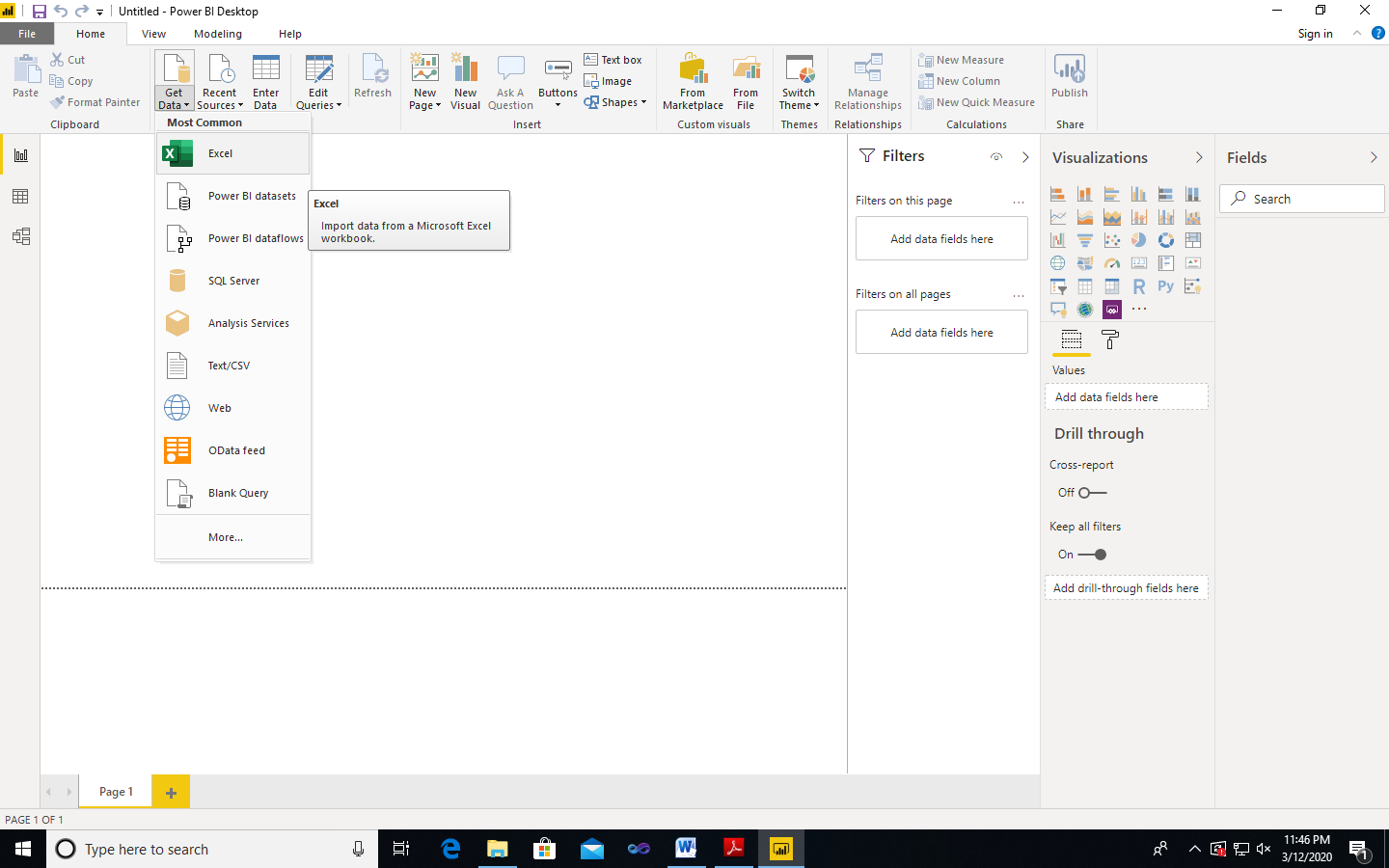
|  |
| --- |
| Description: H:\Uni logo\unilogo1[1].jpg  **UNIVERSITY OF MUMBAI’S**  **Vijayalakshmi Vishwanath Dalvie COLLEGE**  **Talere, dist- sindhudurg**  Department of information technology  2019-2020  **CERTIFICATE**  This to certify that Miss. **Vishakha Santosh Terwankar** of the **T. Y. B.Sc. I. T.** class and Exam seat number **4027409** has satisfactorily carried out the practical work in **Business Intelligence** as laid down by the University of Mumbai.  This Journal represents her bonafied work for **Semester VI** during the year 2019-2020.  **TEACHER INCHARGE HEAD OF DEPARTMENT** |

|  |  |  |  |
| --- | --- | --- | --- |
| **INDEX** | | | |
| **Sr. No.** | **Name of Practical** | **Date** | **Sign** |
| **1.** | **Import the legacy data from different sources such as ( Excel , SQL Server, Oracle etc.) and load in the target system** | **06/12/2019** |  |
| **2.** | **Perform the Extraction Transformation and Loading (ETL) process to construct the database in the SQL server / Power BI** | **13/12/2019** |  |
| **3.** | **Data Visualization from ETL Process** | **20/12/2019** |  |
| **4.** | **Data Modelling and Analytics with Pivot Table and Pivot Chart in Excel.** | **10/01/2020** |  |
| **5.** | **Apply the what – if Analysis for data visualization. Design and generate necessary reports based on the data warehouse data** | **24/01/2020** |  |
| **6.** | **Implementation of Classification algorithm in R Programming** | **07/02/2020** |  |
| **7.** | **Practical Implementation of Decision Tree using R Tool** | **14/02/2020** |  |
| **8.** | **k-means clustering using R** | **21/02/2020** |  |
| **9.** | **Prediction Using Linear Regression** | **28/02/2020** |  |
| **10.** | **Data Analysis using Time Series Analysis** | **06/03/2020** |  |

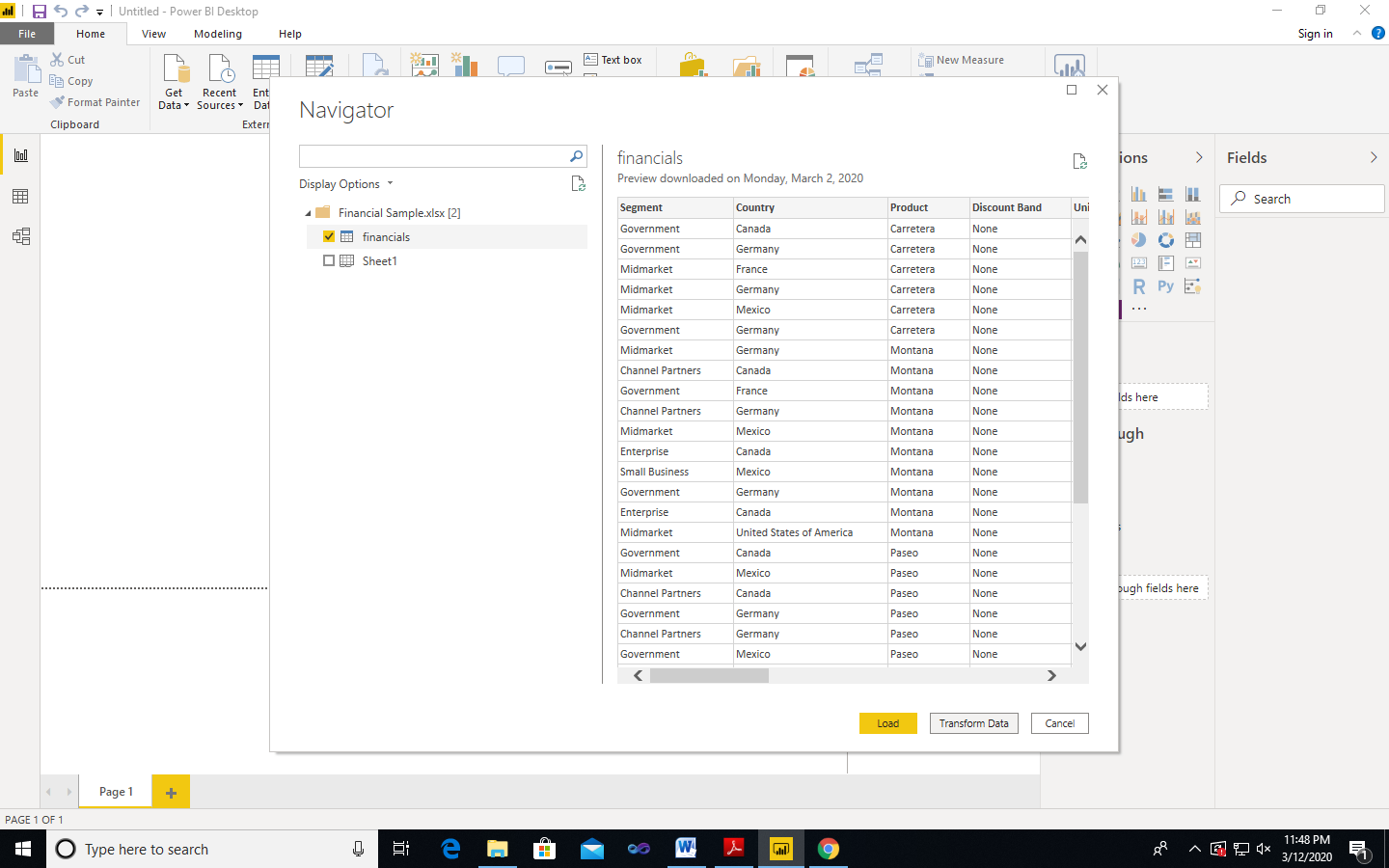
**Practical No. 1**

**Import the legacy data from different sources such as (Excel, SqlServer, Oracle etc.) and load in the target system.**

1. Launch Power BI Desktop.
2. From the Home ribbon, select Get Data. Excel is one of the Most Common data connections, so you can select it directly from the Get Data menu.



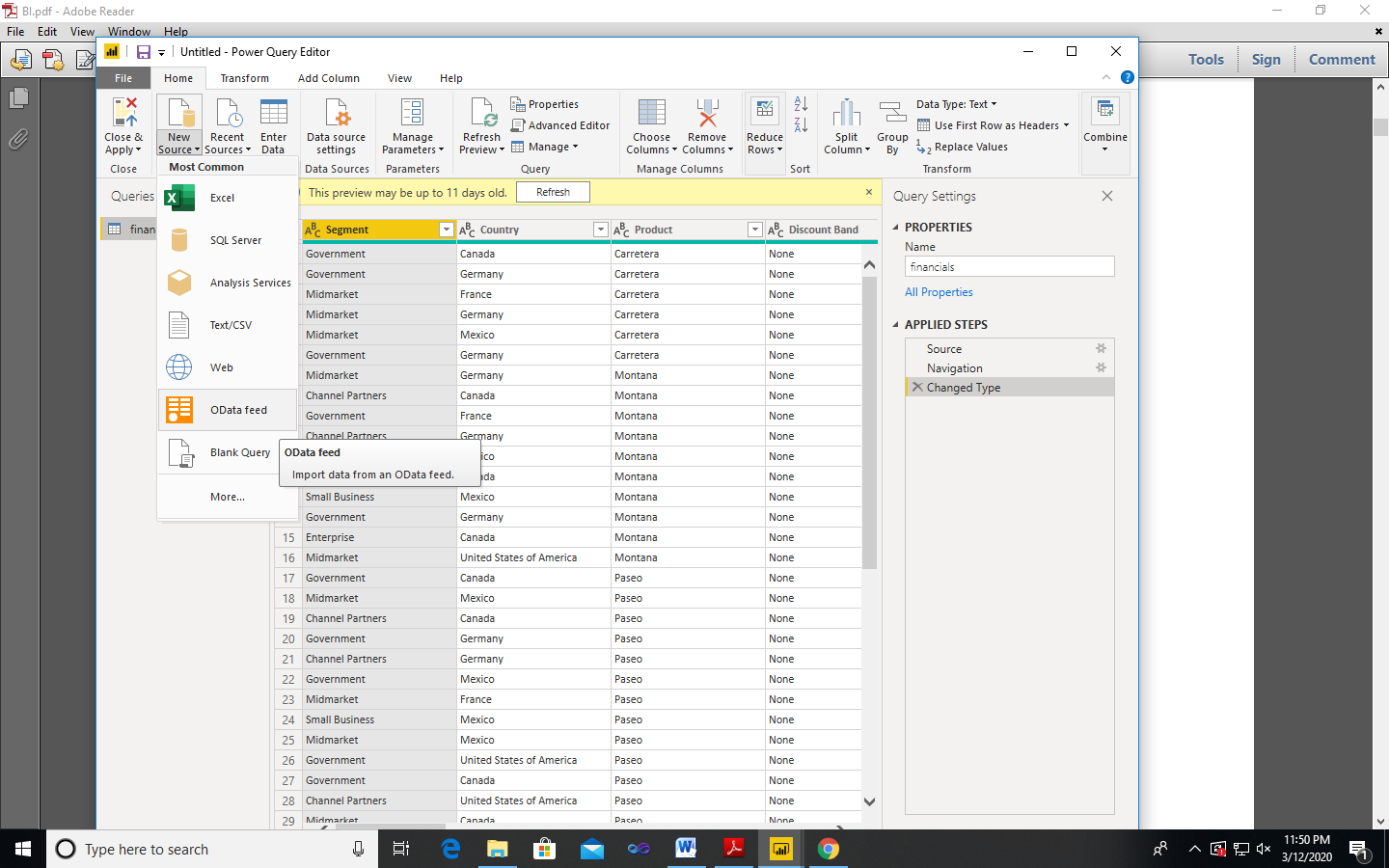
1. If you select the Get Data button directly, you can also select FIle 🡪 Excel and select Connect.
2. In the Open File dialog box, select the Products.xlsx file.
3. In the Navigator pane, select the Products table and then select Edit or Transform Data.



**Importing Data from OData Feed**

Connect to an OData feed

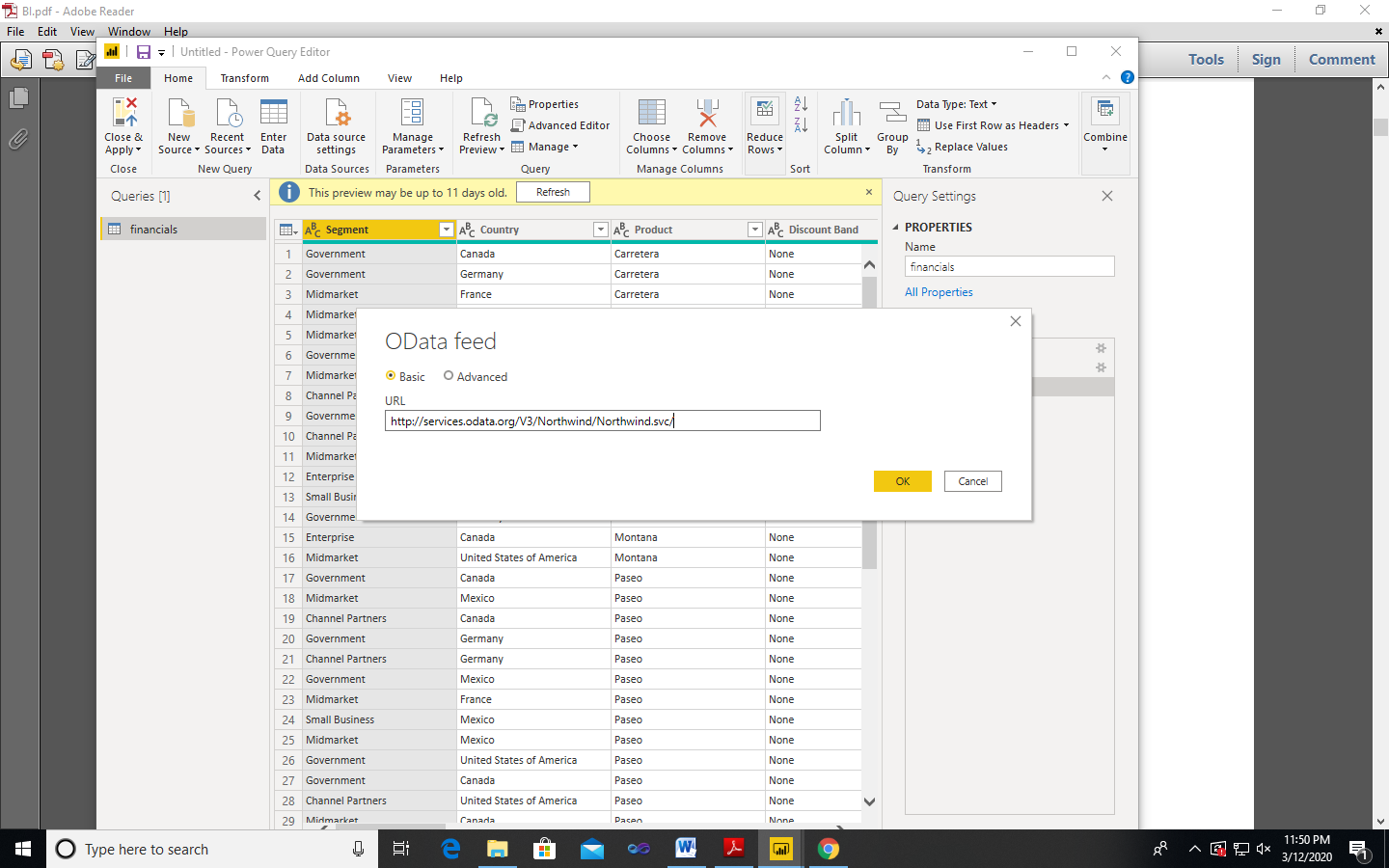
1. From the Home ribbon tab in Query Editor, select Get Data



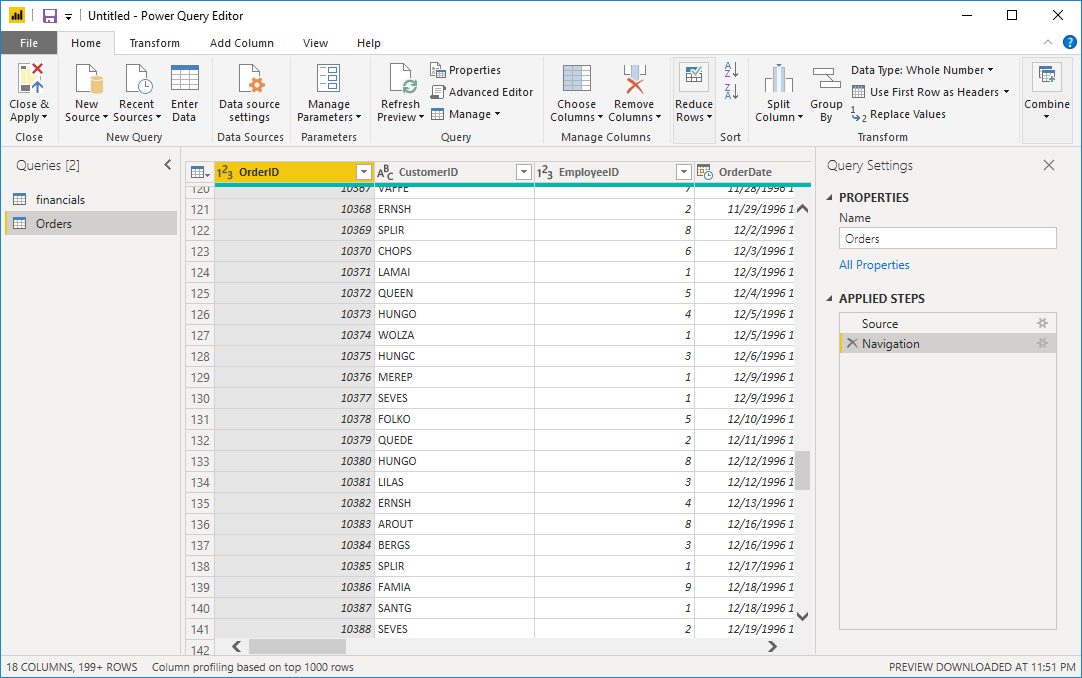
1. Browse to the OData Feed data source.
2. In the OData Feed dialog box, paste the URL for the Northwind OData feed.

<http://services.odata.org/V3/Northwind/Northwind.svc/>

1. Select OK.



1. In the Navigator pane, select the Orders table, and then select Edit.



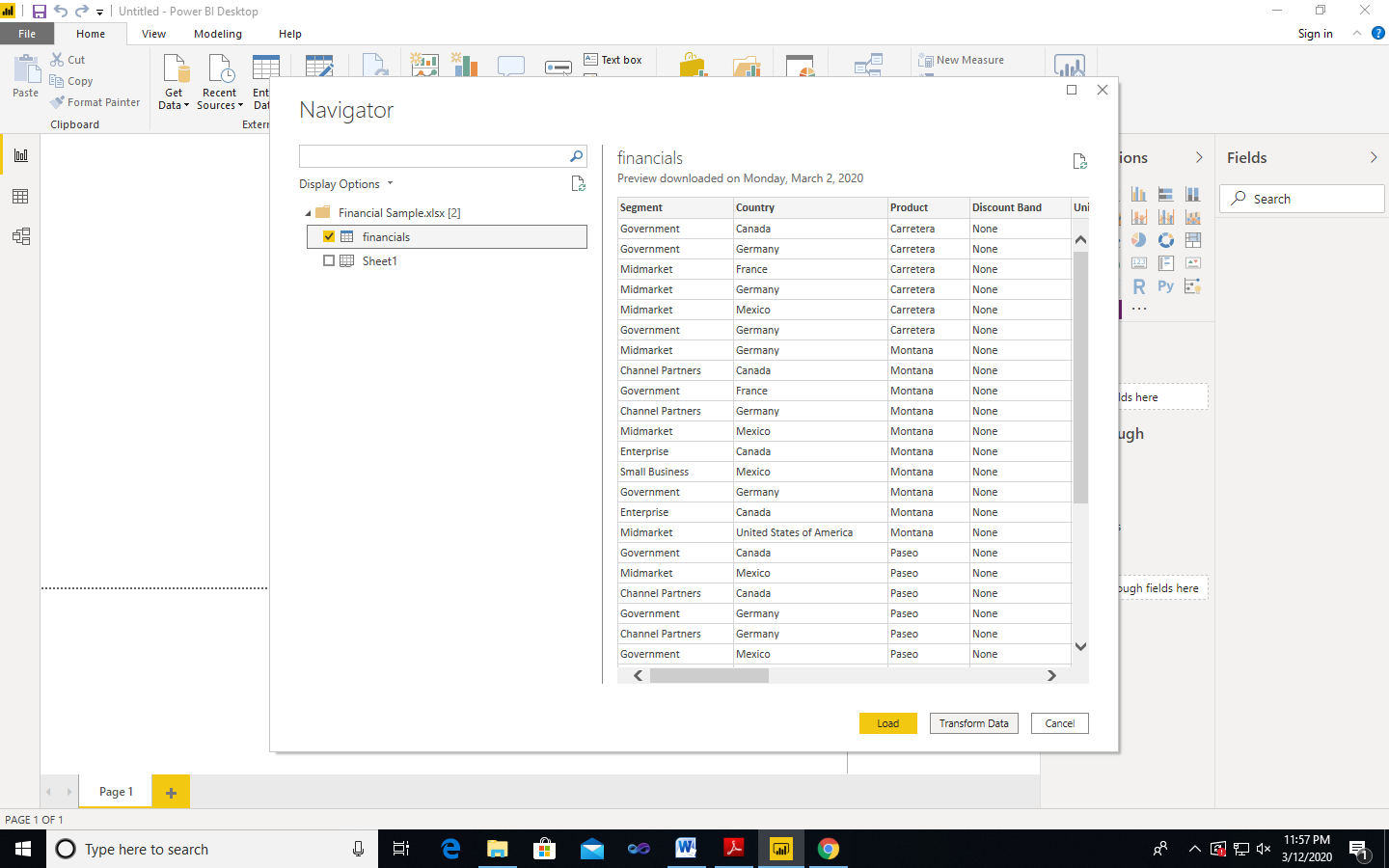
**Practical No. 2**

**Perform the Extraction Transformation and Loading (ETL) process to construct the database in the Power BI.**

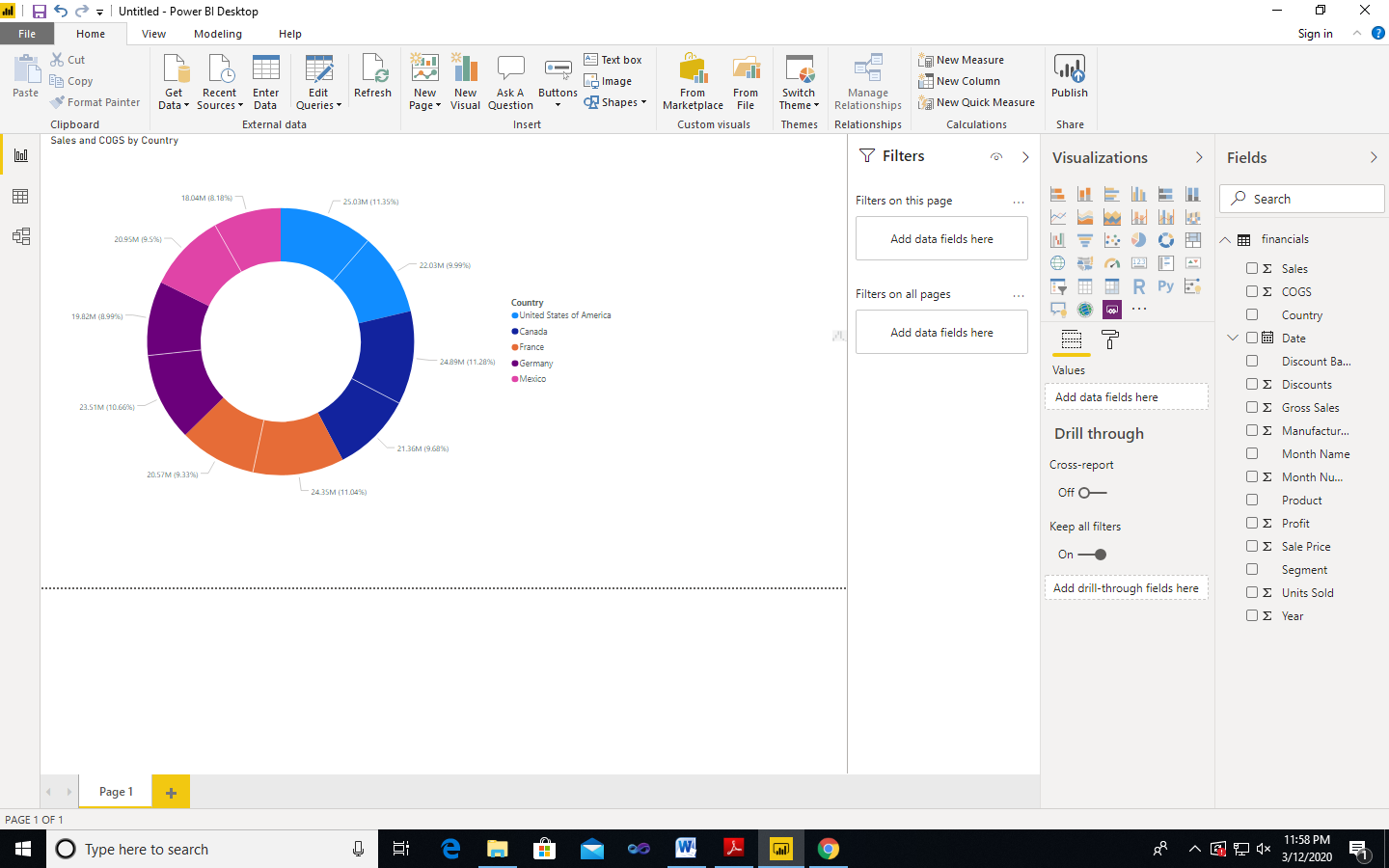
1. From the Home ribbon tab in Query Editor, select Get Data
2. Browse to the OData Feed data source.
3. In the OData Feed dialog box, paste the URL for the Northwind OData feed.

<http://services.odata.org/V3/Northwind/Northwind.svc/>

1. Select OK.
2. In the Navigator pane, select the Orders table, and then select Edit or transform data.
3. Close & Apply



1. In Visualization, select

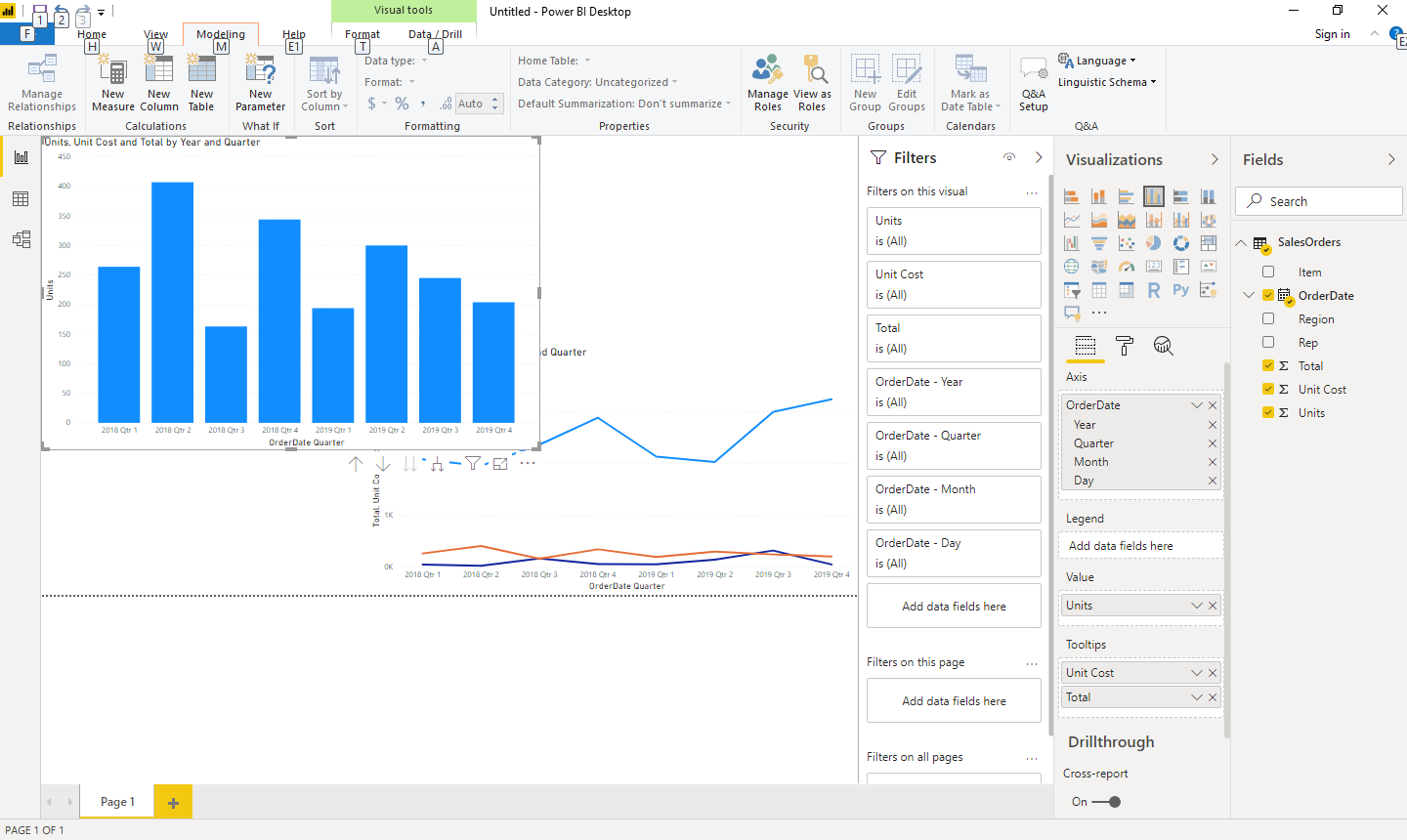


**Practical No.3**

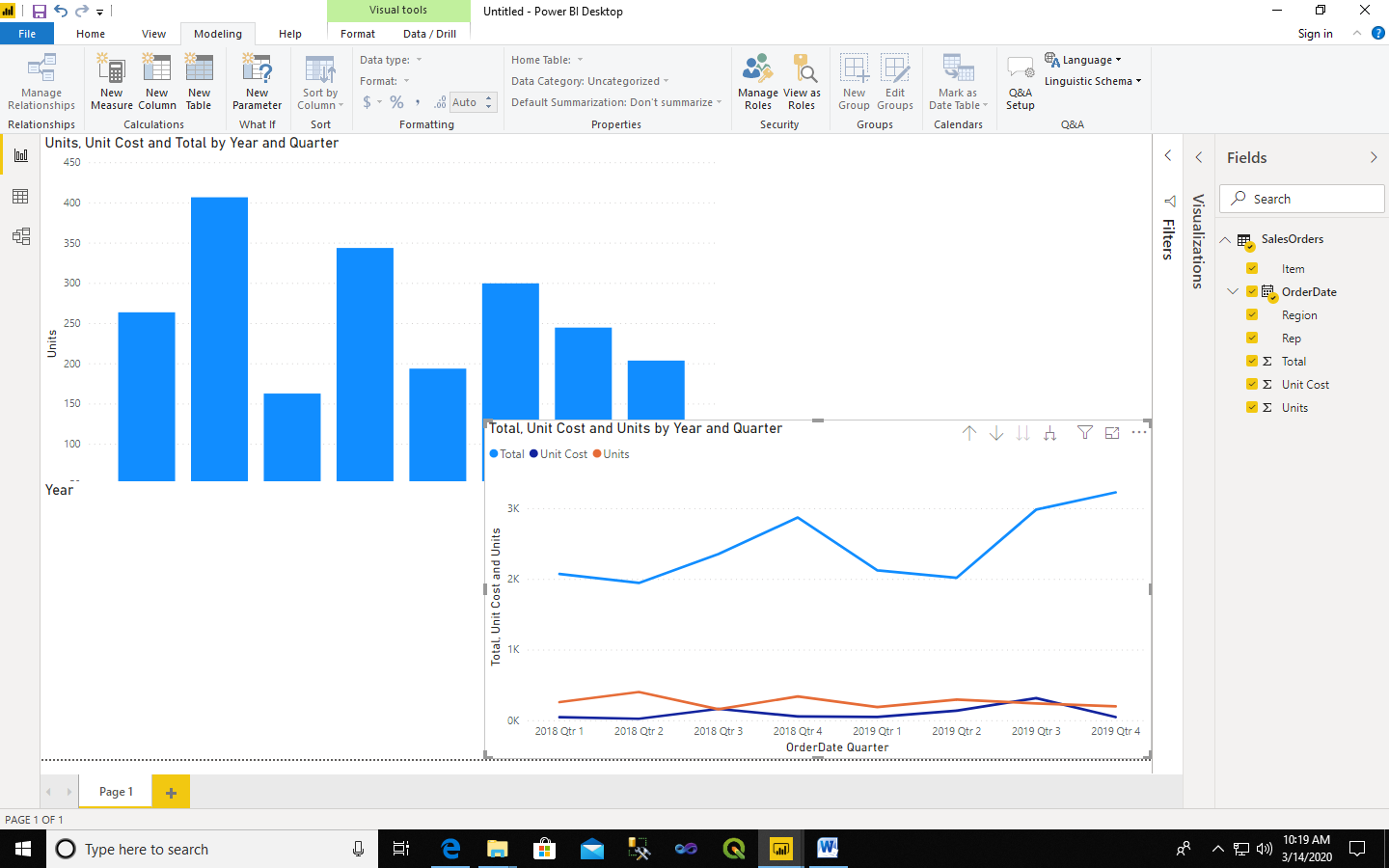
**Data visualization from ETL Process.**

**Step 1:** Create charts showing Units in Stock by Product and Total Sales by Year.

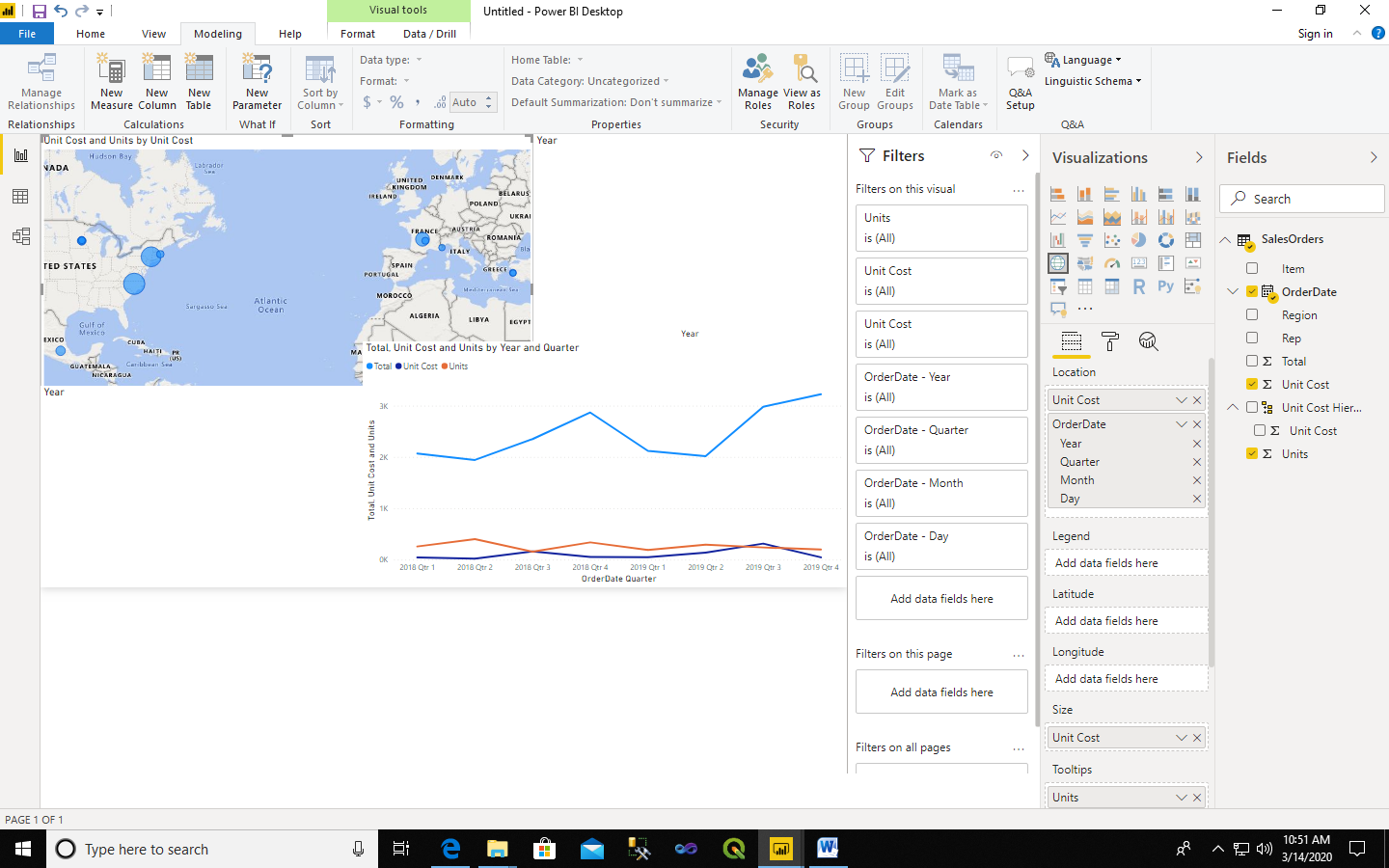
1. Drag **UnitsInStock** from the Field pane (the Fields pane is along the right of the screen) onto a blank space on the canvas. Table visualization is created.
2. Drag Product Name to the Axis box, found in the bottom half of the Visualizations pane. Then we then select Sort By > **UnitsInStock** using the skittles in the top right corner of the visualization.

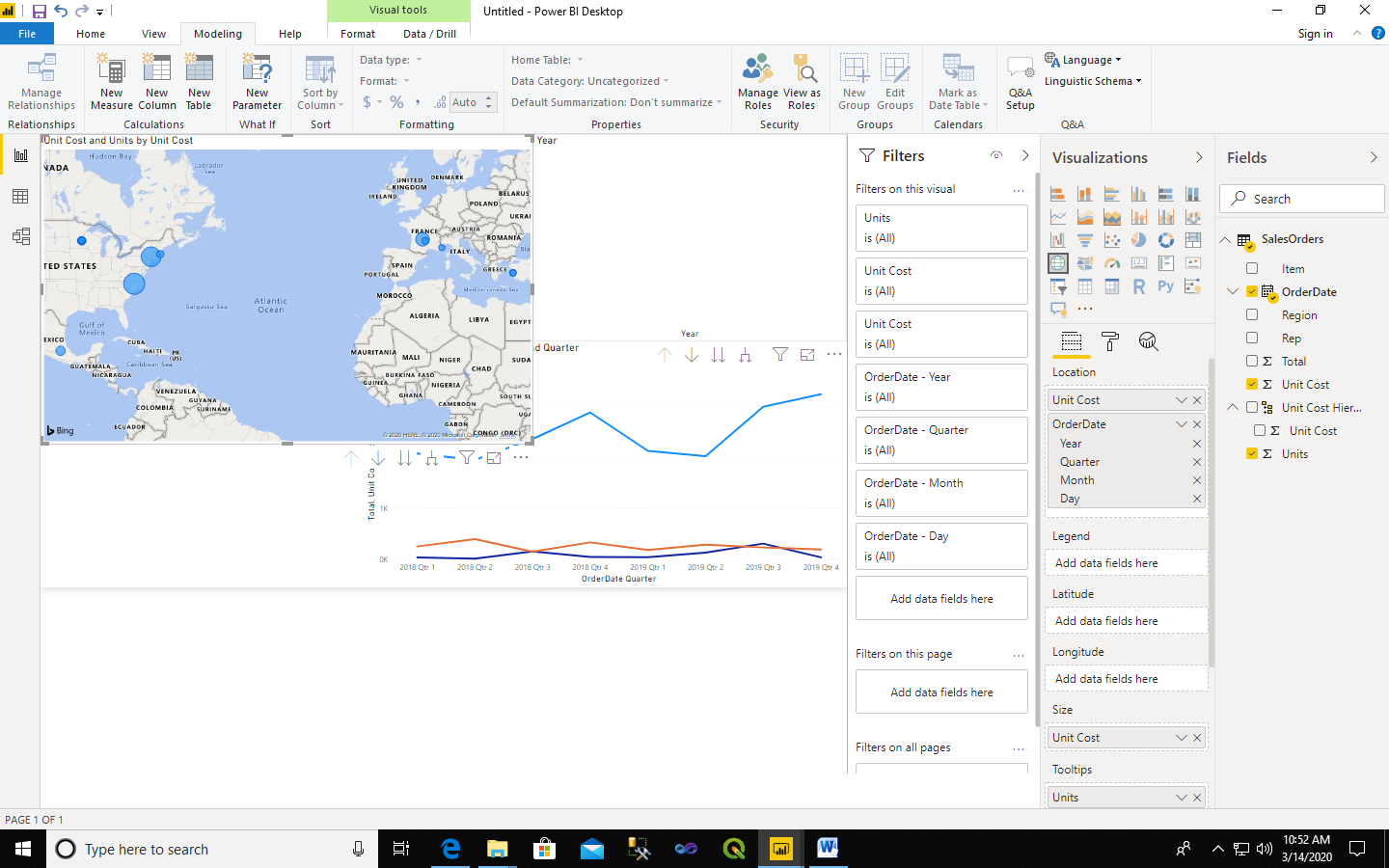


1. Drag **OrderDate** to the canvas beneath the first chart, then drag **LineTotal** (again, from the Fields pane) onto the visual, then select **Line Chart**. The following visualization is created.



4) Drag **orderdate** to a space on the canvas in the top right. Because you selected a geographic field, a map was created automatically. Now drag **LineTotal** to the Values field; the circles on the map for each country are now relative in size to the **LineTotal** for orders shipped to that country.

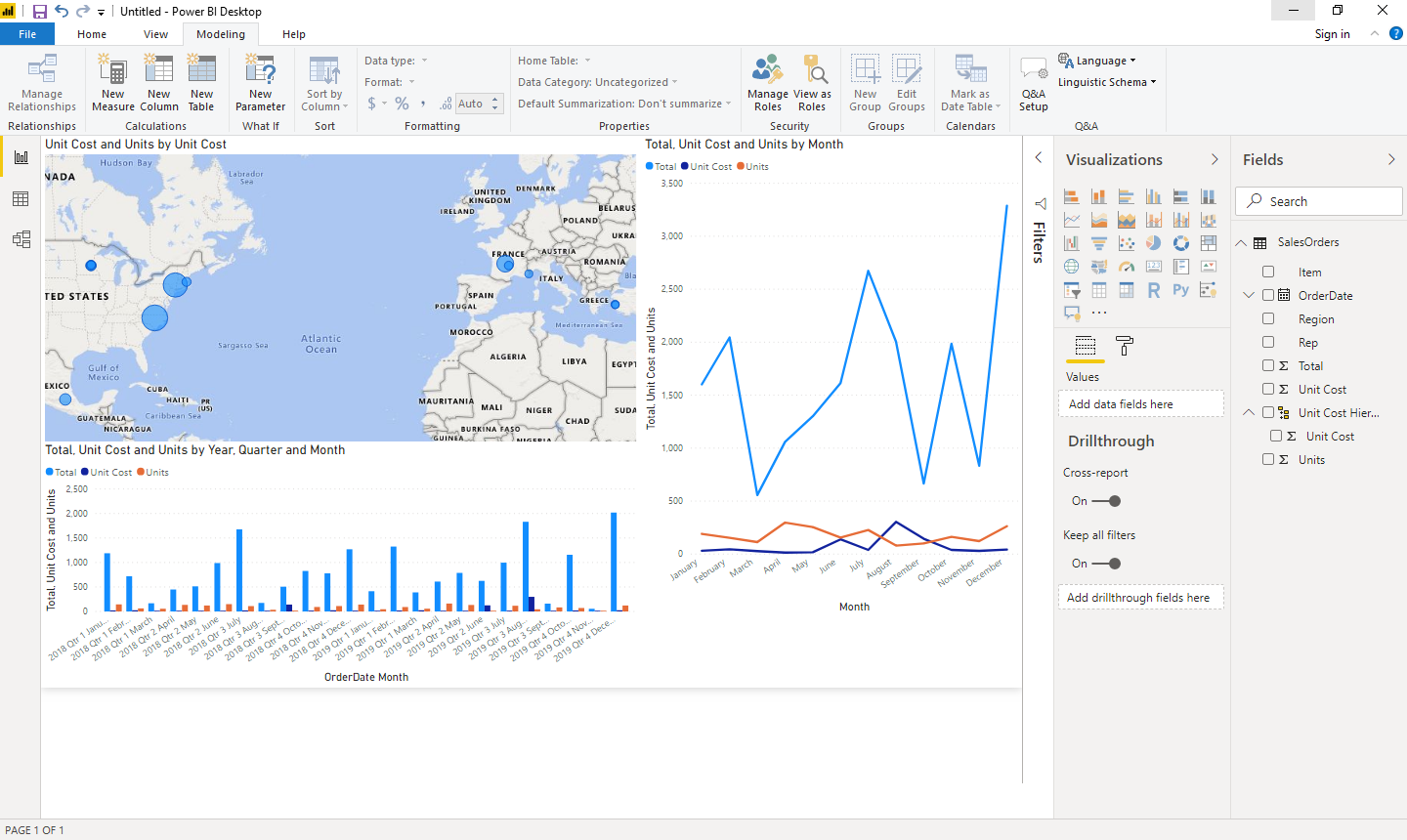




**Step 2:** Interact with your report visuals to analyze further.

Power BI Desktop lets you interact with visuals that cross-highlight and filter each other to uncover further trends.

1. Click on the light blue circle centered in Canada. Note how the other visuals are filtered to show Stock (ShipCountry) and Total Orders (LineTotal) just for Canada.



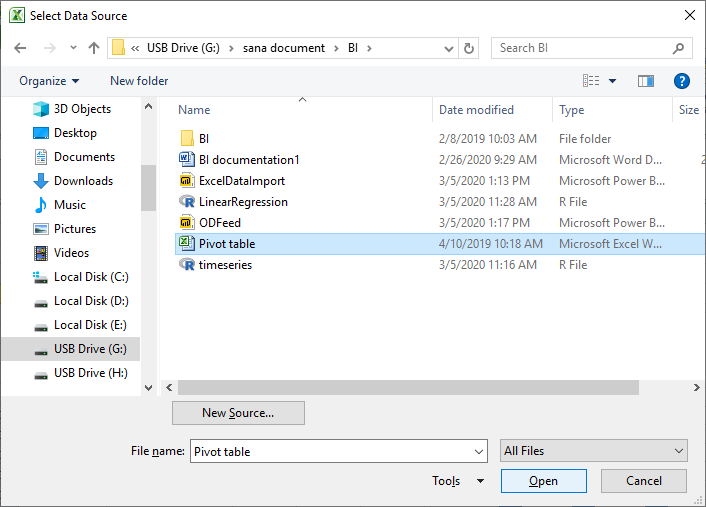
**Practical No. 4**

**Data Modeling and Analytics with Pivot Table and Pivot Chart in Excel.**

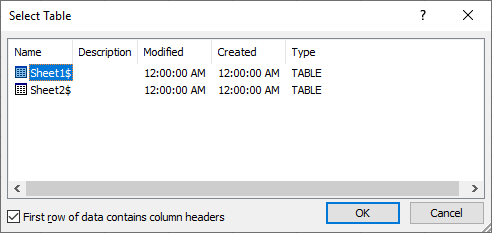
**Step 1** − Open a new blank Workbook in Excel.

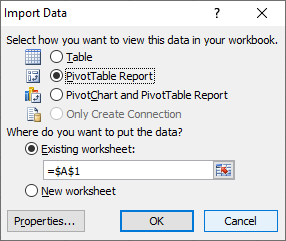
**Step 2** − Click on the **DATA** tab.

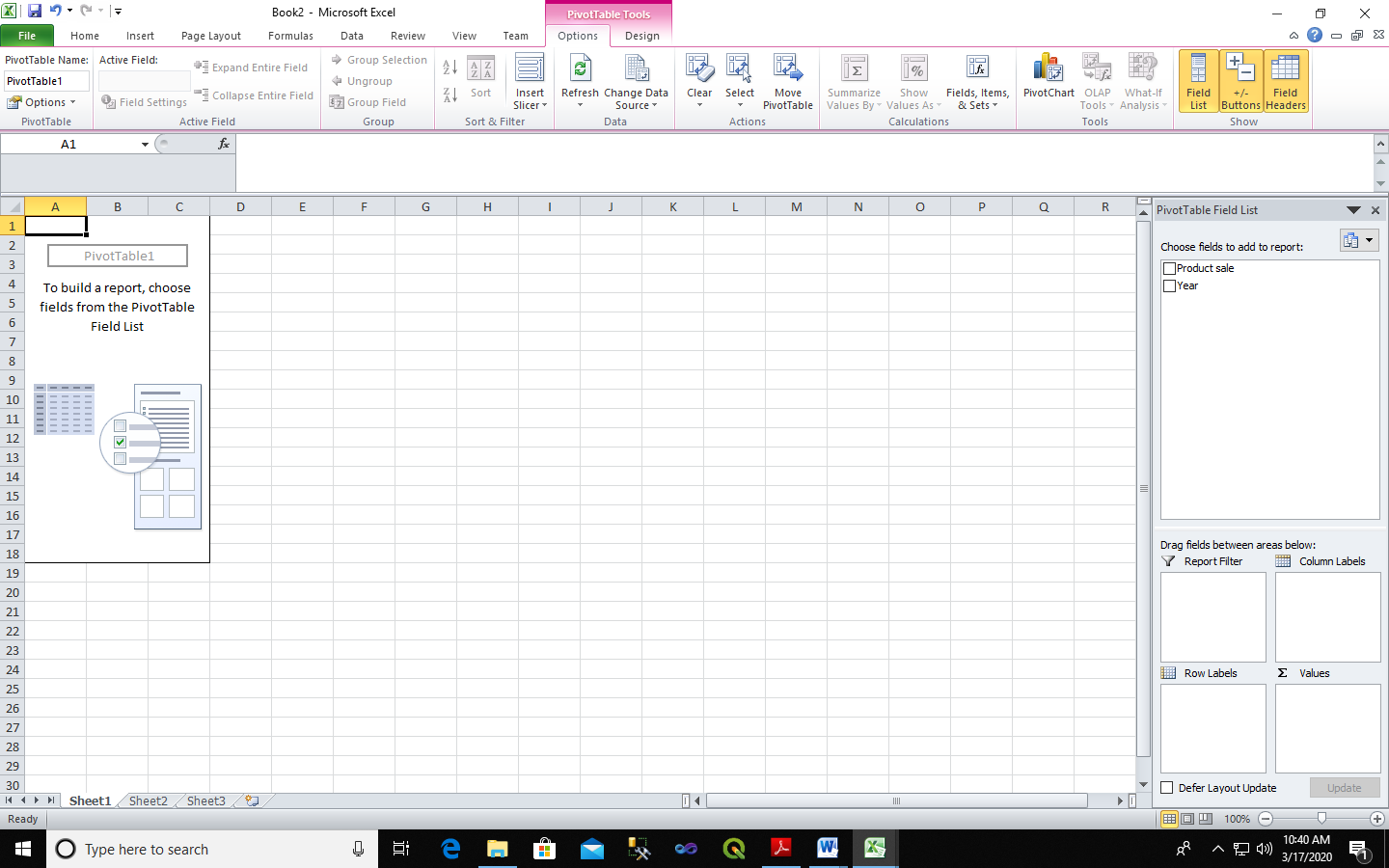
**Step 3** − In the **Get External Data** group, click on the option **From Access**. The **Select Data Source** dialog box opens.



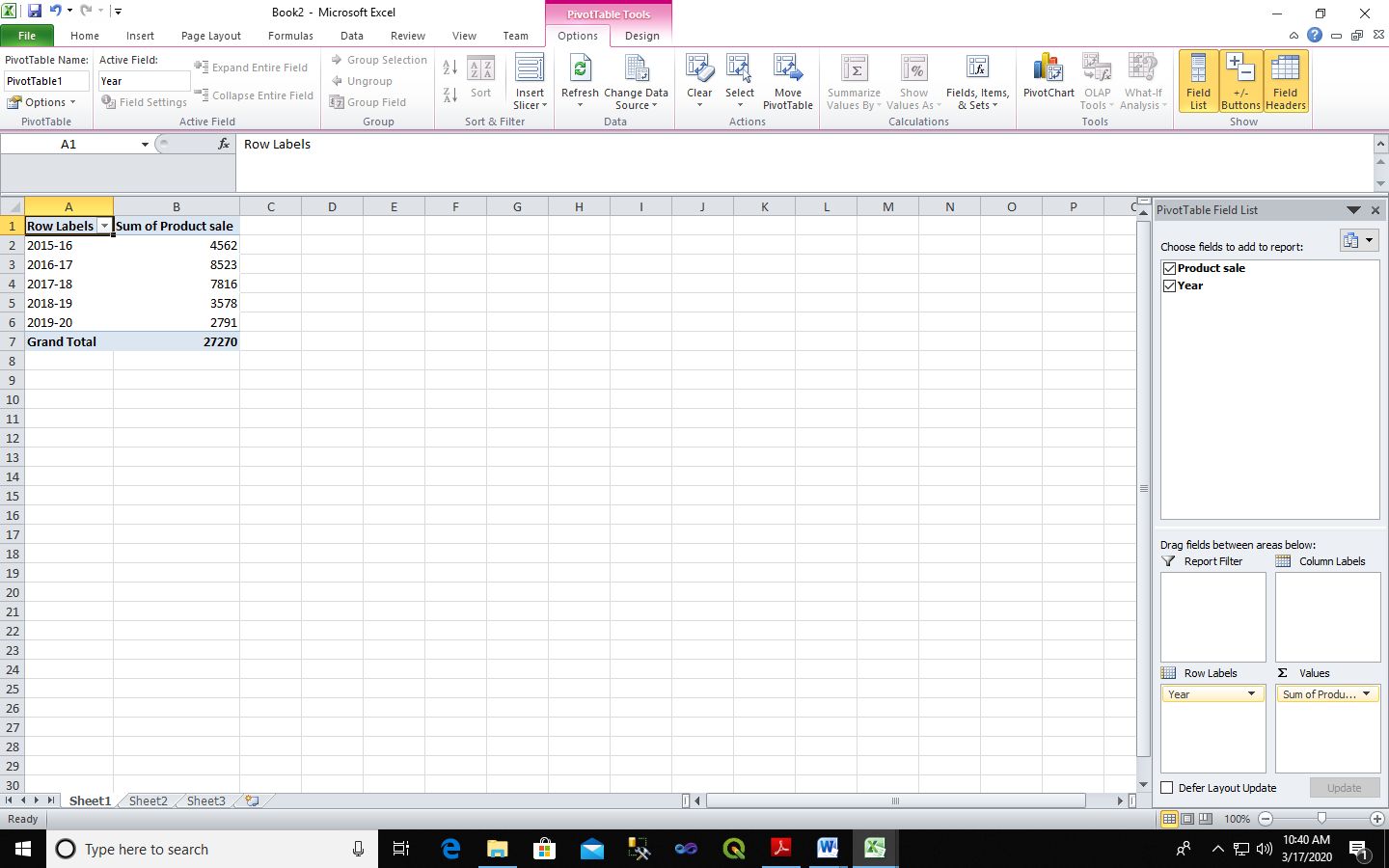
**Step 4** –Choose sheet number 1and click ok.



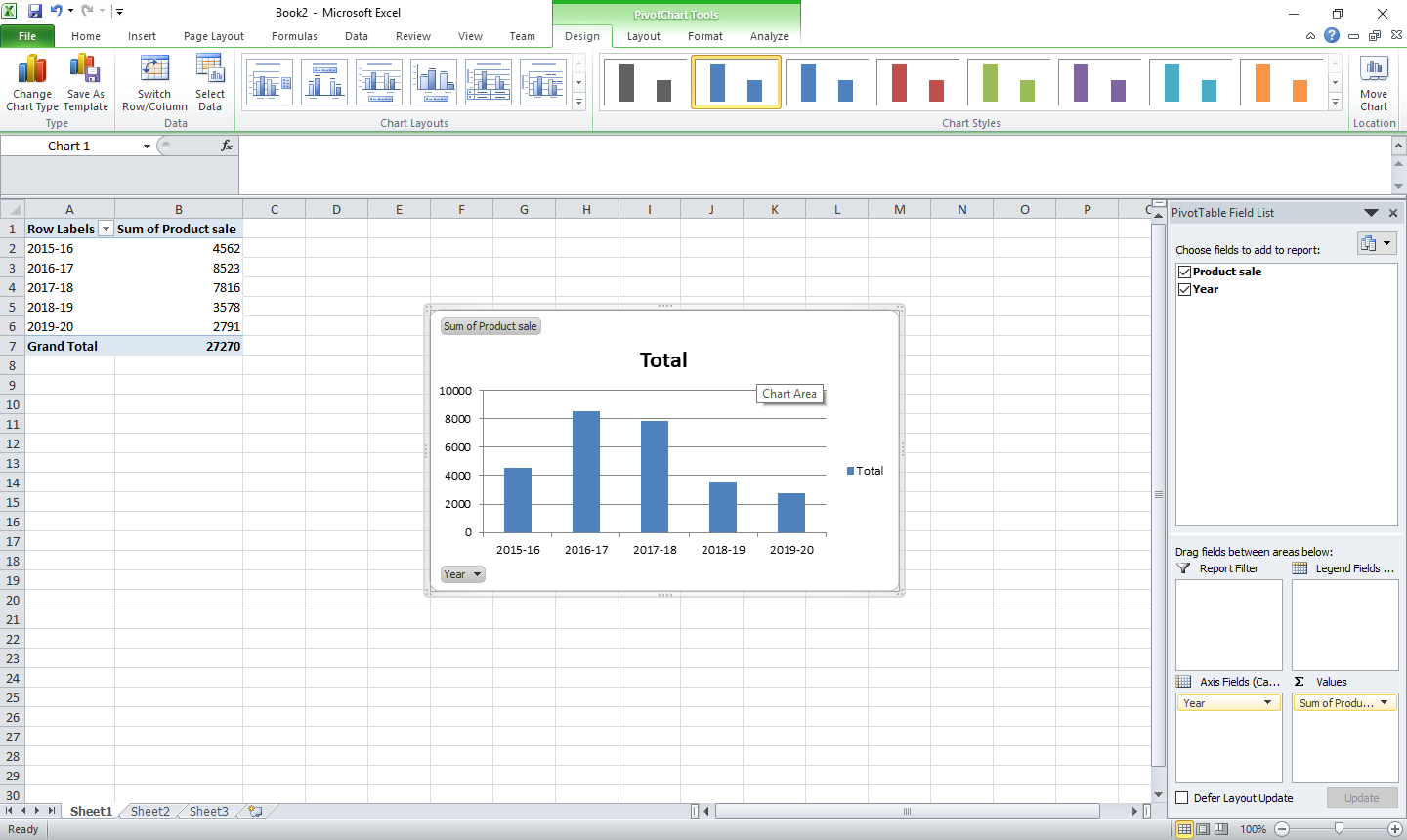




**Pivot Table:**



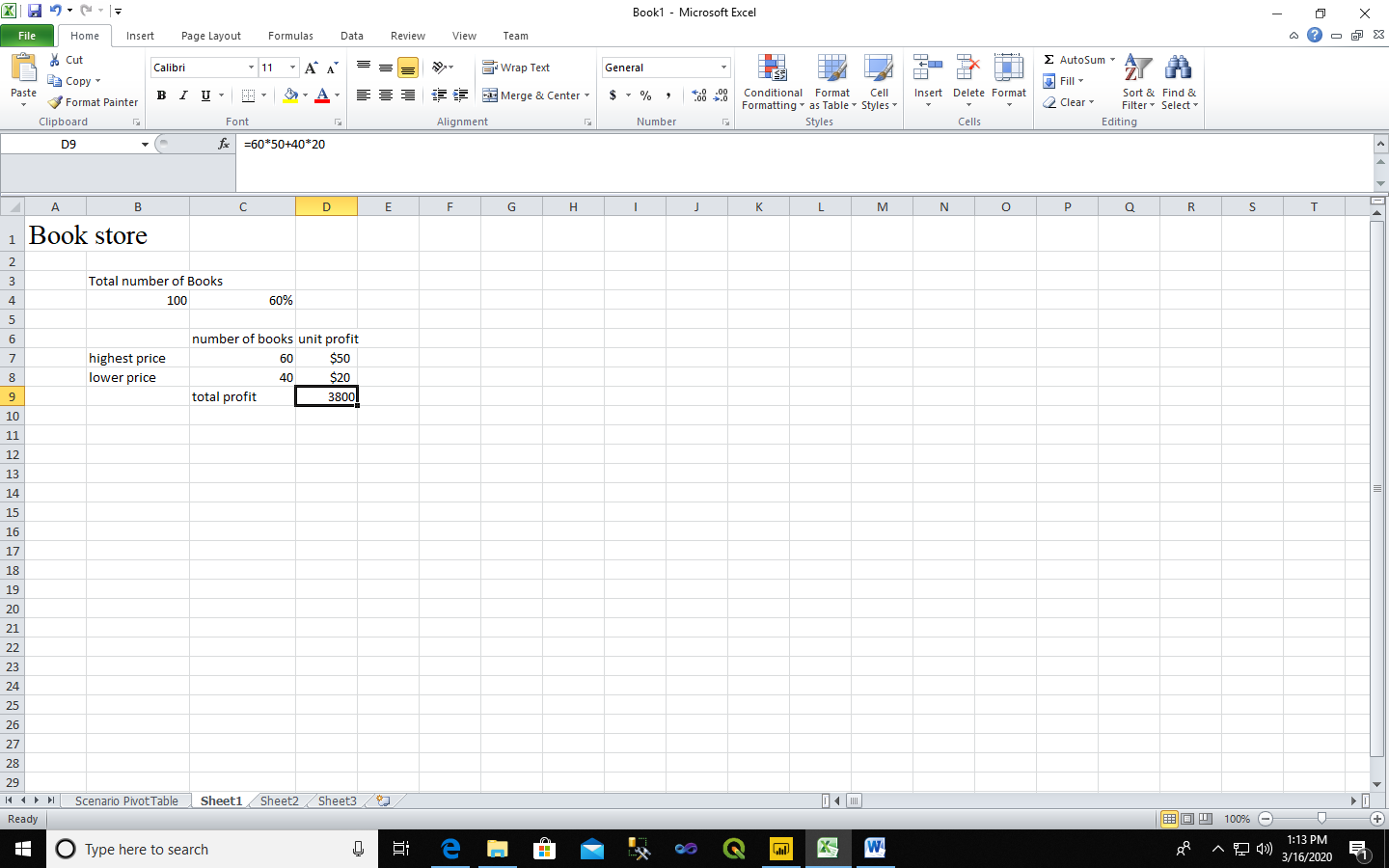
**Pivot Chart:**



**Practical No.5**

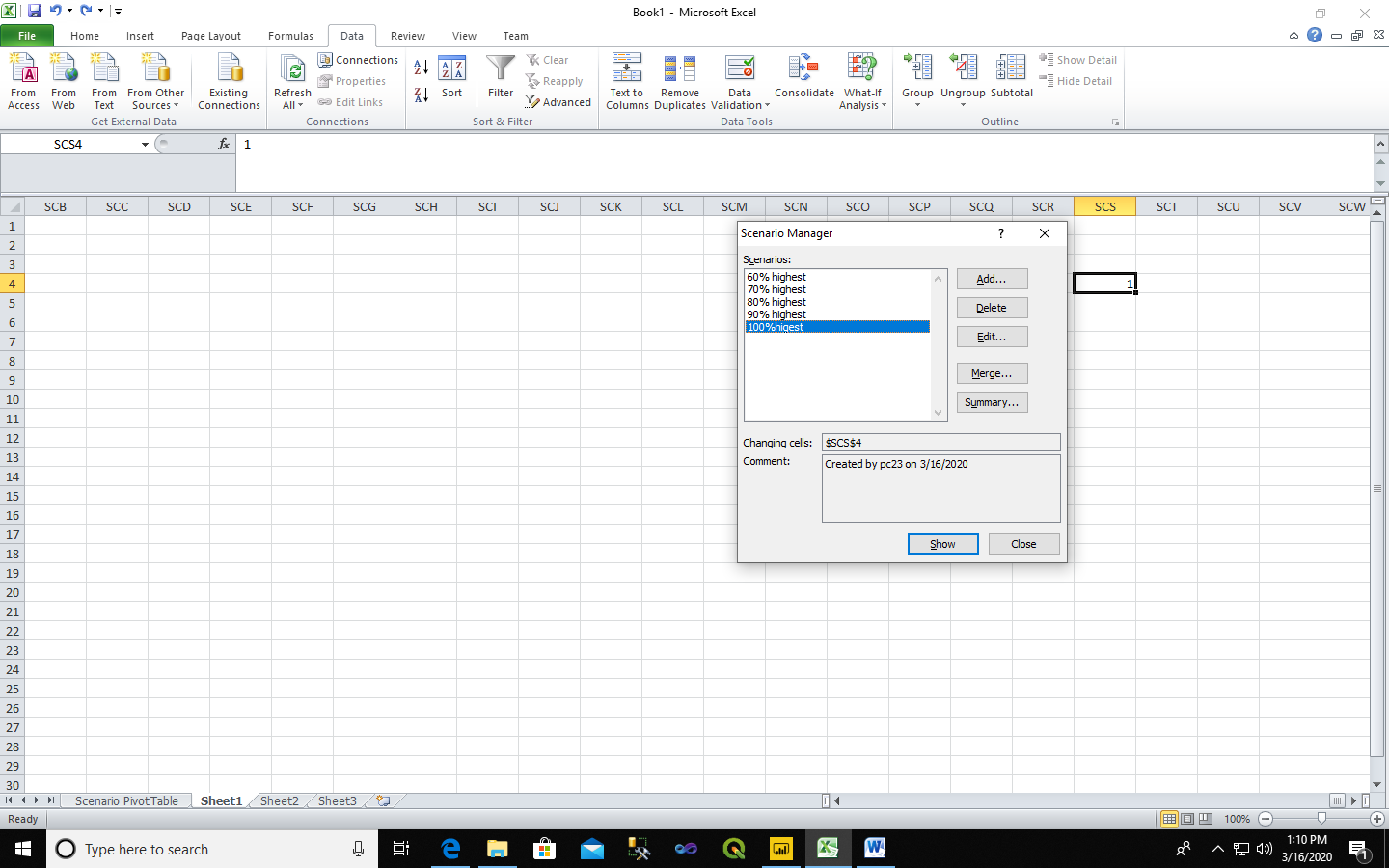
**Apply the what – if Analysis for data visualization. Design and generate necessary reports based on the data warehouse data.**

A book store and have 100 books in storage. You sell a certain % for the highest price of $50 and a certain % for the lower price of $20.

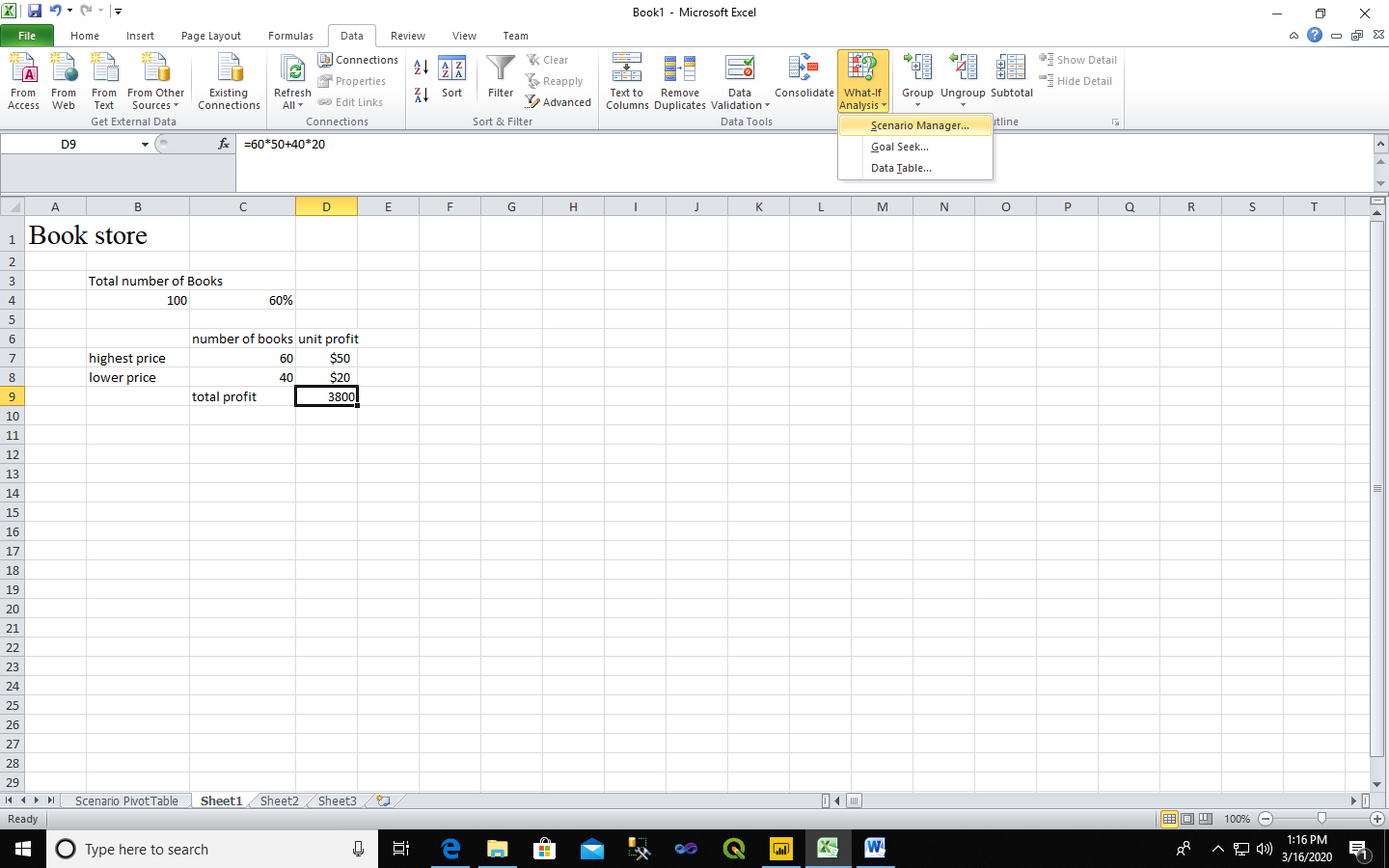


If you sell 60% for the highest price, cell D10 calculates a total profit of 60 \* $50 + 40 \* $20 = $3800. Create Different Scenarios But what if you sell 70% for the highest price? And what if you sell 80% for the highest price? Or 90%, or even 100%? Each different percentage is a different scenario. You can use the Scenario Manager to create these scenarios. Note: You can simply type in a different percentage into cell C4 to see the corresponding result of a scenario in cell D10. However, what-if analysis enables you to easily compare the results of different scenarios. Read on.

1. On the Data tab, in the Forecast group, click What-If Analysis.

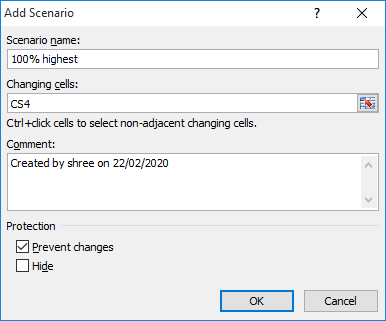


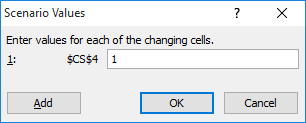
1. Click Scenario Manager.

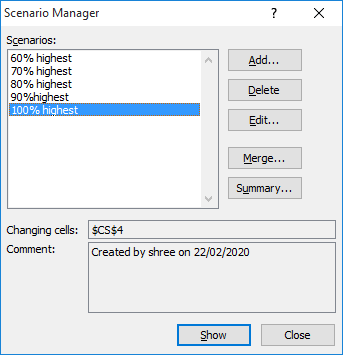


The Scenario Manager dialog box appears.

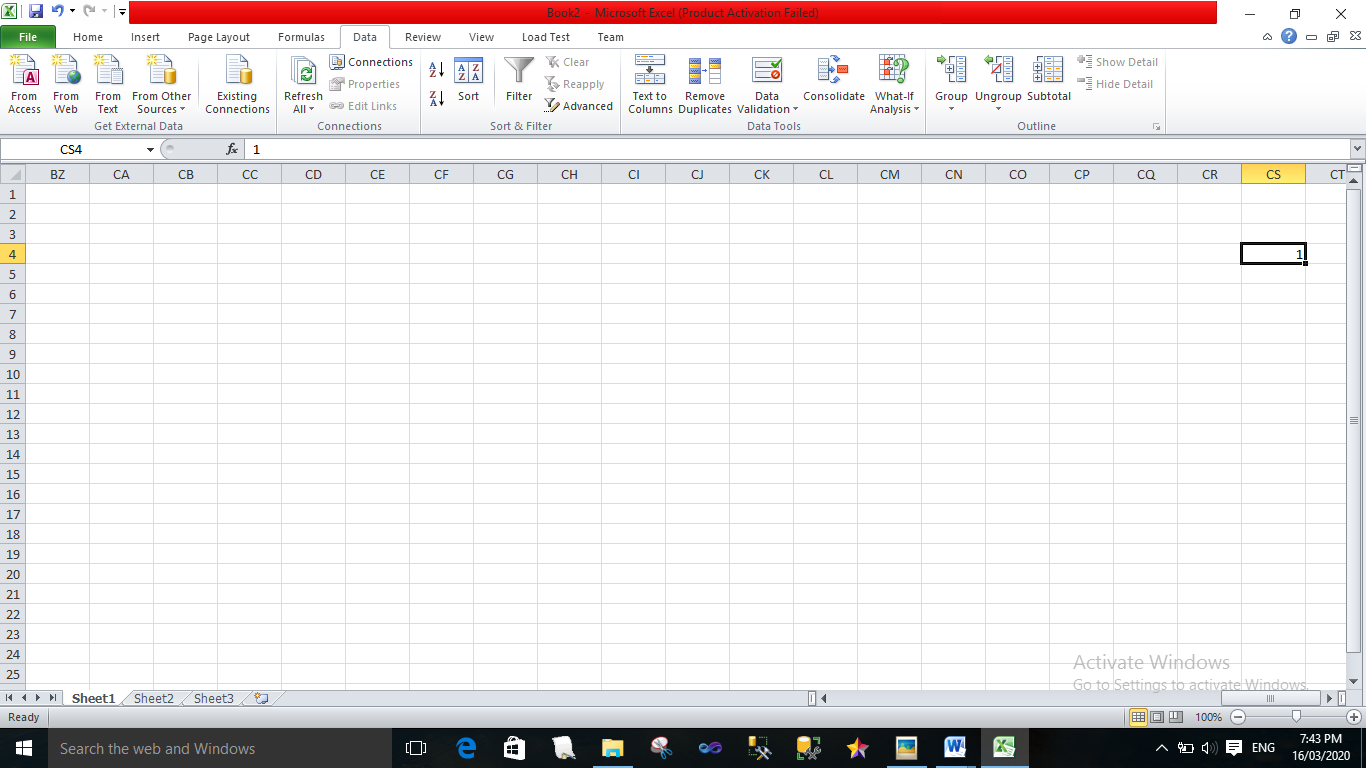
1. Add a scenario by clicking on Add.
2. Type a name (60% highest), select cell C4 (% sold for the highest price) for the Changing cells and click on OK.
3. Enter the corresponding value 0.6 and click on OK again.
4. Next, add 4 other scenarios (70%, 80%, 90% and 100%).

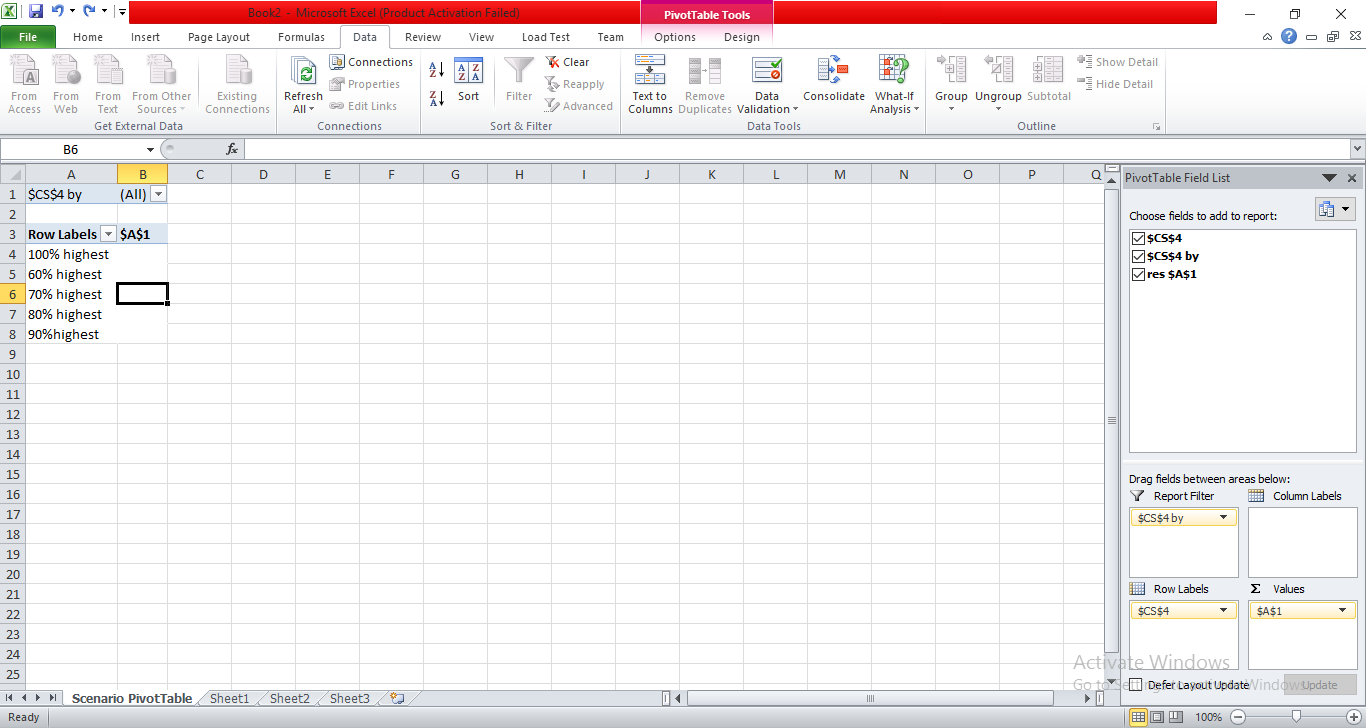






**Output:**





**Practical No.6**

**Implementation of Classification algorithm in R Programming.**

Consider the annual rainfall details at a place starting from January 2012. We create an R time series object for a period of 12 months and plot it.

**Code:**

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)

png(file = "E:rainfall.png")

plot(rainfall.timeseries)

dev.off(0)

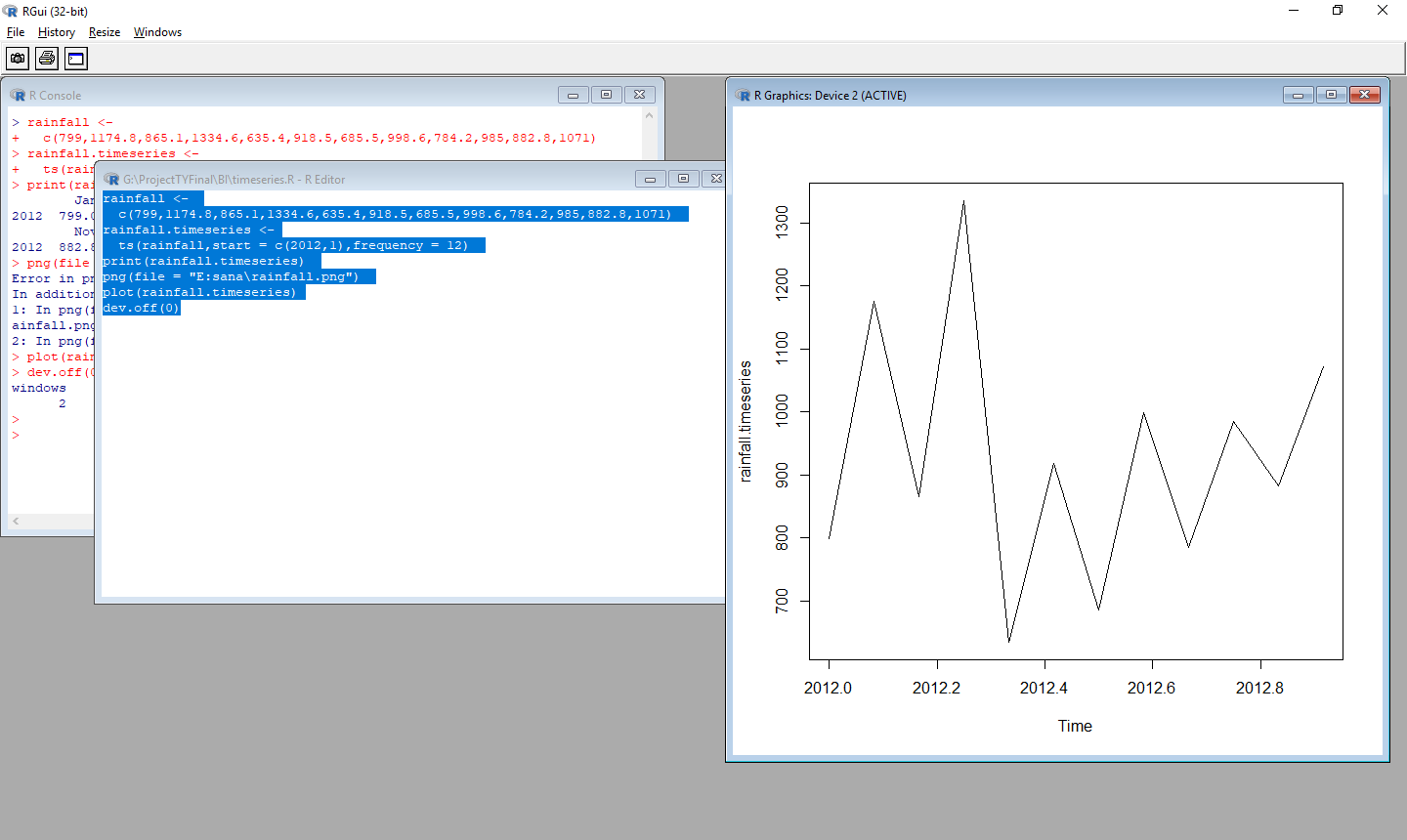
**Output:**

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



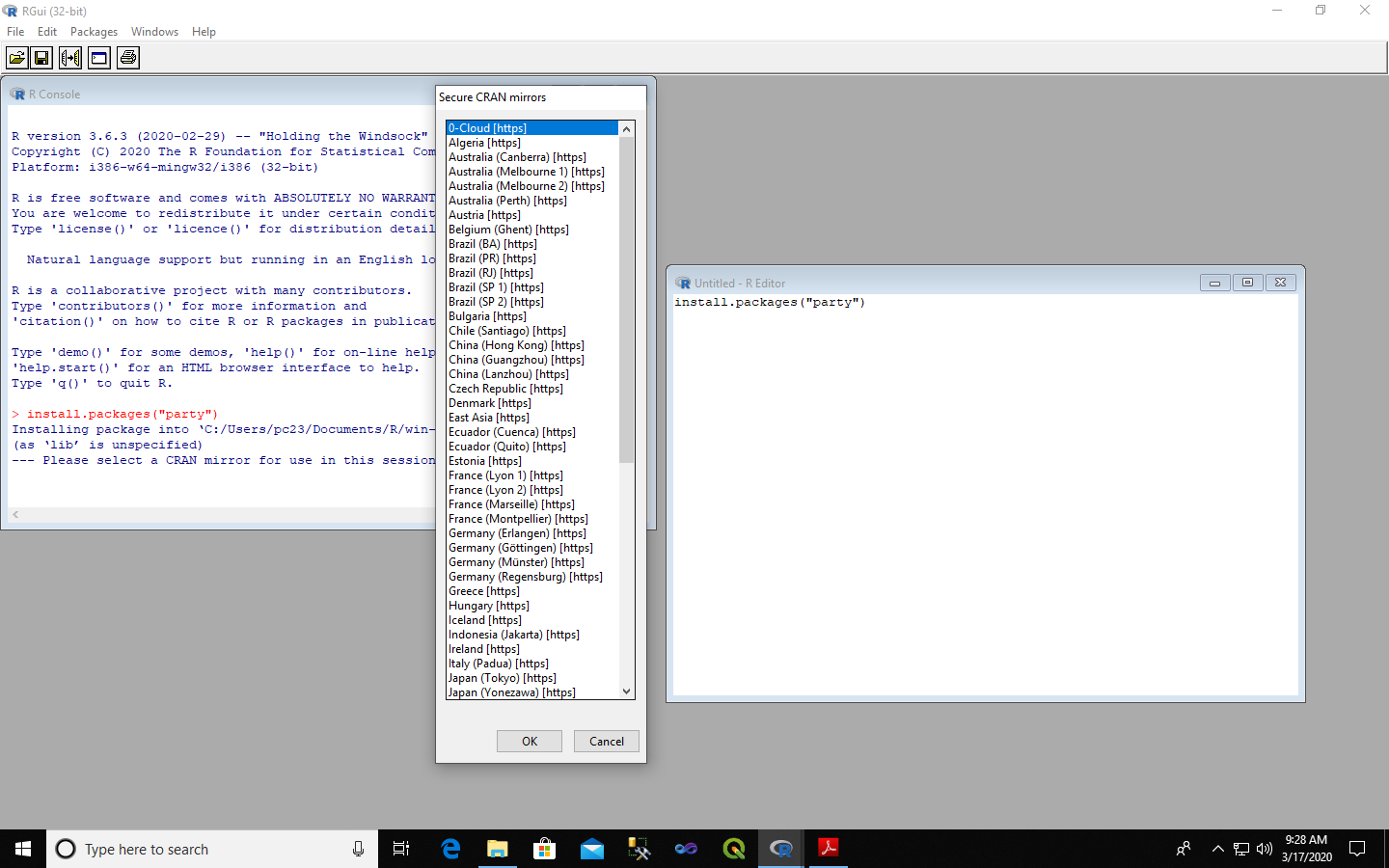
**Practical No.7**

**Practical Implementation of Decision Tree using R Tool.**

The package "party" has the function **ctree()** which is used to create and analyze decison tree.

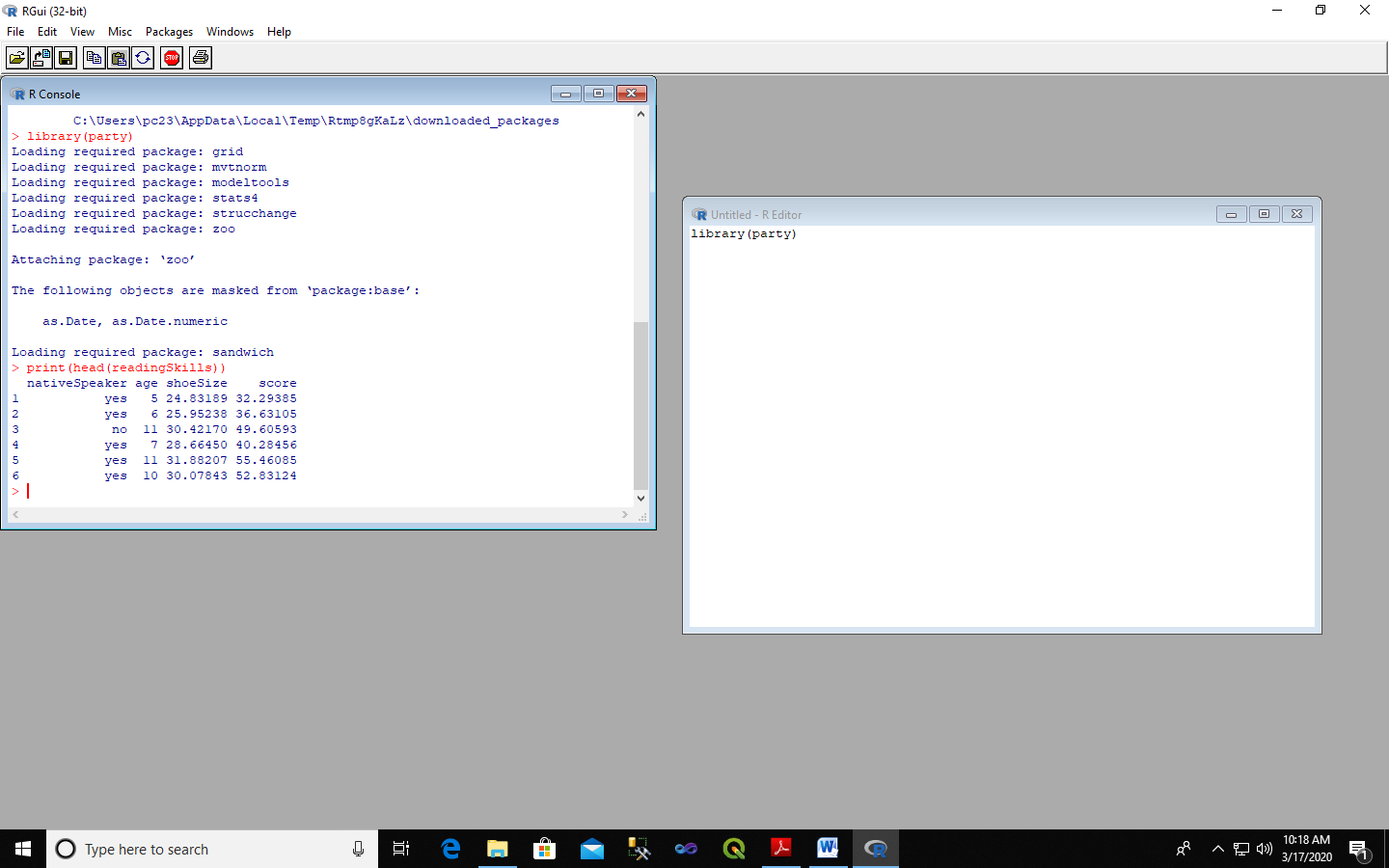
**Code:**

Install.packages(“party”).



library(party)

print(head(readingSkills))



library(party)

input.dat <- readingSkills[c(1:105),]

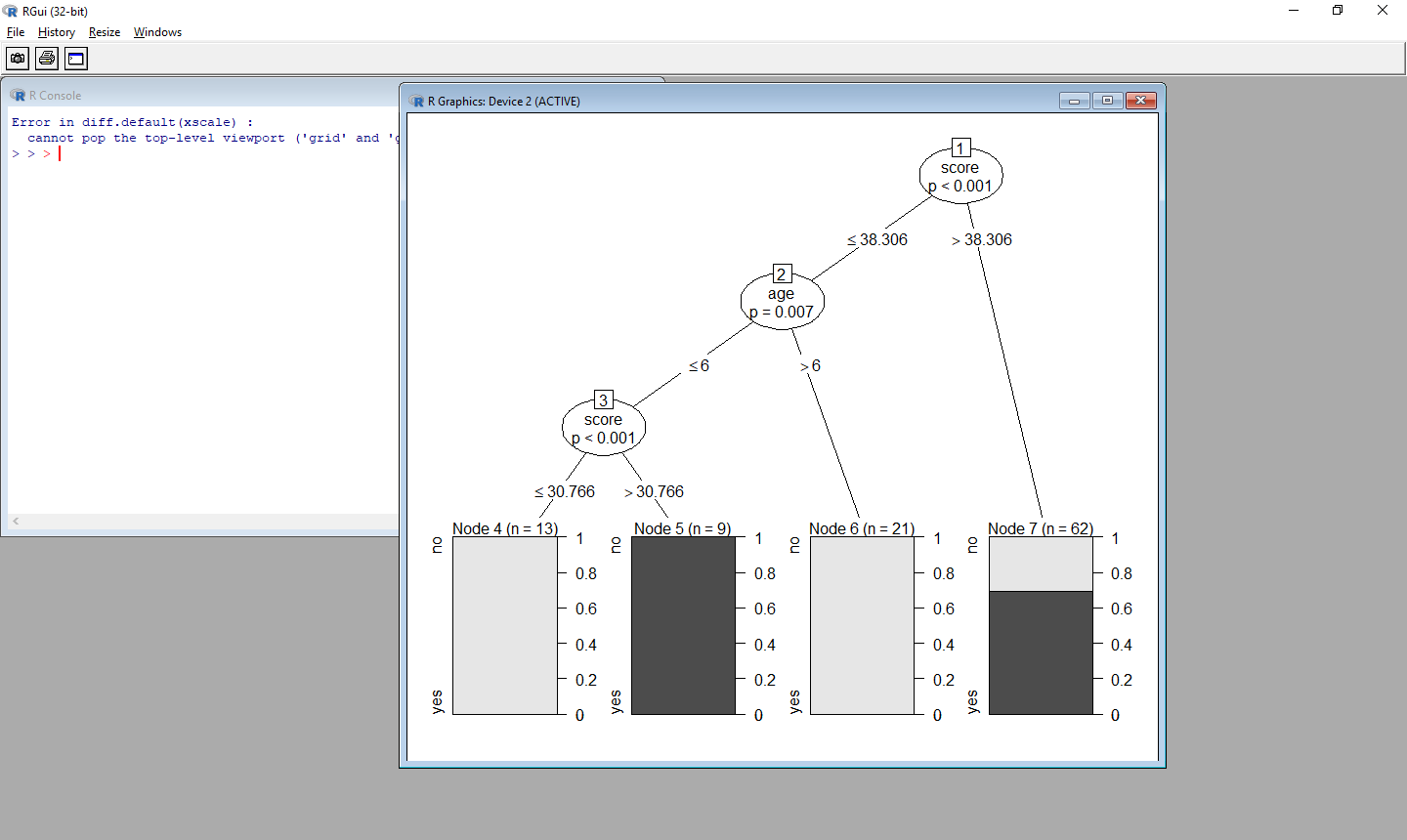
png(file = "E:\decision\_tree.png")

output.tree <- ctree(

nativeSpeaker ~ age + shoeSize + score,data = input.dat)

plot(output.tree)

dev.off(0)



**Practical No. 8**

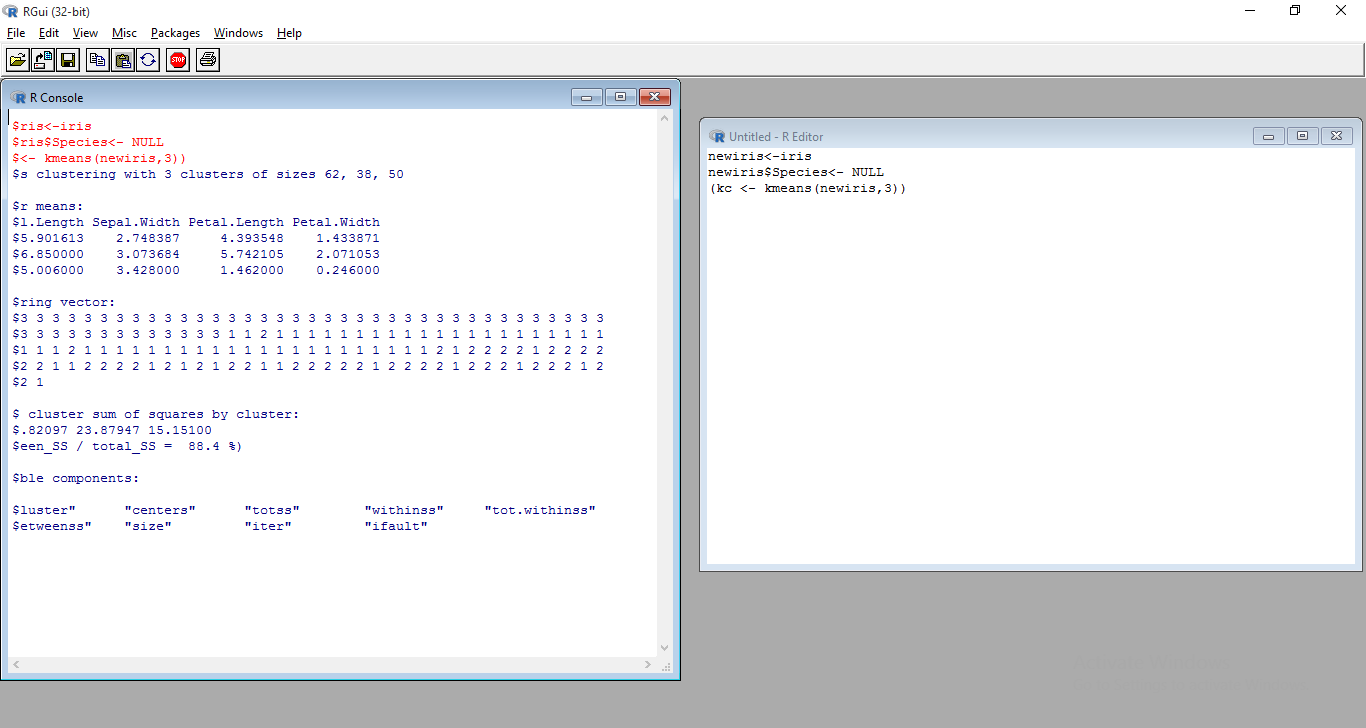
**k-means clustering using R**

**Step 1: Apply K mean to iris and store result.**

newiris<-iris

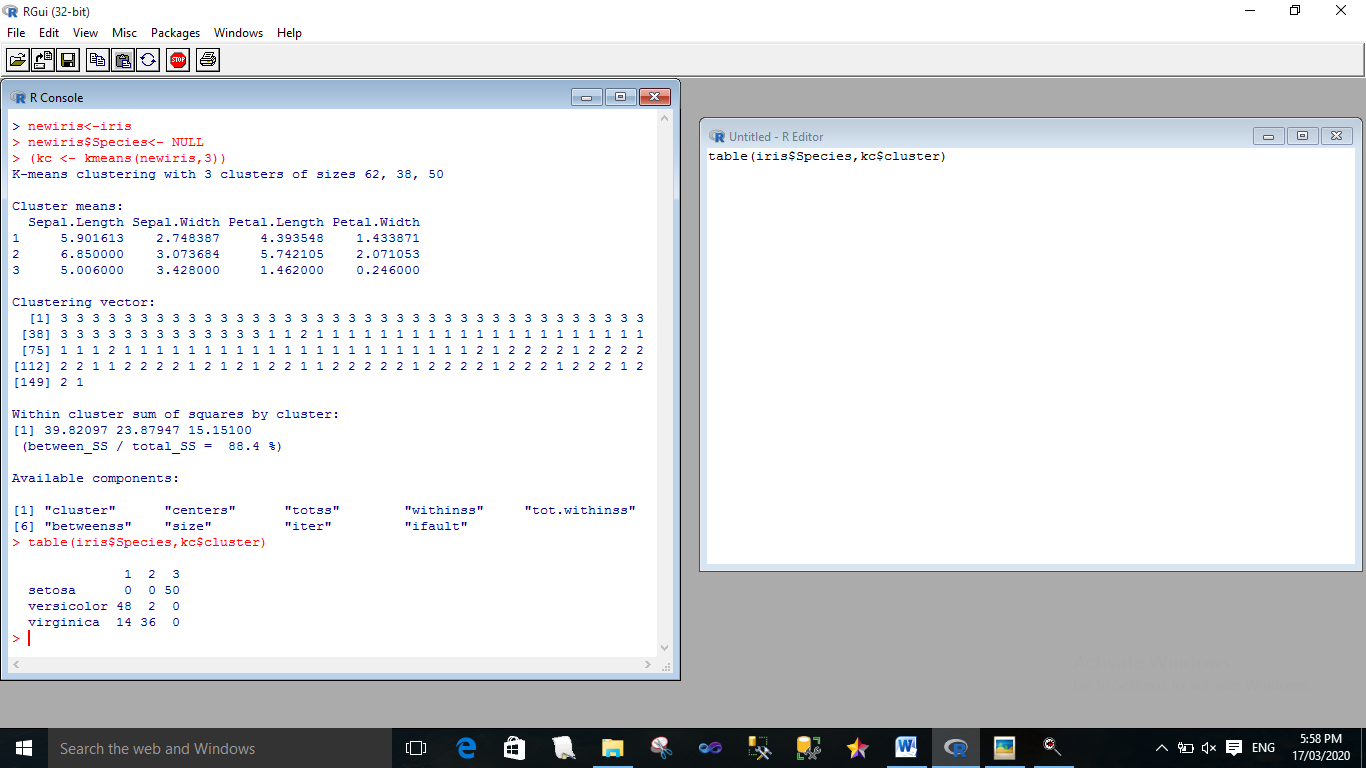
newiris$Species<- NULL

(kc <- kmeans(newiris,3))



**Step 2: Compare the Species label with the clustering result**

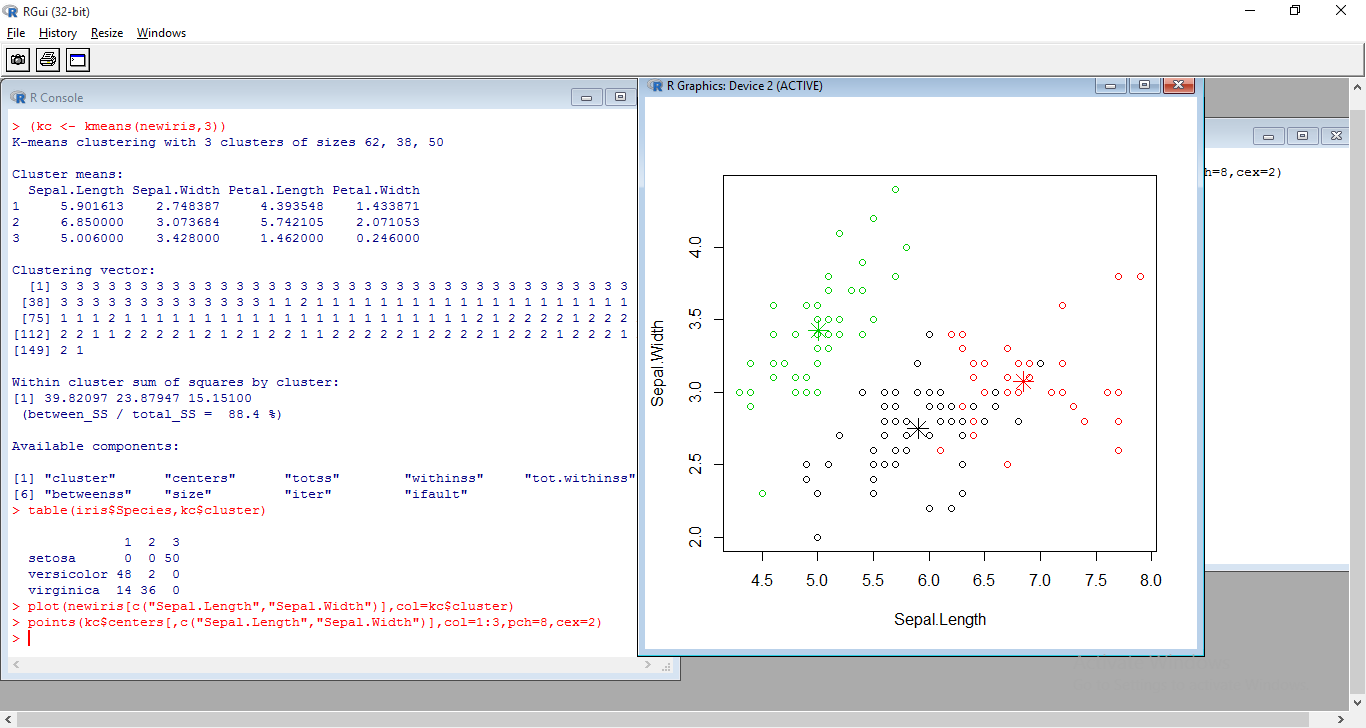
table(iris$Species,kc$cluster)



**Step 3: Plot the clusters and their center**

plot(newiris[c("Sepal.Length","Sepal.Width")],col=kc$cluster)

points(kc$centers[,c("Sepal.Length","Sepal.Width")],col=1:3,pch=8,cex=2)



**Practical No. 9**

**Prediction Using Linear Regression.**

**Code:**

x <- c(151, 174, 138, 186, 128, 136, 179, 163, 152, 131)

y <- c(63, 81, 56, 91, 47, 57, 76, 72, 62, 48)

relation <- lm(y~x)

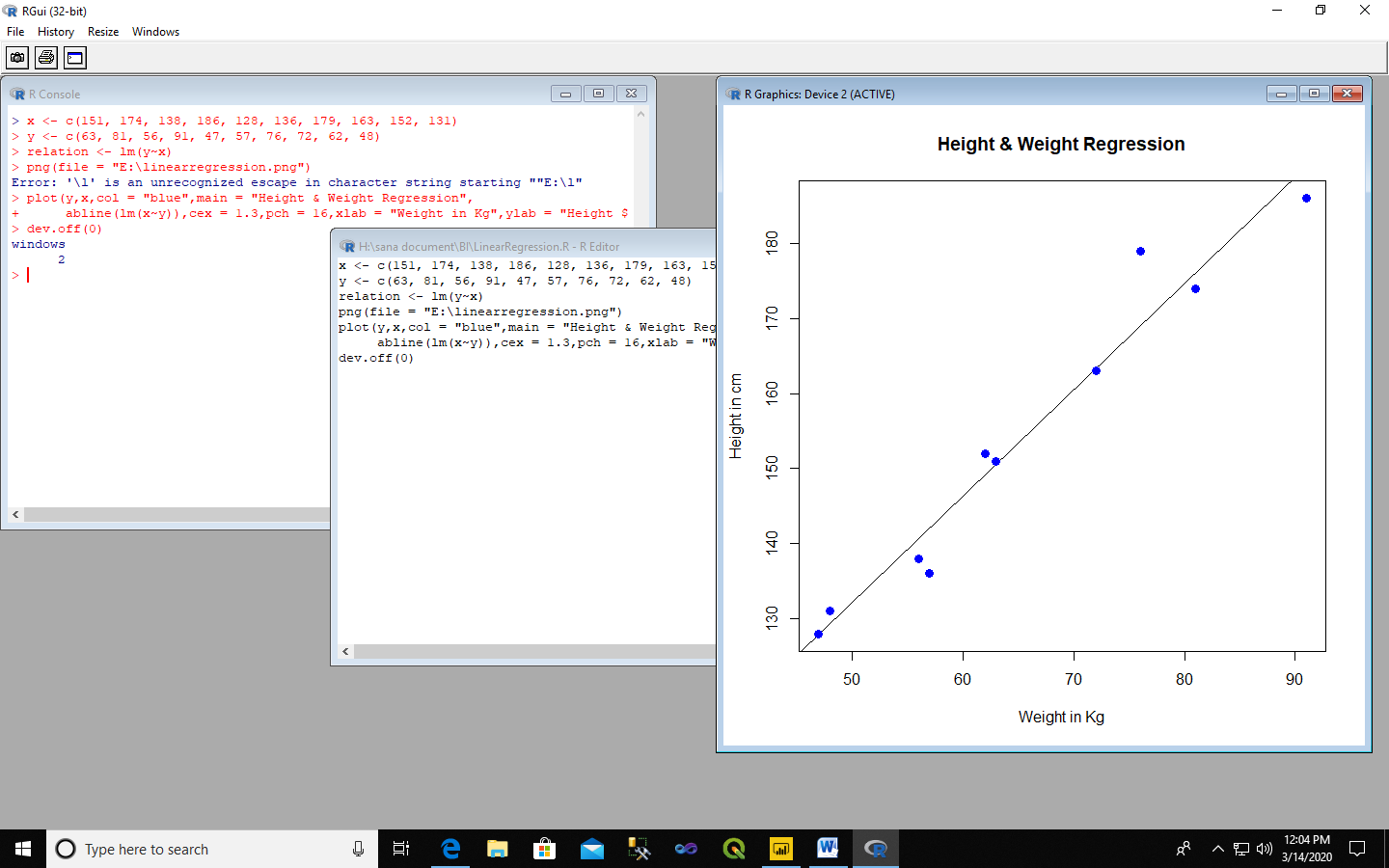
png(file = "E:\linearregression.png")

plot(y,x,col = "blue",main = "Height & Weight Regression",

abline(lm(x~y)),cex = 1.3,pch = 16,xlab = "Weight in Kg",ylab = "Height in cm")

dev.off(0)

**Output:**



**Practical No. 10**

**Data Analysis using Time Series Analysis.**

**Code:**

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)

dev.off(0)

**Output:**

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