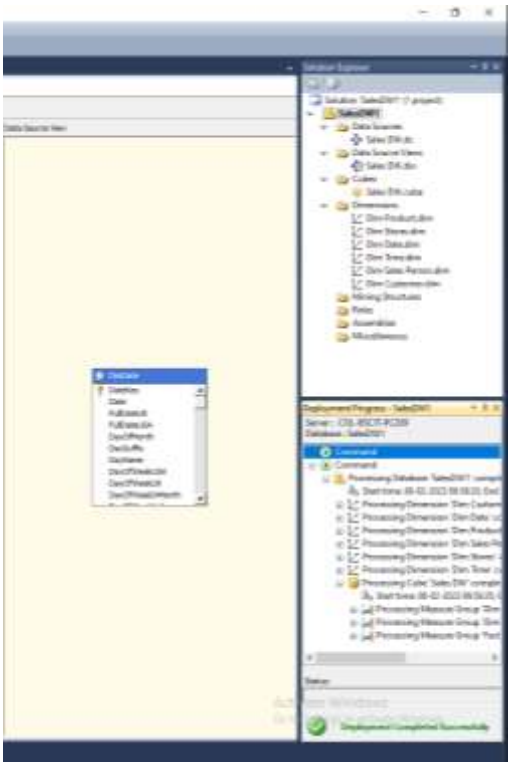
In Deployment, make Processing Option as Full instead of Default. Server Mode as Deploy All instead of Deploy Changes Only. Server as the server name used till now instead of localhost. Apply these changes.



Right click on the file name and click on Deploy to deploy the project.



At the bottom of the screen it is shown that the deployment is successful.

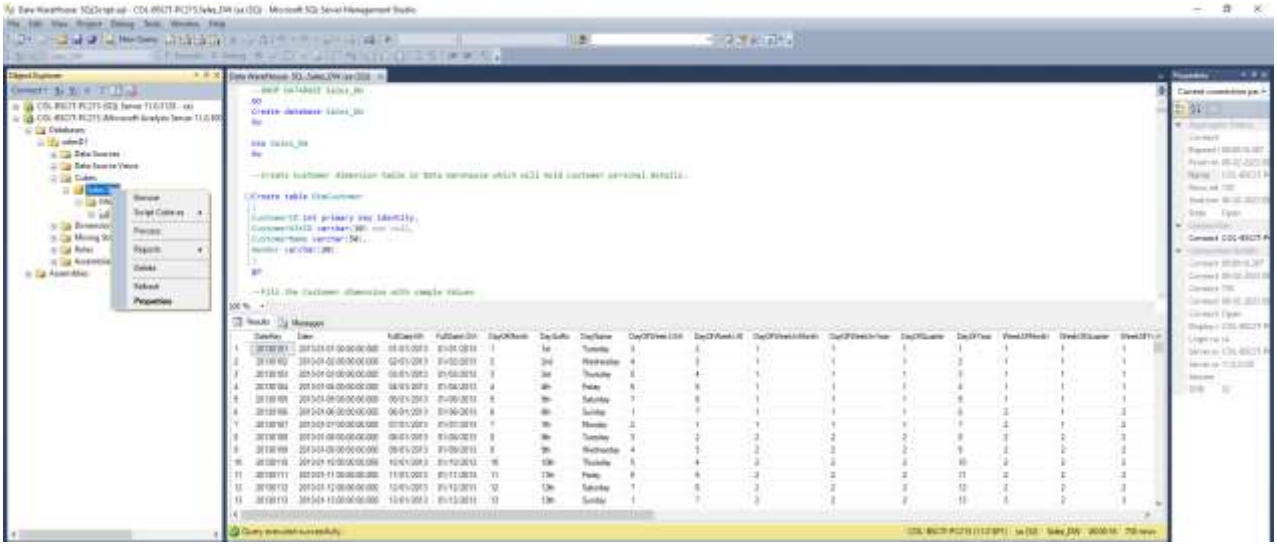


KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

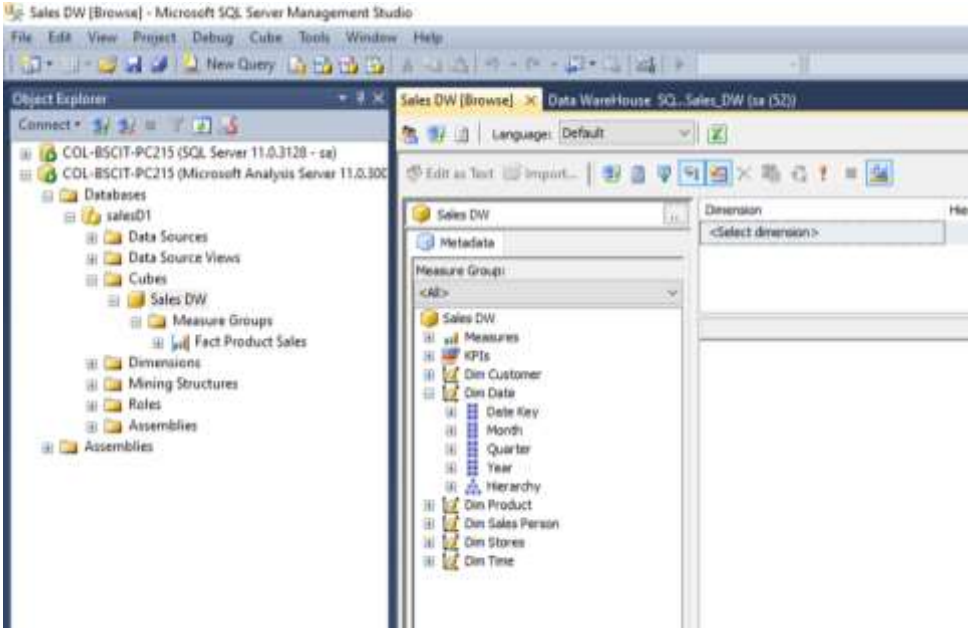
Practical 4B – Perform Multi-Dimensional Expressions (MDX) queries for OLAP database

Continue from the above practical.

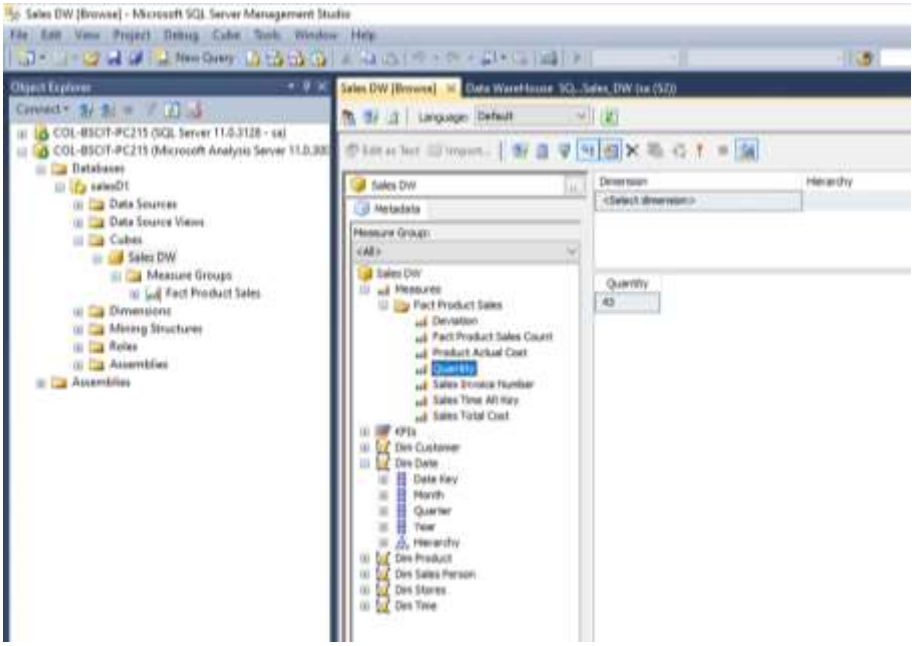
To verify, in SQL Server Management right click on the file name in Object Explorer and click on Browse.



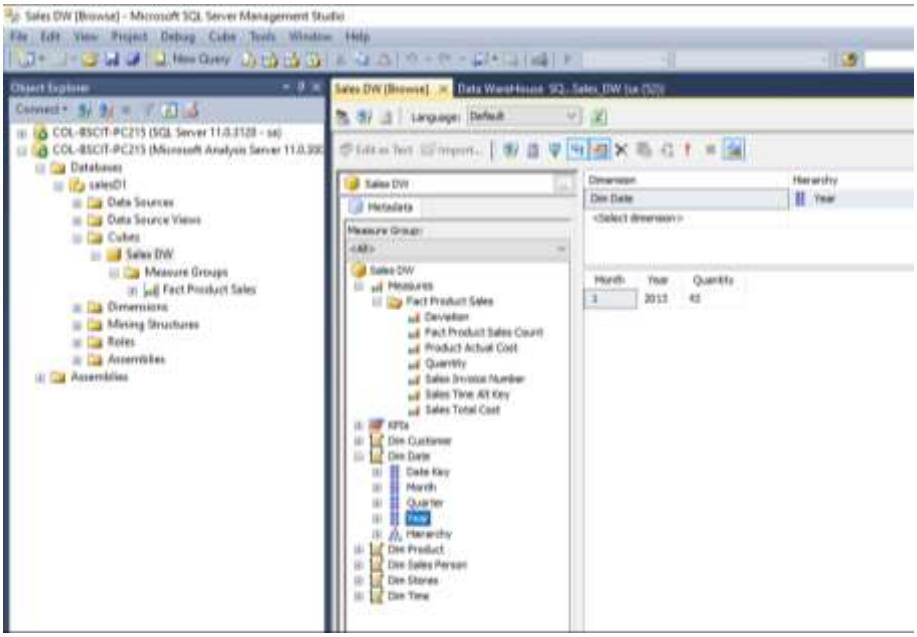
Enlarge the Measures of SalesDW.



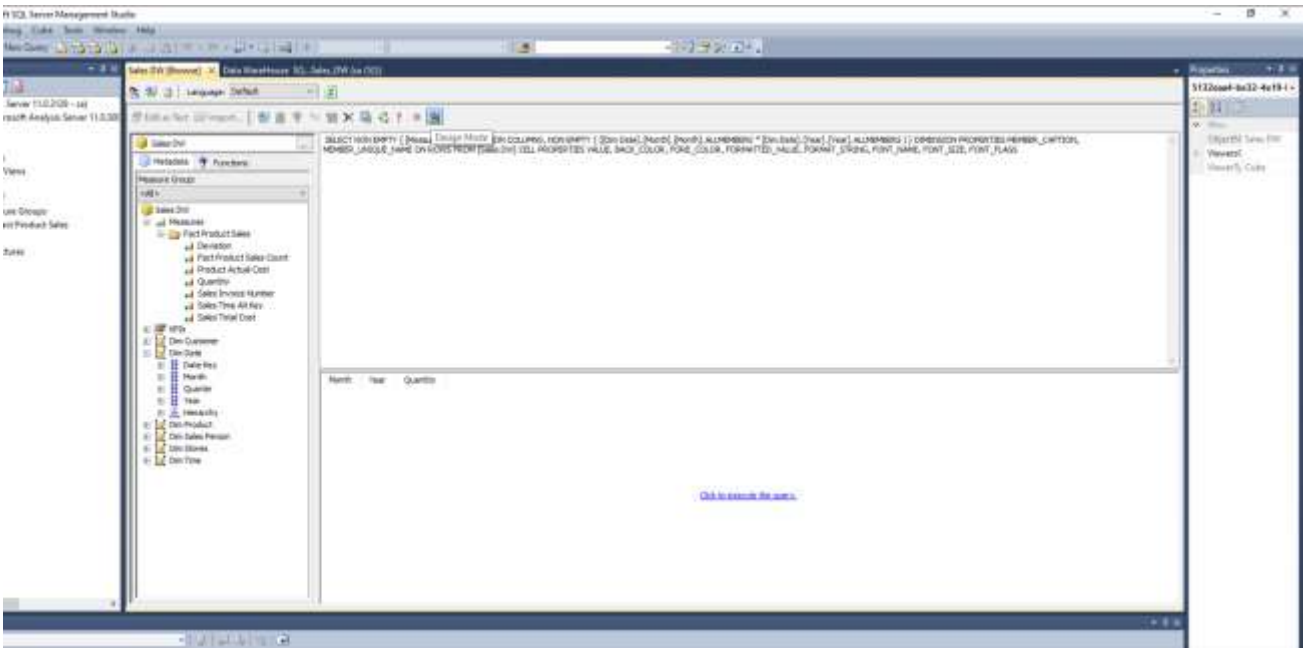
On selecting Quantity under Fact Product Sales the value shown will be 43.



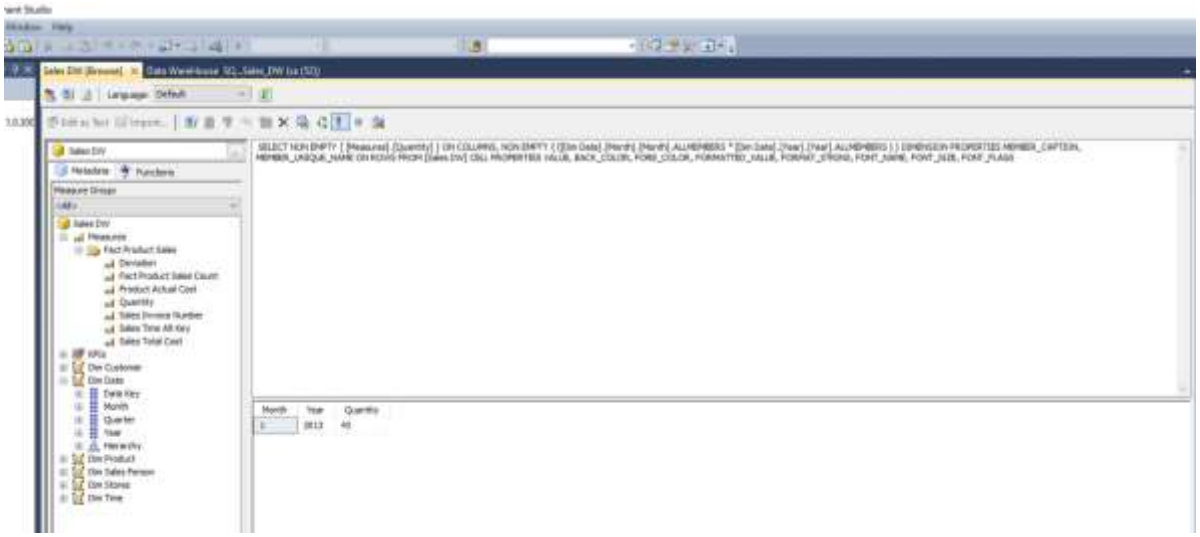
Select Month and Year as well from Dim Date.



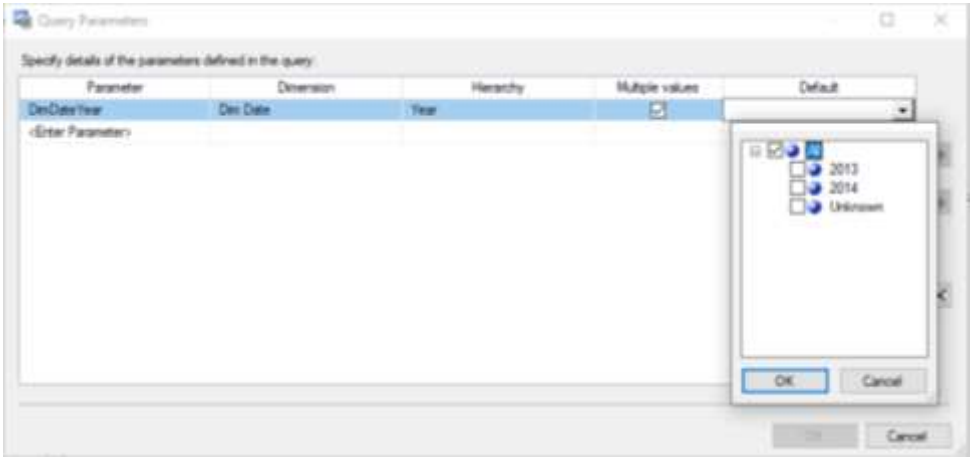
Go in Design Mode and click on the line that says 'Click to execute the query.'

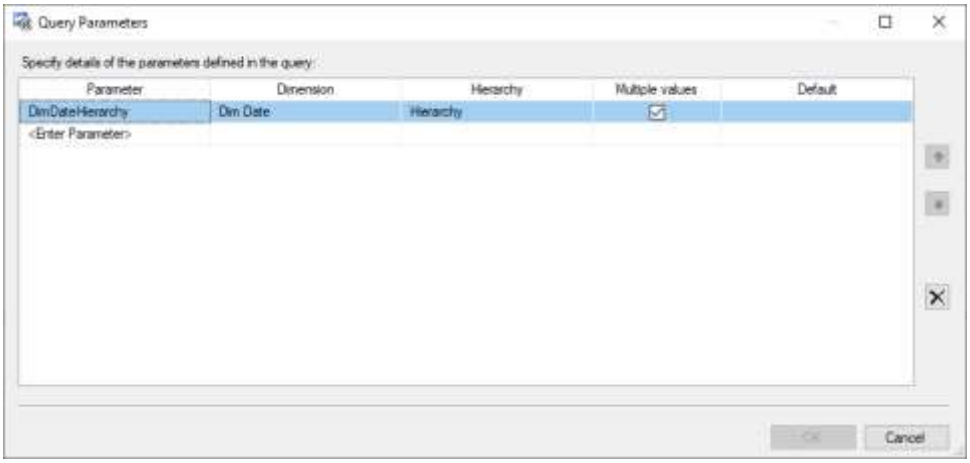


Click on the error symbol.

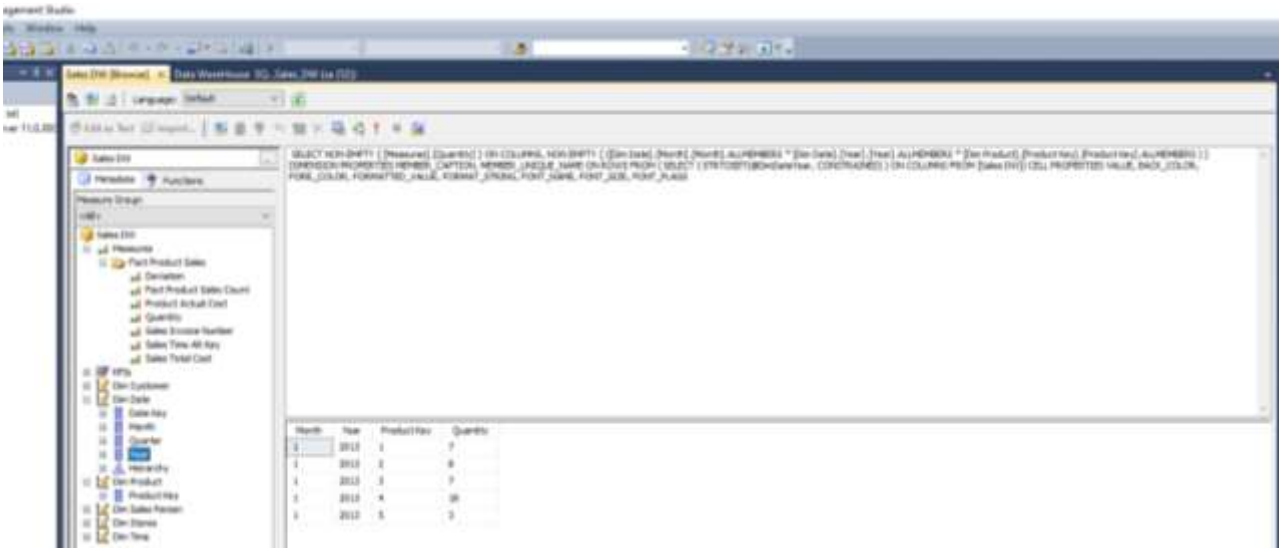


Click on OK.

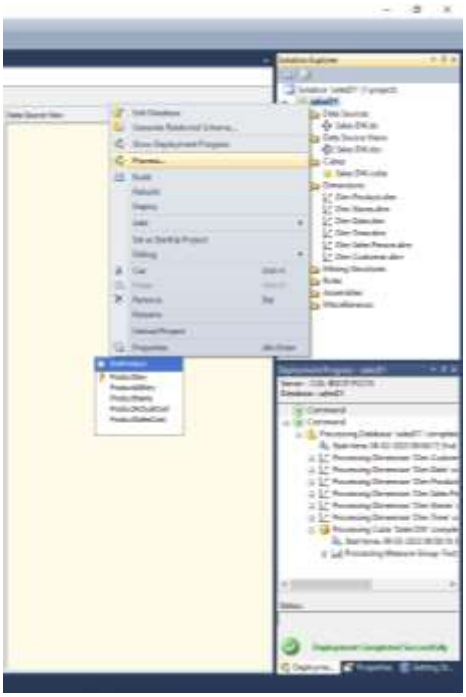




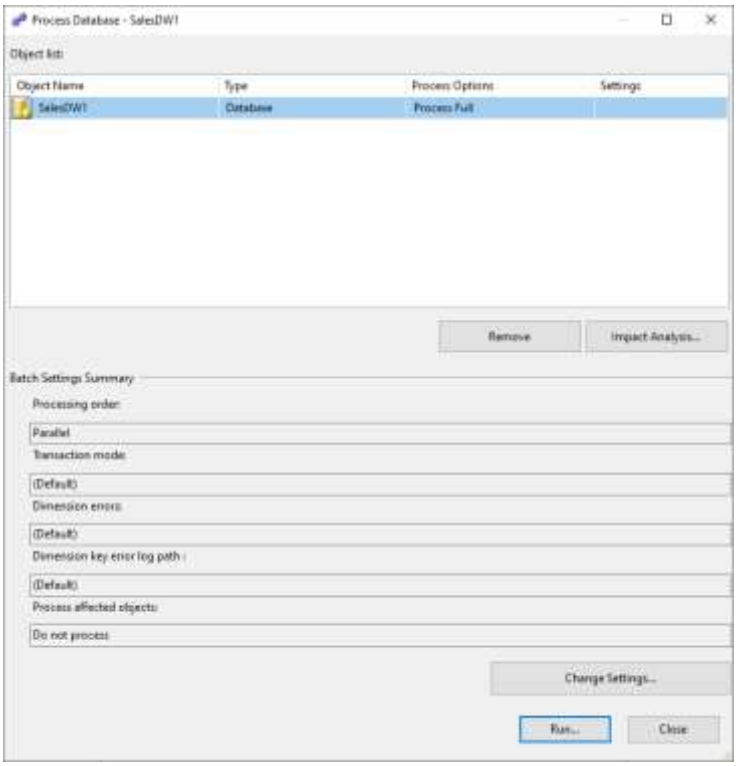
The entries will be shown on the screen.



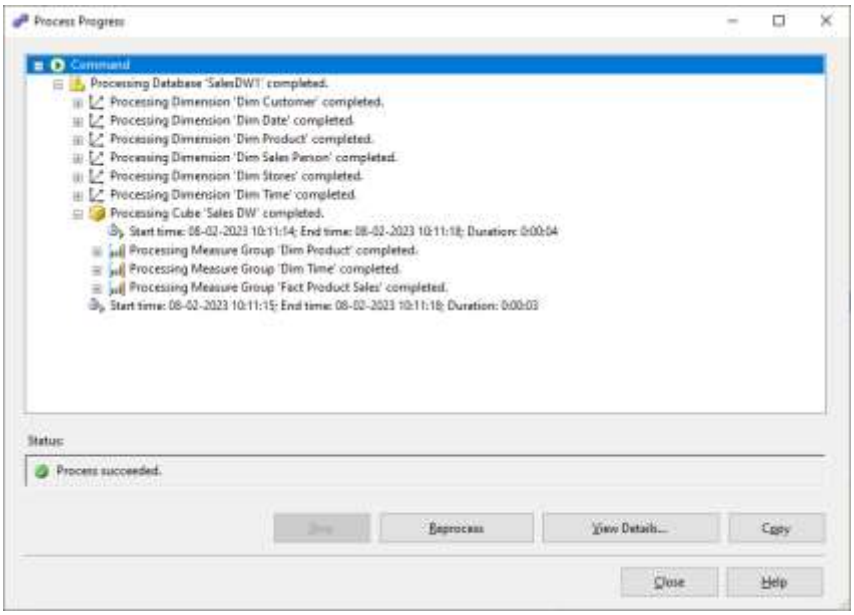
Right click on the file name and select Process.



Select the file and click on Run.



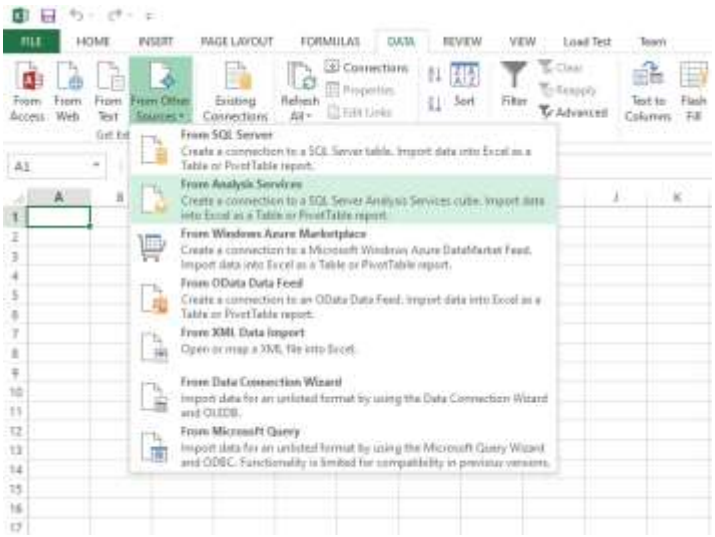
Process succeeded is the status.



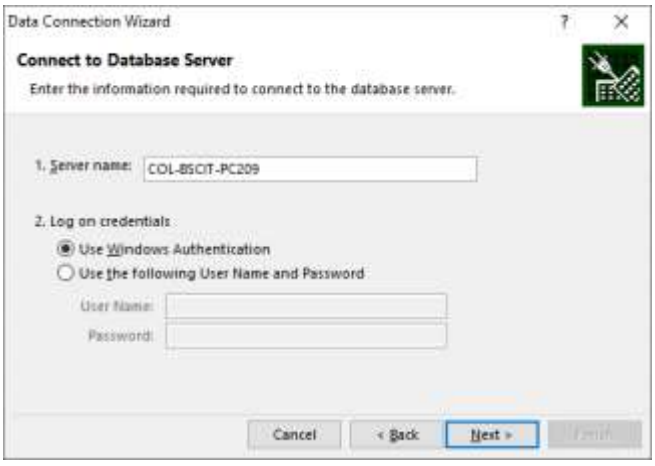
KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

Practical 5A – Import the data warehouse in Microsoft Excel and create Pivot Table to perform Data Analytics

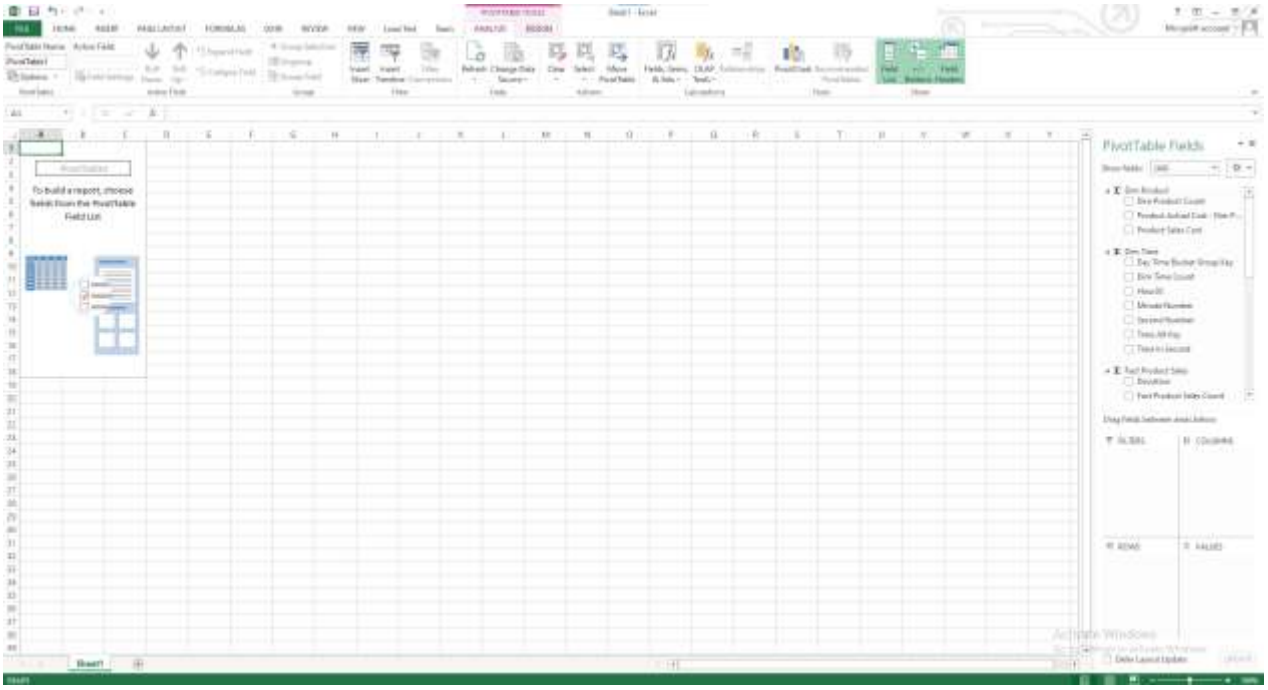
Open Microsoft Excel, go to Data, in From Other Sources select From Analysis Services.



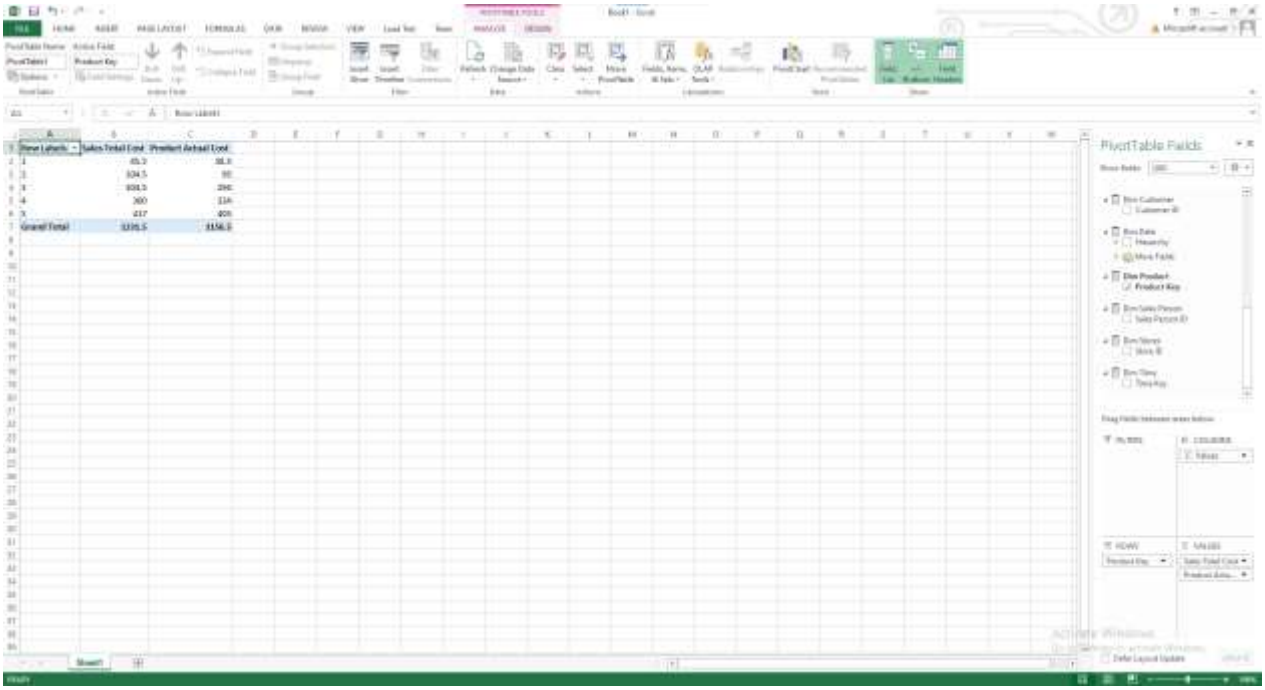
Enter your Server name and click on Next.



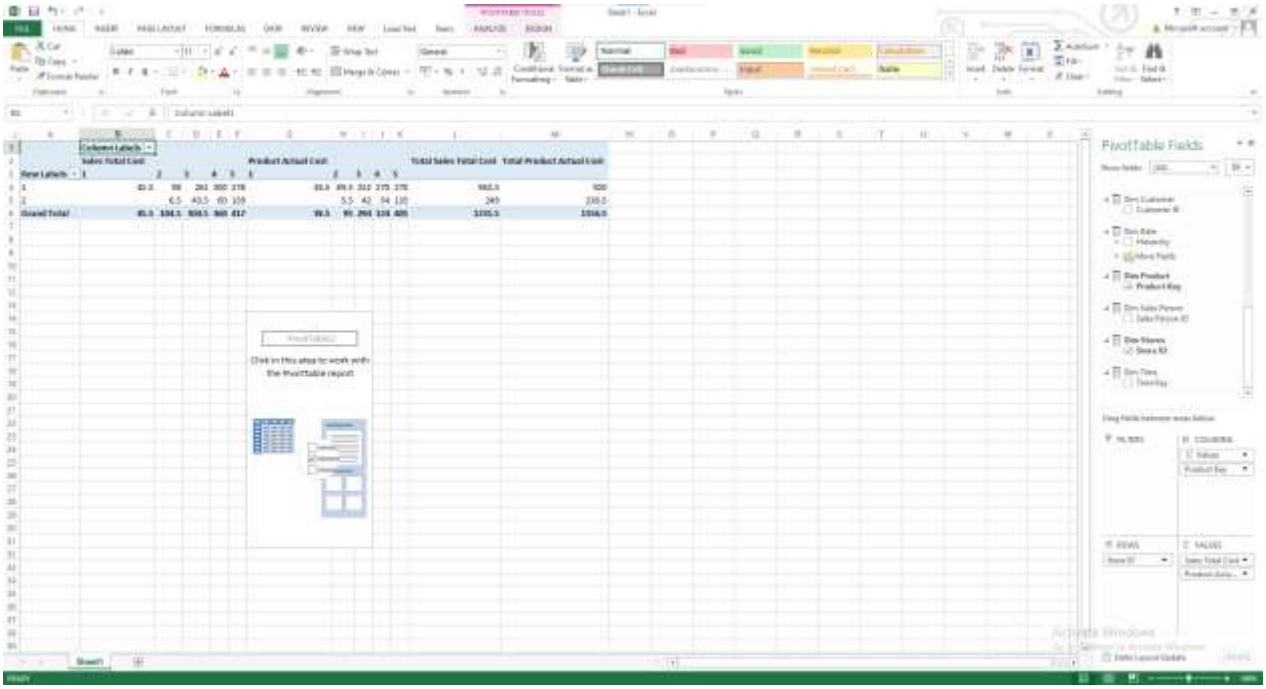
Insert a Pivot Table.



Select the row labels for the pivot table.



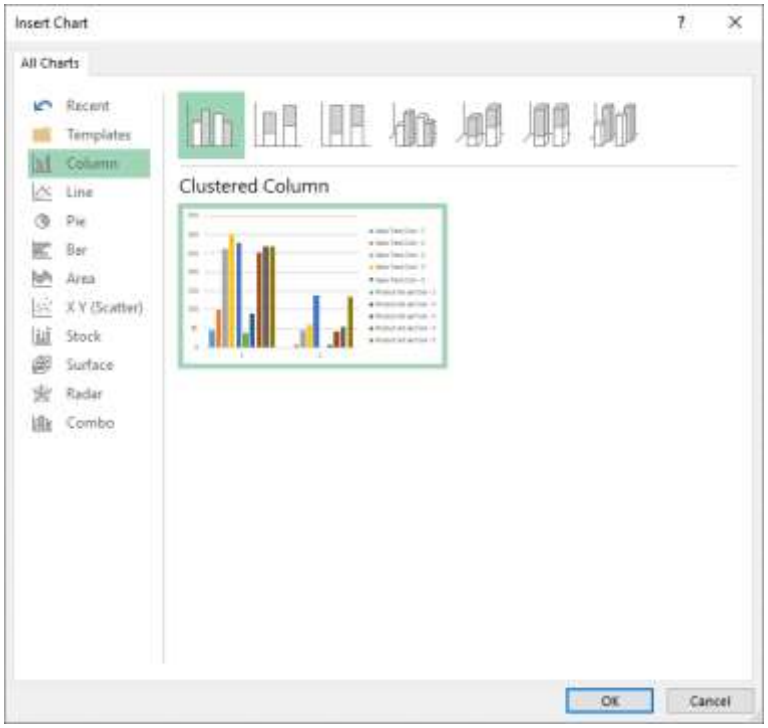
Select the column labels for the pivot table.



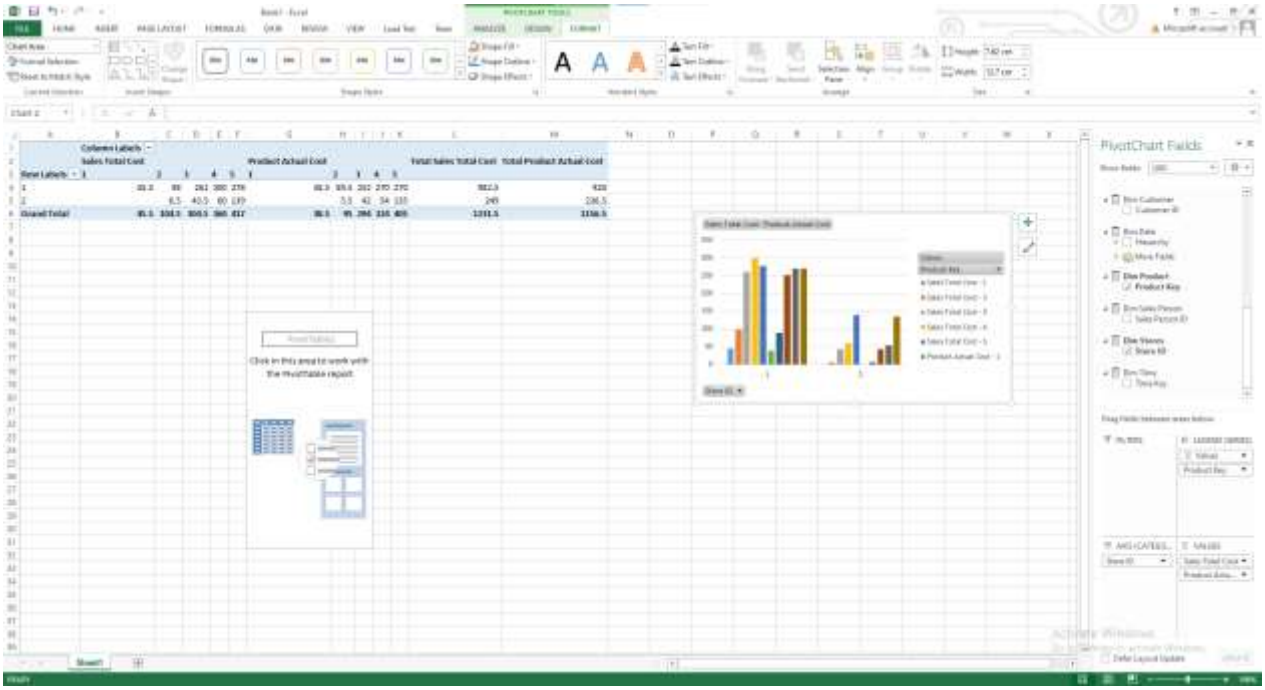
KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

Practical 5B – Import the cube in Microsoft Excel and create the Pivot Chart to perform Data Modelling

Open Microsoft Excel containing a Pivot Table and insert a clustered column chart.



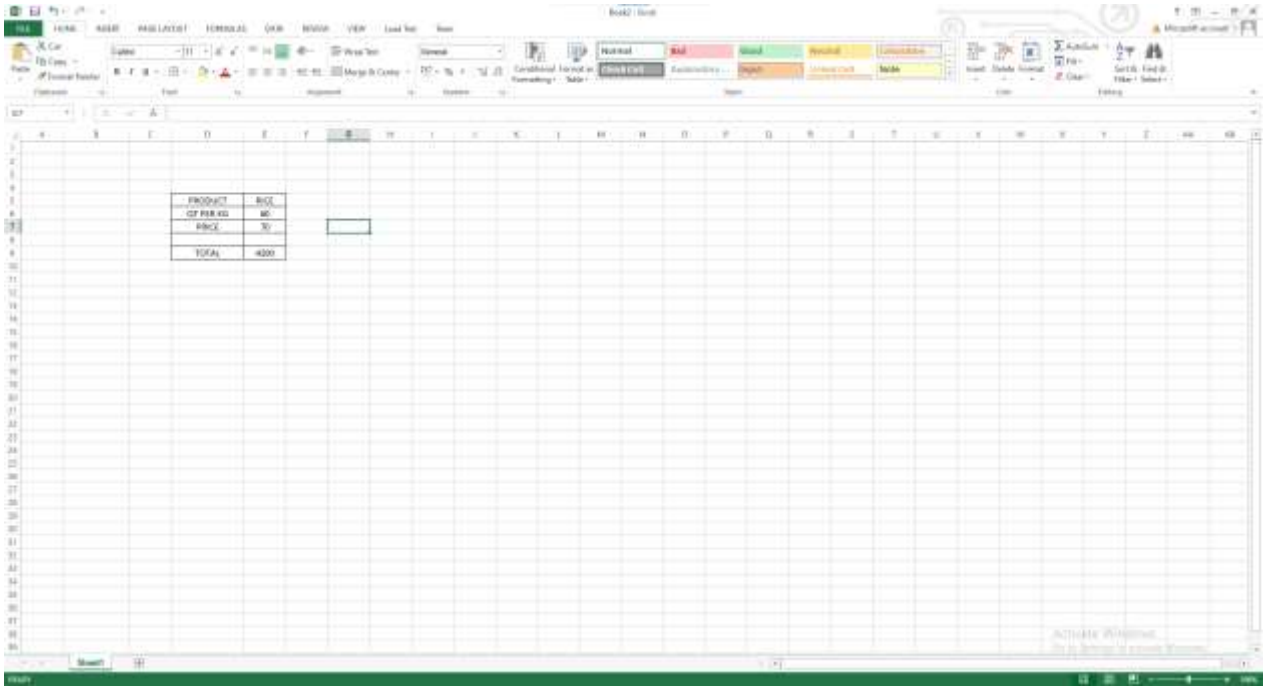
The clustered column chart is shown with the selected database.



KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

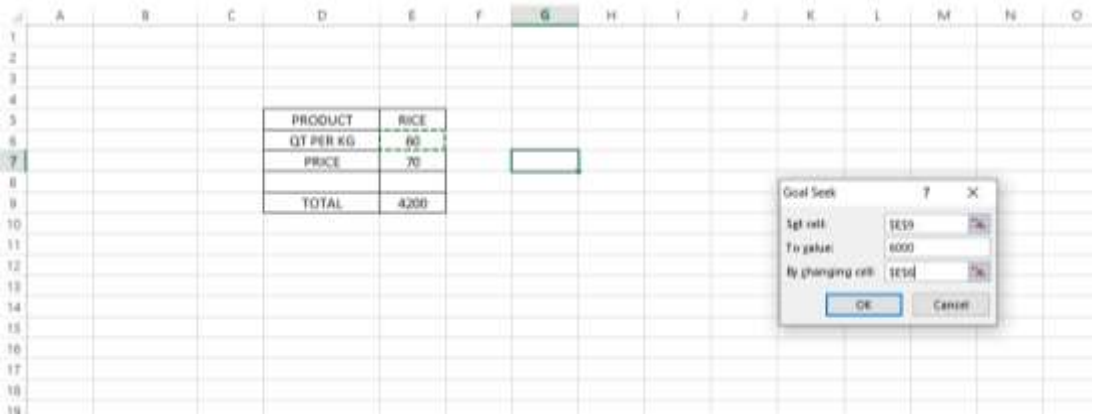
Practical 6 – Perform the What-If Analysis for Data Visualization

Open Microsoft Excel. Create a table.

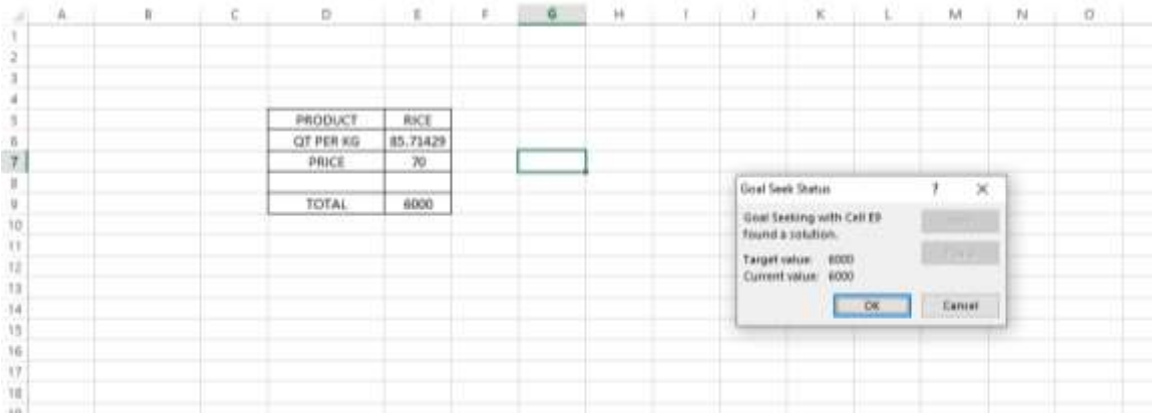


In Data tab, select Goal Seek under What-If Analysis.

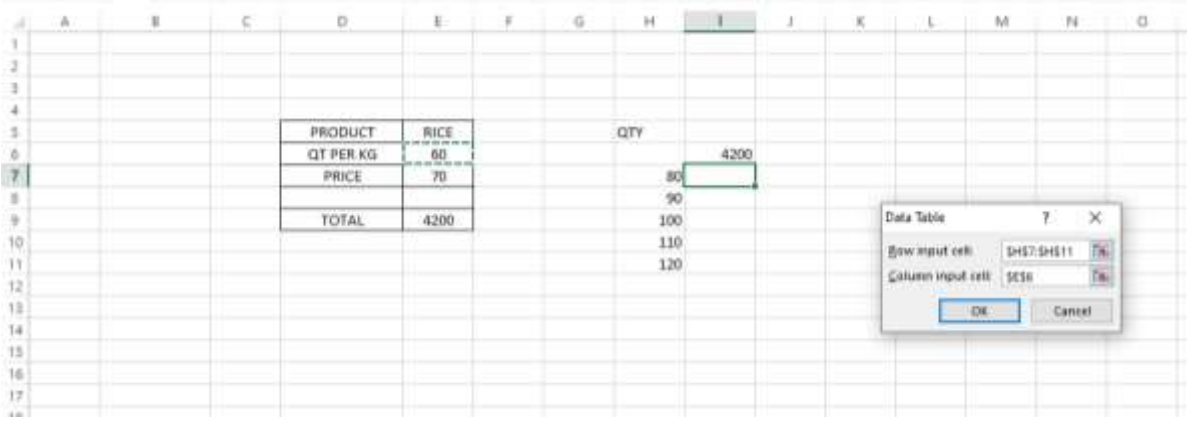
Enter the Set Cell value to where the changed value will be displayed. Enter the value that should be changed in To Value and the cell which should be changed in By changing cell.



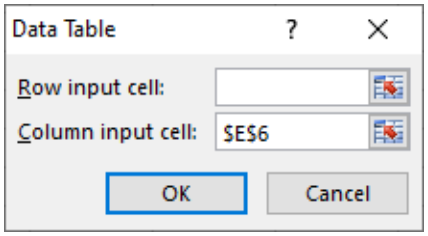
Click on OK in Goal Seek Status.



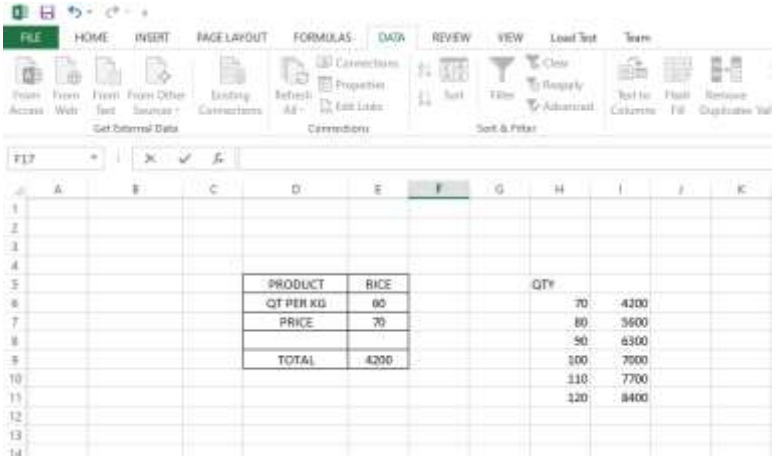
Select Data Table under What-If Analysis.



Enter the input cells and click on OK.



The result will be displayed.



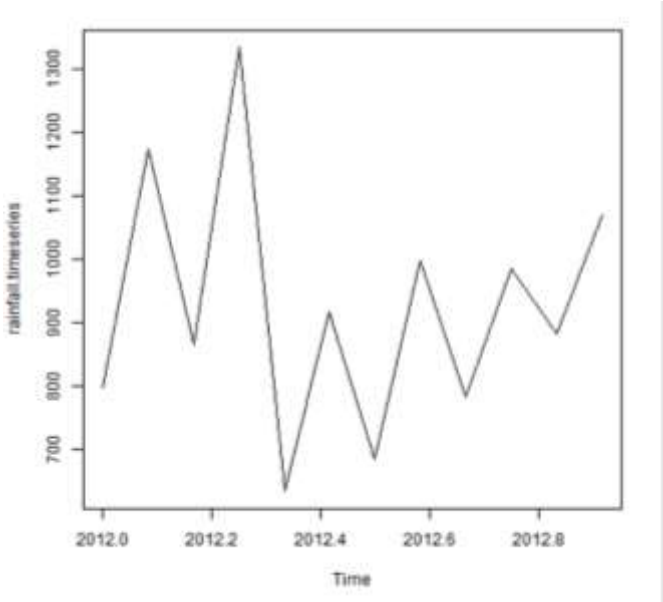
KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

Practical 7 – Data Analysis using Time Series Analysis

Open RGui. Type the code and run it.

```
RGui (64-bit) - [R Console]
File  Edit  View  Misc  Packages  Windows  Help

R version 4.2.2 (2022-10-31 ucrt) -- "Innocent and Trusting"
Copyright (C) 2022 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

  Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

Warning: namespace 'modeltools' is not available and has been replaced
by .GlobalEnv when processing object 'output.tree'
Warning: namespace 'party' is not available and has been replaced
by .GlobalEnv when processing object 'output.tree'
[Previously saved workspace restored]

> rainfall <-
+ c(799,1174.8,865.1,1334.6,635.4,918.5,685.5,998.6,784.2,985,882.8,1071)
> rainfall.timeseries <- ts(rainfall,start = c(2012,1),frequency = 12)
> print(rainfall.timeseries)
       Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct
2012 799.0 1174.8 865.1 1334.6 635.4 918.5 685.5 998.6 784.2 985.0
       Nov   Dec
2012 882.8 1071.0
> png(file = "rainfall.png")
> # Get the data points in form of a R vector.
> rainfall <-
+ c(799,1174.8,865.1,1334.6,635.4,918.5,685.5,998.6,784.2,985,882.8,1071)
> # Convert it to a time series object.
> rainfall.timeseries <- ts(rainfall,start = c(2012,1),frequency = 12)
> # Print the timeseries data.
> print(rainfall.timeseries)
       Jan   Feb   Mar   Apr   May   Jun   Jul   Aug   Sep   Oct
2012 799.0 1174.8 865.1 1334.6 635.4 918.5 685.5 998.6 784.2 985.0
       Nov   Dec
2012 882.8 1071.0
> # Give the chart file a name.
> png(file = "rainfall.png")
> # Plot a graph of the time series.
> plot(rainfall.timeseries)
> # Save the file.
> dev.off()
png
  2
> |
```

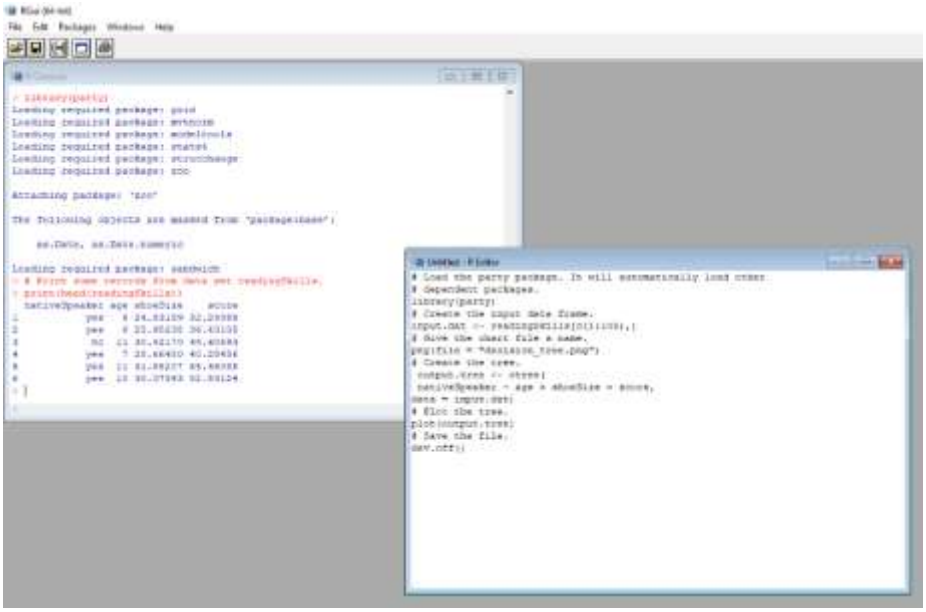
The output will be shown as this.



KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

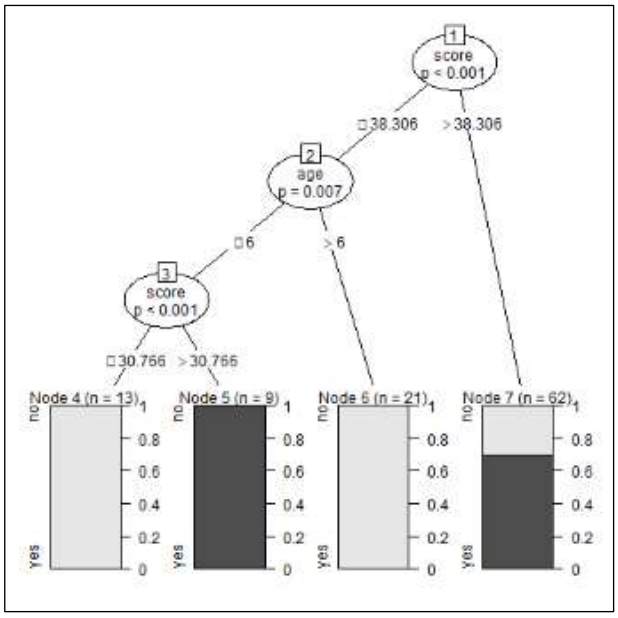
Practical 8 – Implementation of Decision Tree using R Tool

Open RGui. Type the code and run it.



```
RGui (64-bit) - [R Console]
File  Edit  View  Misc  Packages  Windows  Help

trying URL 'https://cloud.r-project.org/bin/windows/contrib/4.2/strucchange_1.5-3.zip'
Content type 'application/zip' length 947691 bytes (925 KB)
downloaded 925 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/4.2/coin_1.4-2.zip'
Content type 'application/zip' length 1423472 bytes (1.4 MB)
downloaded 1.4 MB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/4.2/zoo_1.8-11.zip'
Content type 'application/zip' length 1019155 bytes (995 KB)
downloaded 995 KB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/4.2/sandwich_3.0-2.zip'
Content type 'application/zip' length 1458564 bytes (1.4 MB)
downloaded 1.4 MB

trying URL 'https://cloud.r-project.org/bin/windows/contrib/4.2/party_1.3-11.zip'
Content type 'application/zip' length 696117 bytes (679 KB)
downloaded 679 KB

package 'TH.data' successfully unpacked and MD5 sums checked
package 'libcoin' successfully unpacked and MD5 sums checked
package 'matrixStats' successfully unpacked and MD5 sums checked
package 'multcomp' successfully unpacked and MD5 sums checked
package 'mvtnorm' successfully unpacked and MD5 sums checked
package 'modeltools' successfully unpacked and MD5 sums checked
package 'strucchange' successfully unpacked and MD5 sums checked
package 'coin' successfully unpacked and MD5 sums checked
package 'zoo' successfully unpacked and MD5 sums checked
package 'sandwich' successfully unpacked and MD5 sums checked
package 'party' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\COL-BSCIT-PC41\AppData\Local\Temp\RtmpmE3pGw\downloaded_packages
> library(party)
Loading required package: grid
Loading required package: mvtnorm
Loading required package: modeltools
Loading required package: stats4
Loading required package: strucchange
Loading required package: zoo

Attaching package: 'zoo'

The following objects are masked from 'package:base':

    as.Date, as.Date.numeric

Loading required package: sandwich
> print(head(readingSkills))
  nativeSpeaker age shoeSize    score
1           yes   5 24.83189 32.29385
2           yes   6 25.95238 36.63105
3            no  11 30.42170 49.60593
4           yes   7 28.66450 40.28456
5           yes  11 31.81207 55.46085
6           yes  10 30.07843 52.83124
>
```

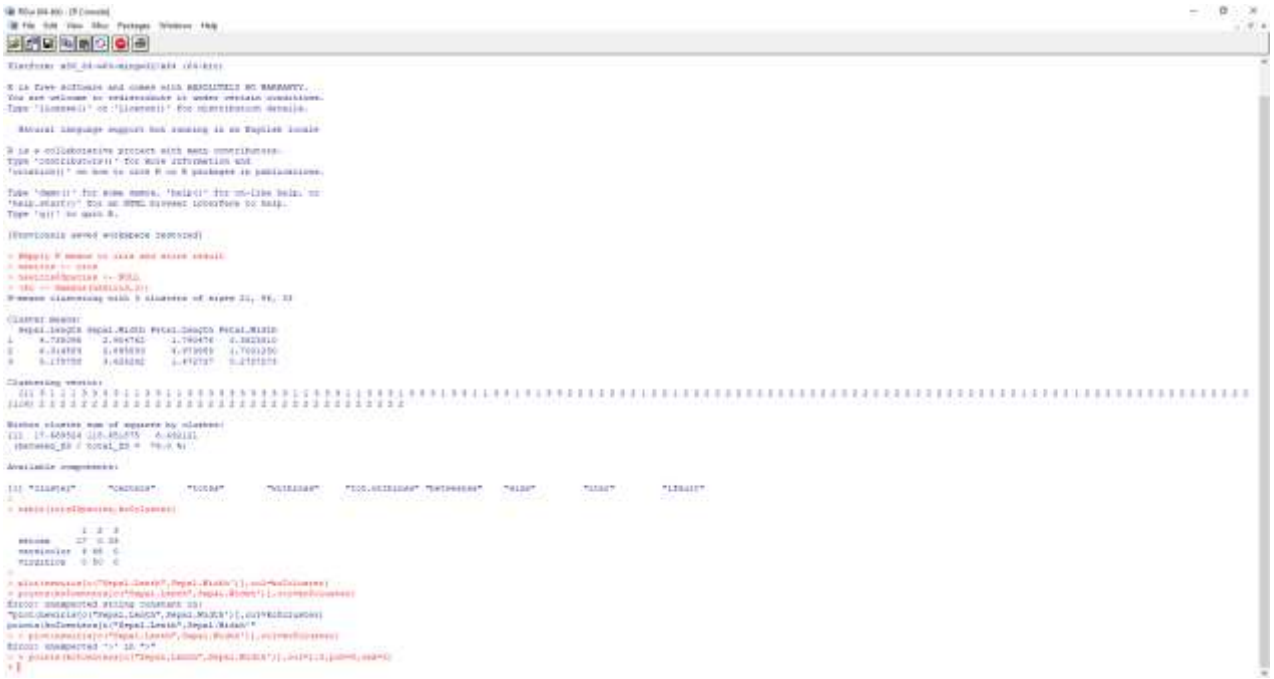
The output will be shown as this.



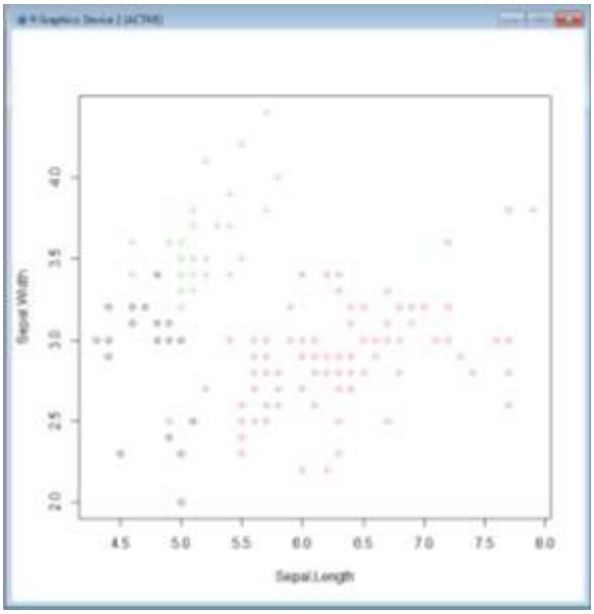
KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

Practical 9 – K-Means Clustering using R Programming

Open RGui. Type the code and run it.



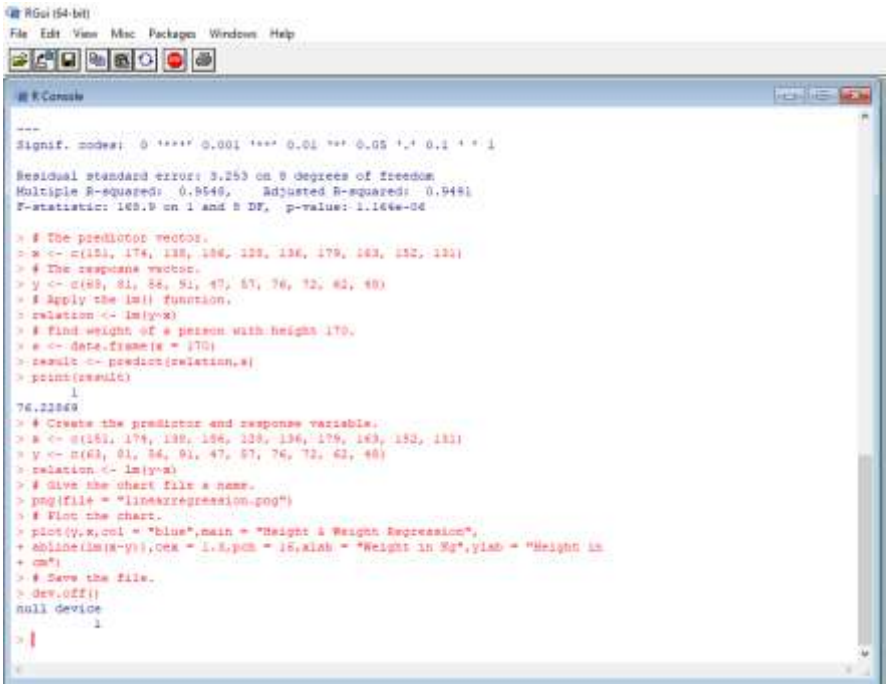
The output will be shown as this.



KISHINCHAND CHELLARAM COLLEGE, MUMBAI - 20
T.Y. B.Sc. I.T. Semester VI

Practical 10 – Prediction using Linear Regression

Open RGui. Type the code and run it.



The output will be shown as this.