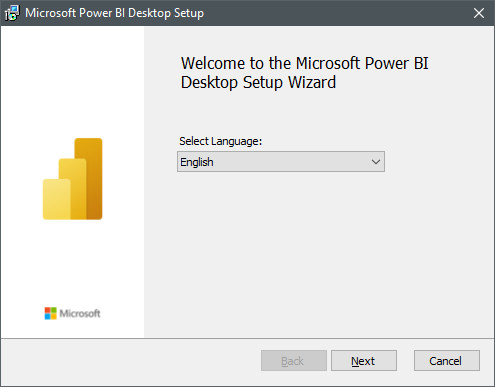
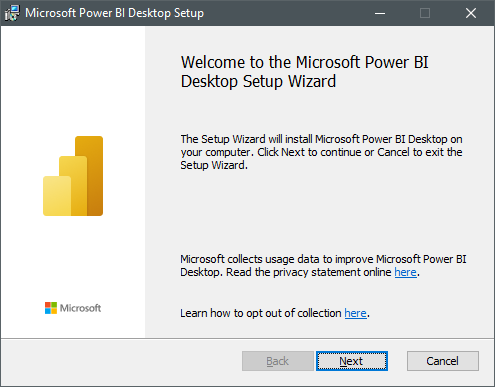
**Practical 0 :- Installing Power BI**

**Steps :**

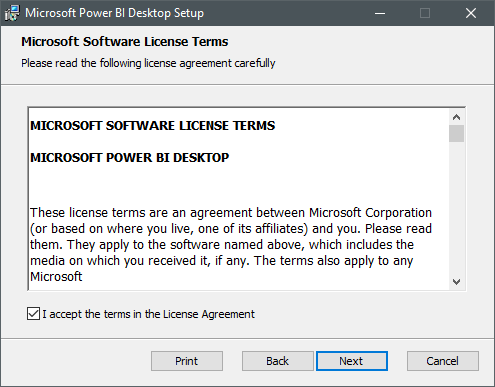
1. **Selct preferred language and Click Next .**



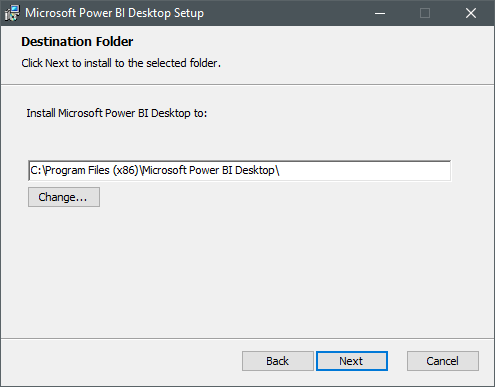
1. **Click Next .**



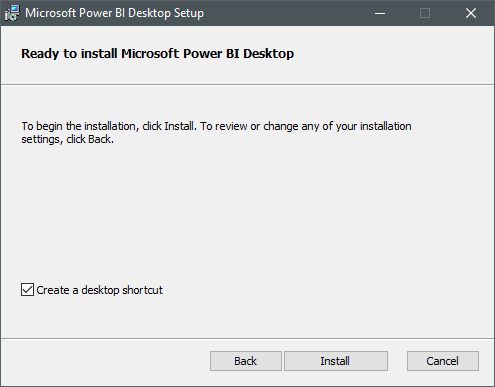
1. **Read rhe agreement and click on I agree**



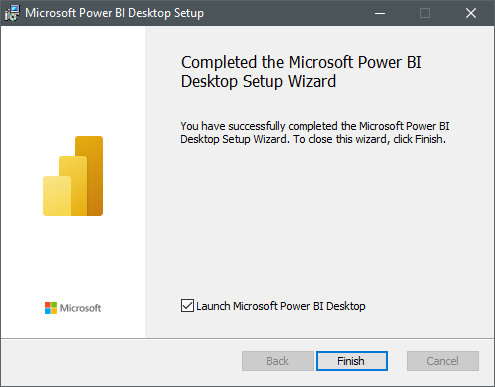
1. **Select Destination To save the Application**



1. **Click on Install .**

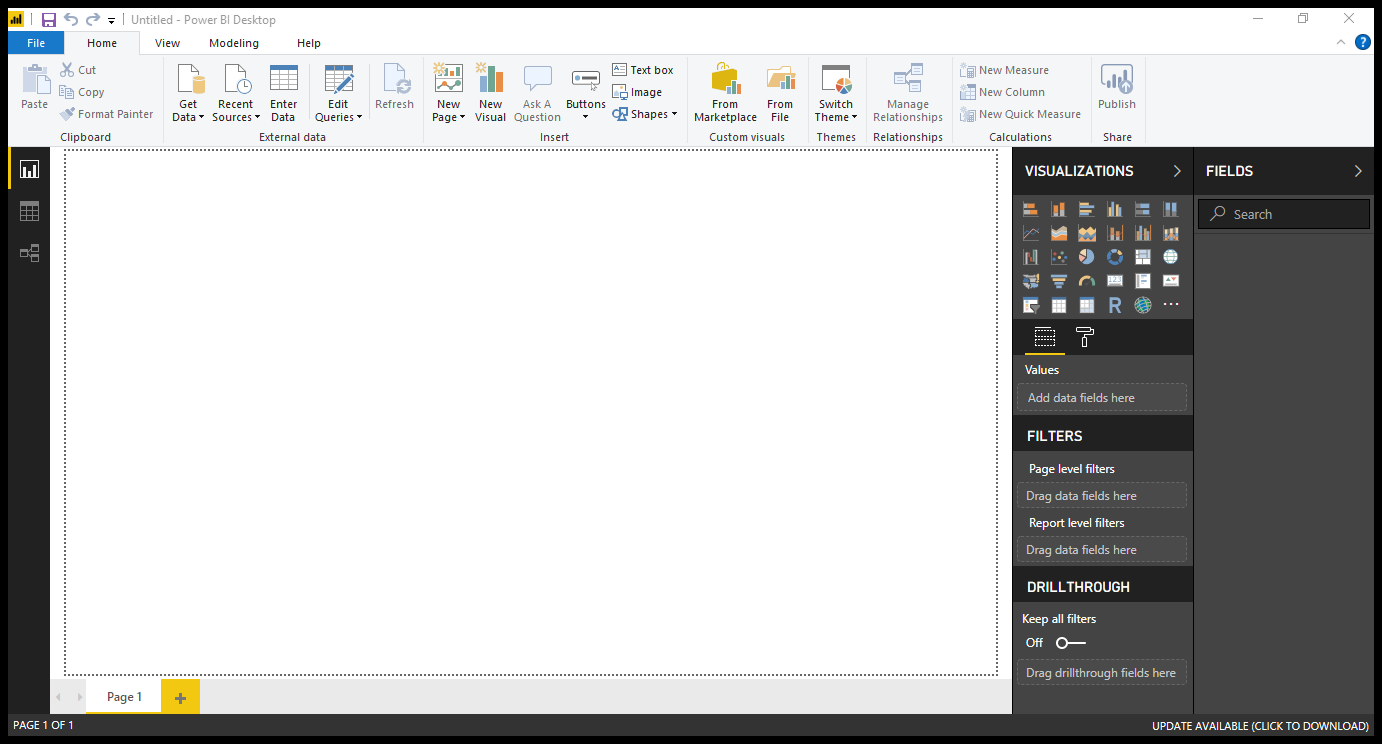


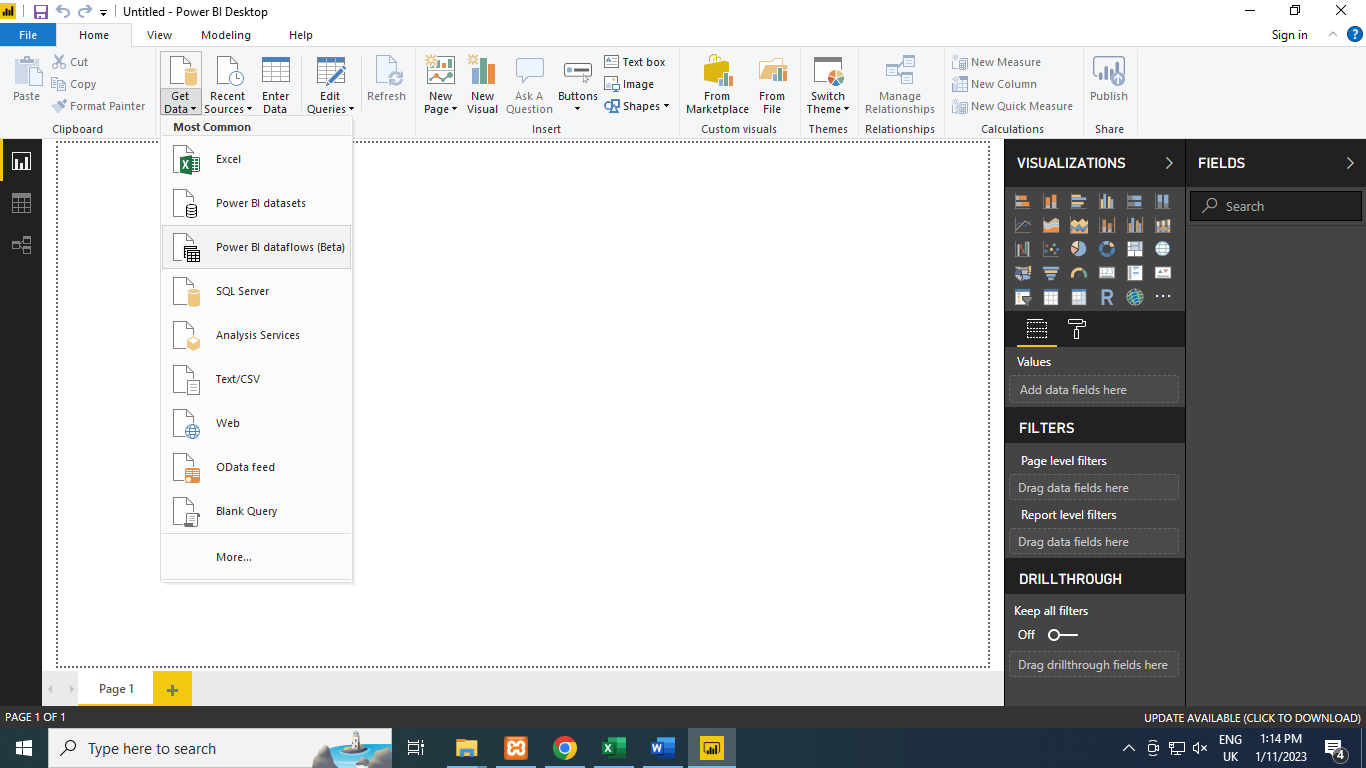
1. **The setup is finsished .**



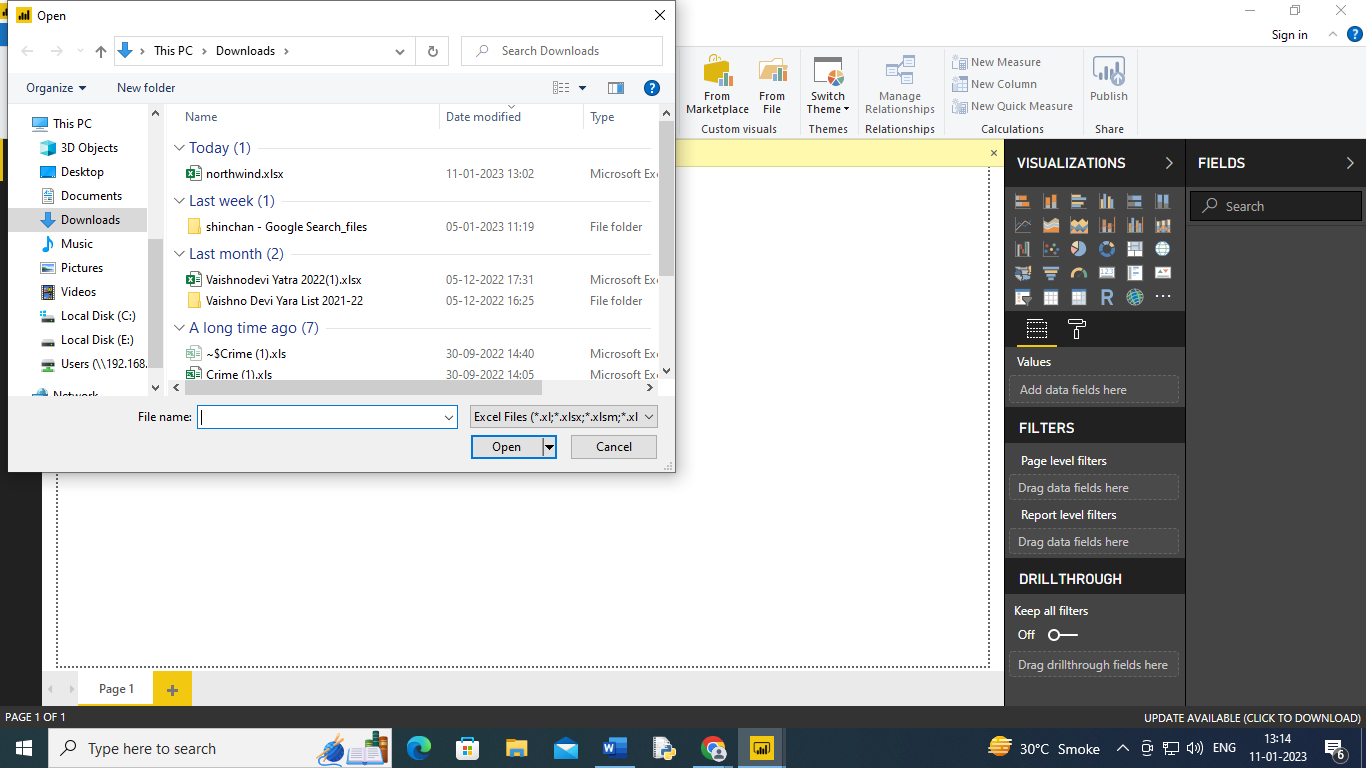
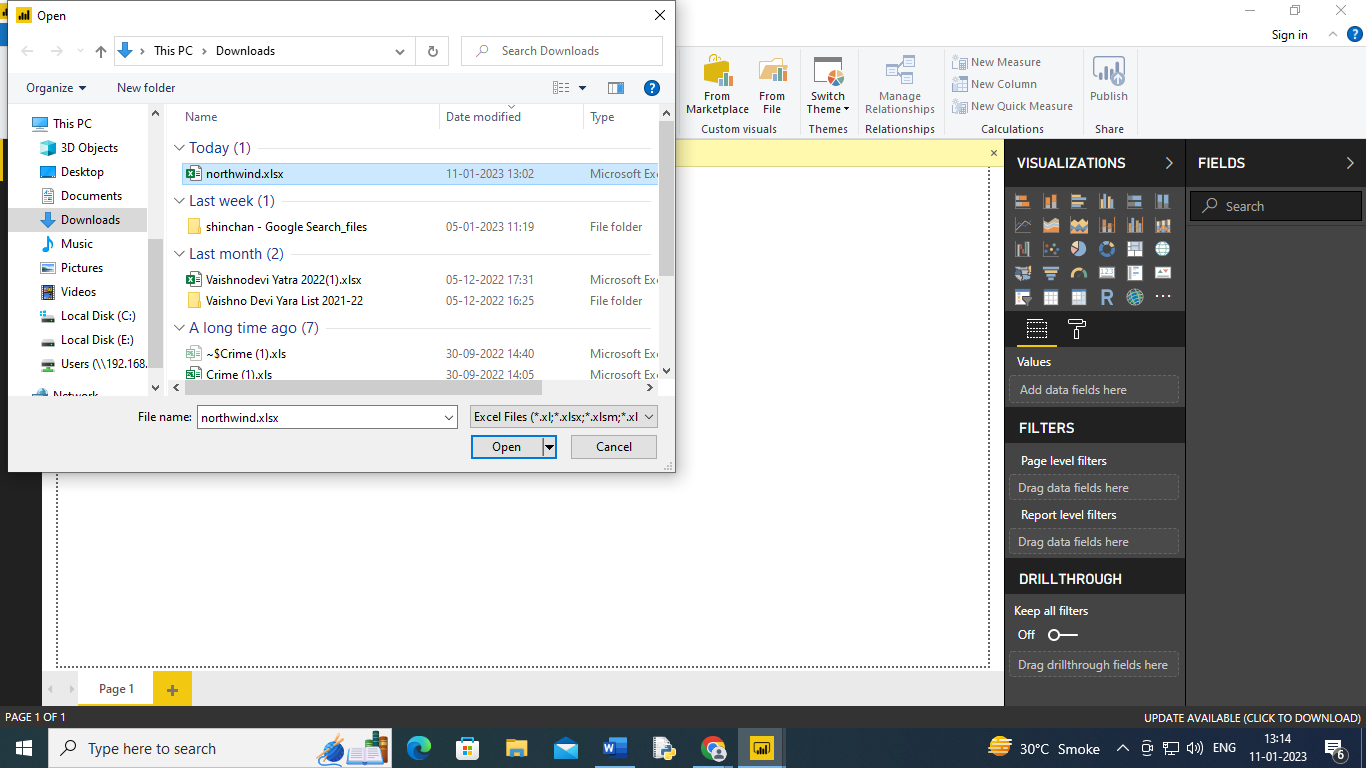
**Practical 1: Import the legacy data from different sources such as (Excel, SQL Server ,Oracle etc.) and load in the target system.**

**Step1: Open Power BI**

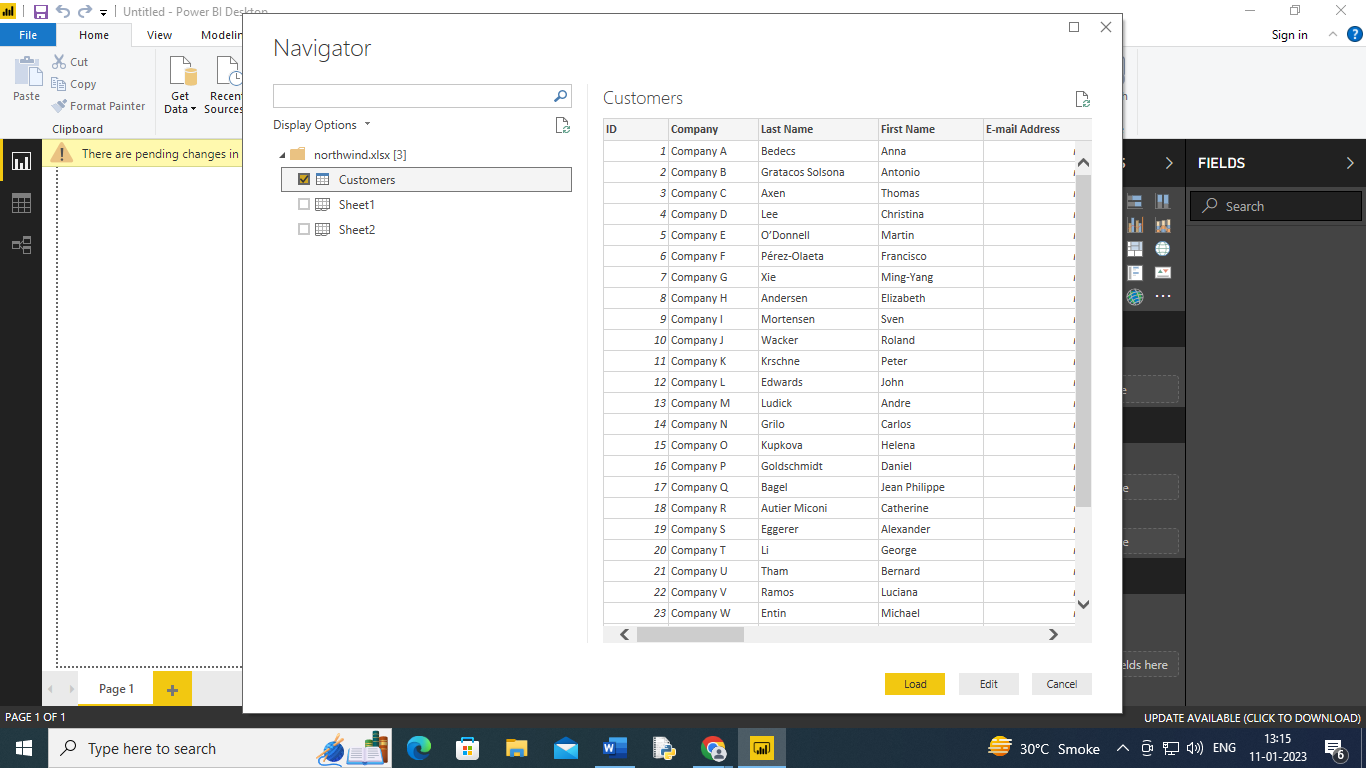


**Step2: Click on get data .Following list will be displayed -> select Excel** 

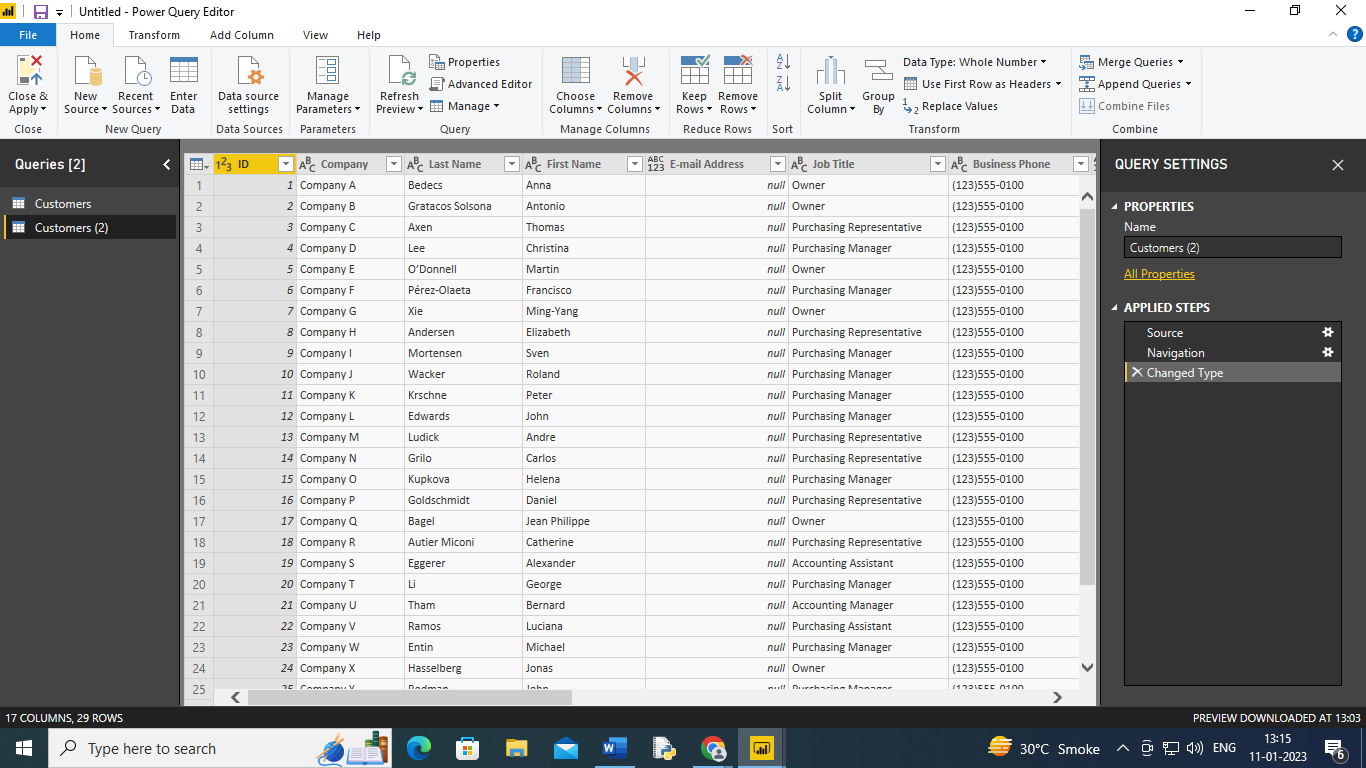
**Step 3: Select required file and click on Open, Navigator screen appears**



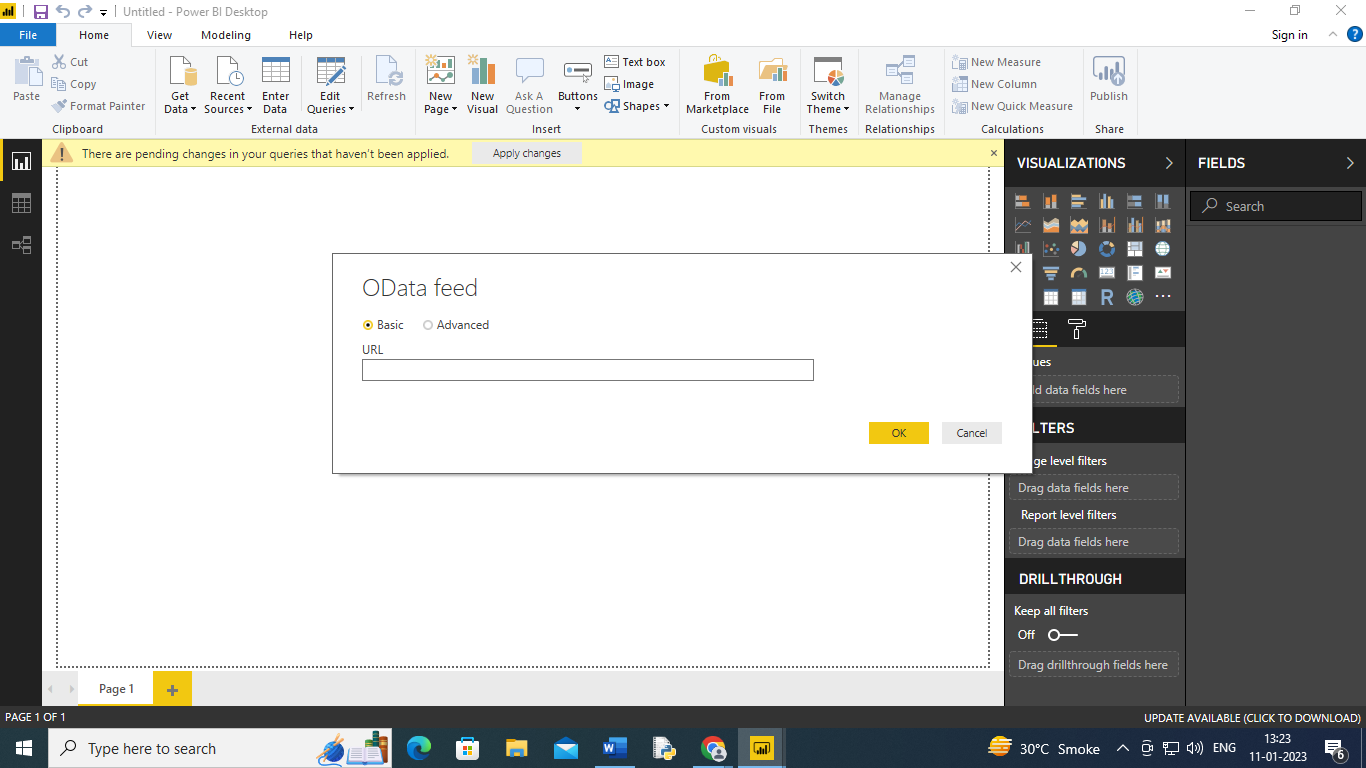
**Step 4: Select file and click on edit**



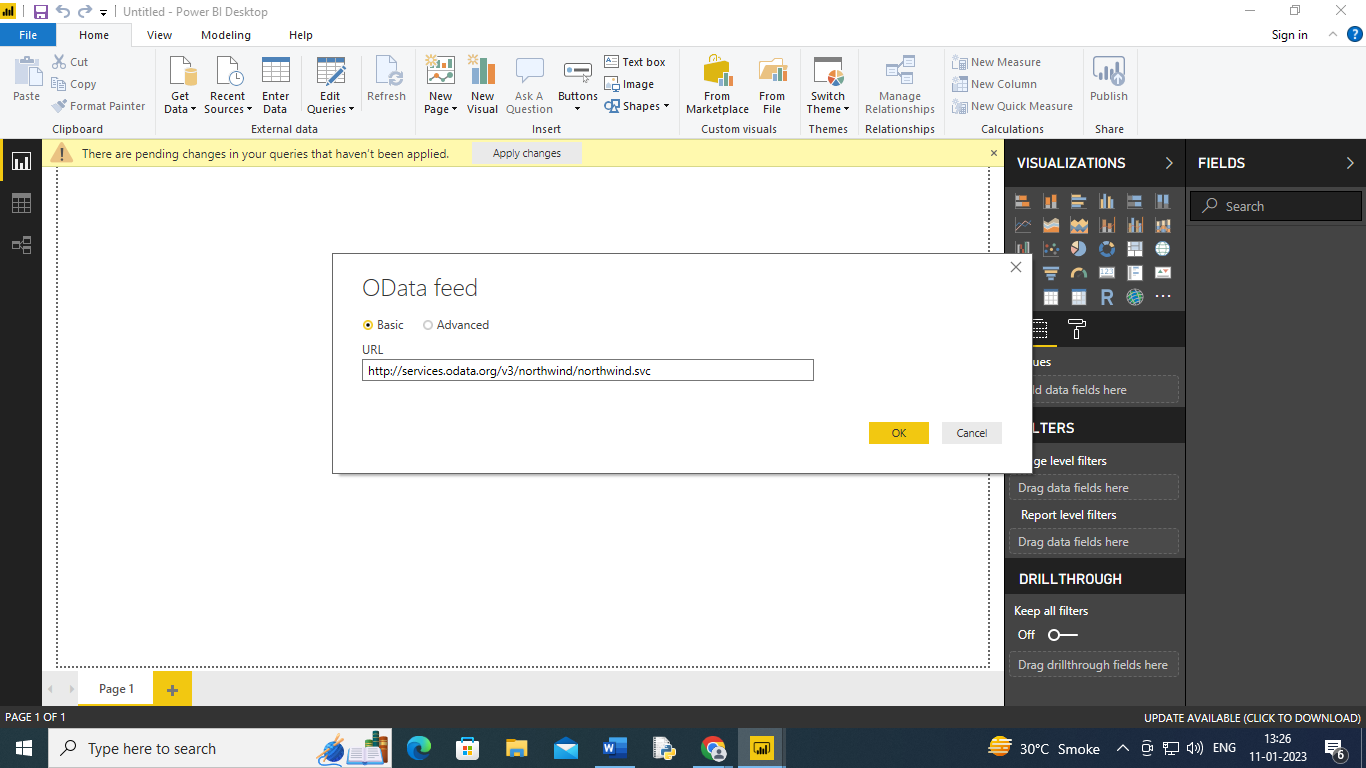
**Step5 :Power query editor appears**



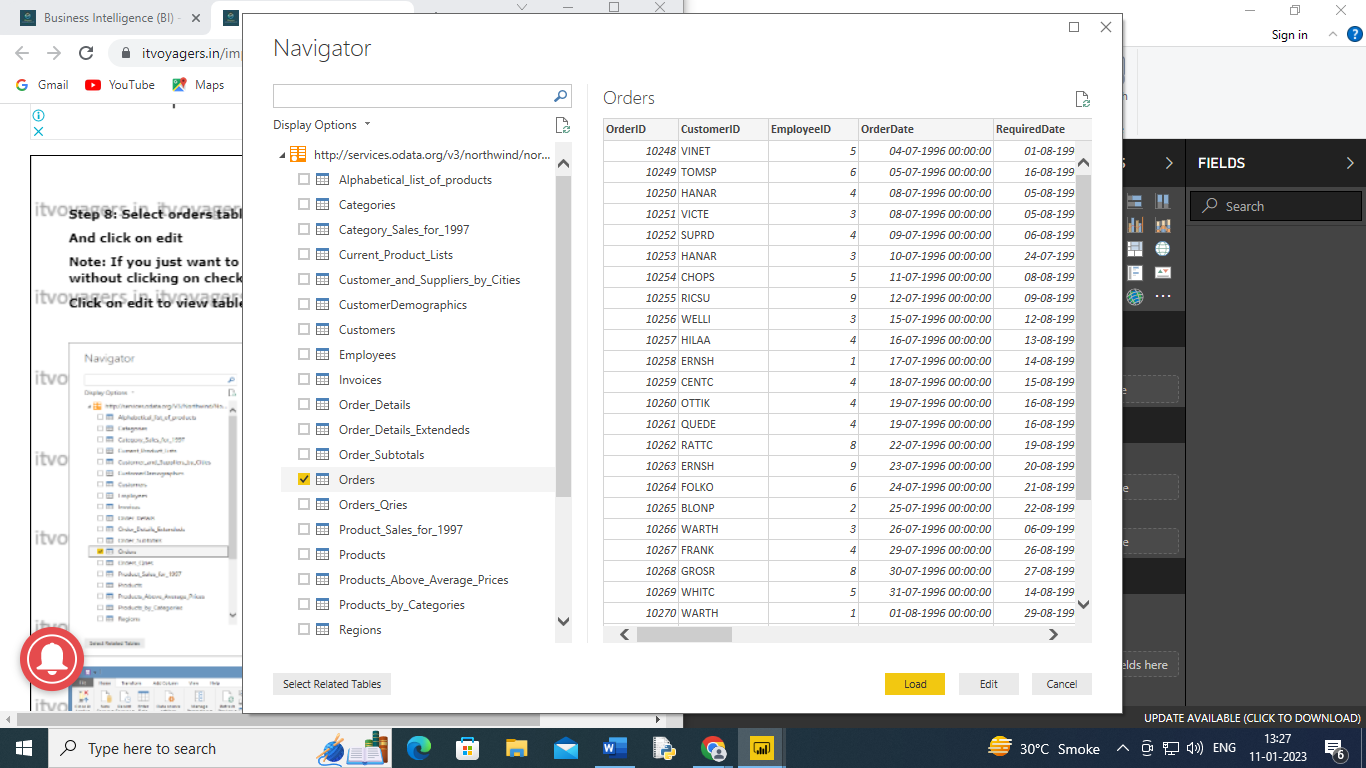
**Step 6 : Again click Get Data and select Odata feed**



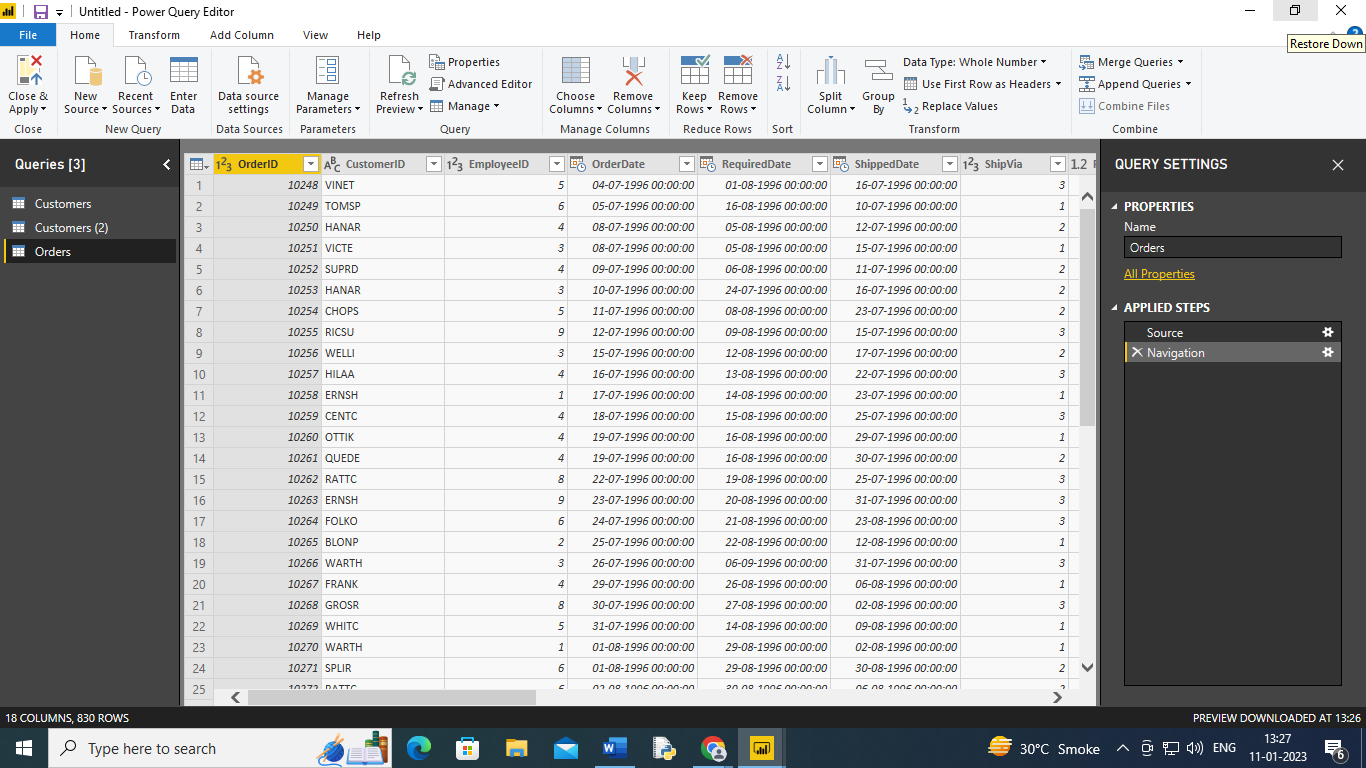
**Step 7: Paste url as** [**http://services.odata.org/v3/northwind/northwind.svc**](http://services.odata.org/v3/northwind/northwind.svc) **. Click on ok**



**Step 8: Select Orders table . And click on edit**

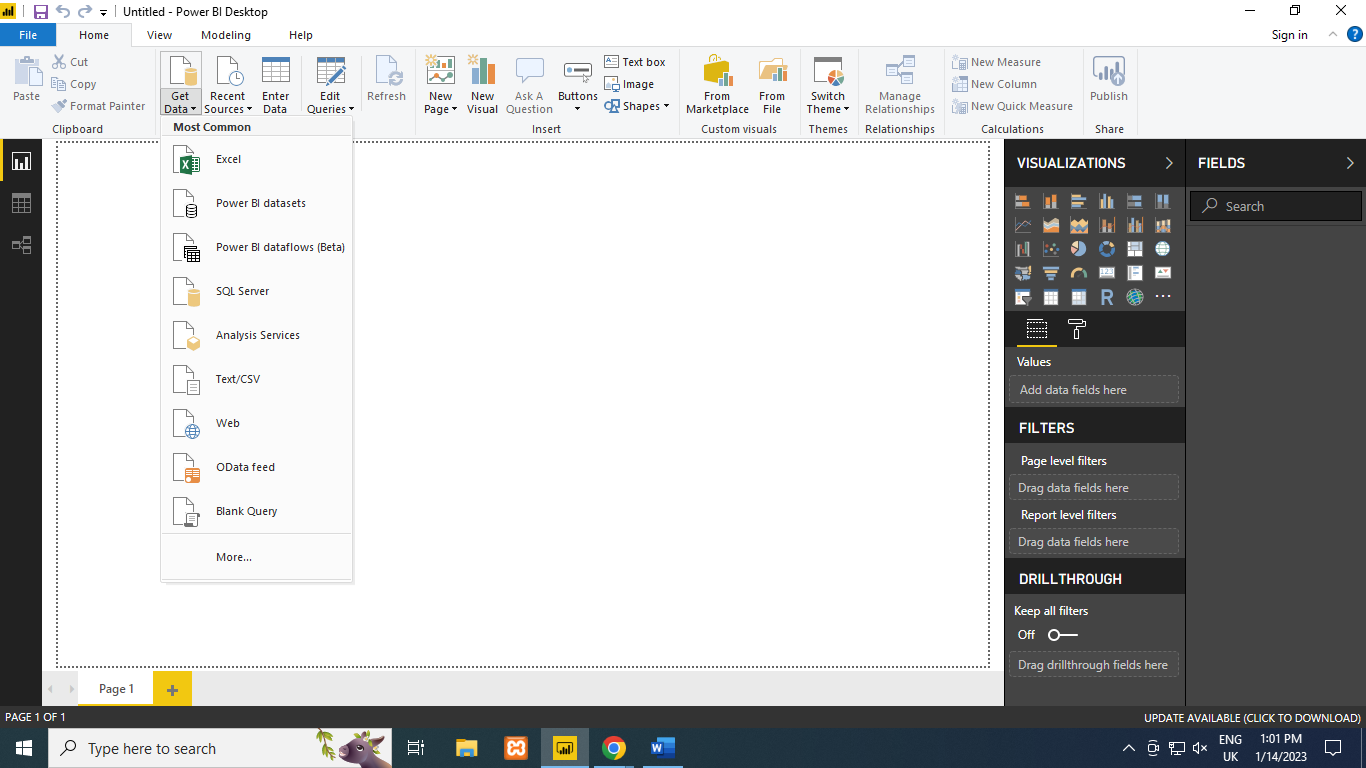


**Output :**

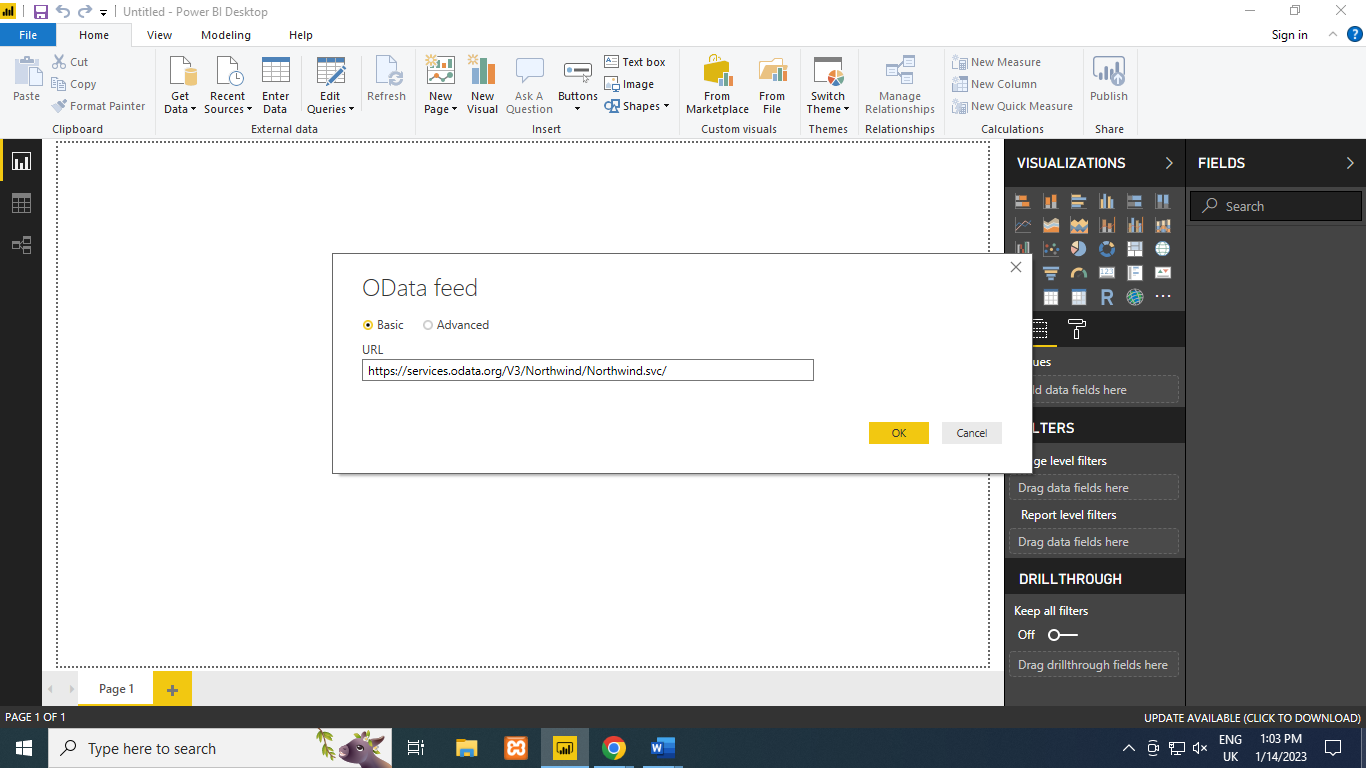


**Practical 2: Perform the Extraction Transformation and Loading (ETL) process to construct the database in the Power BI**

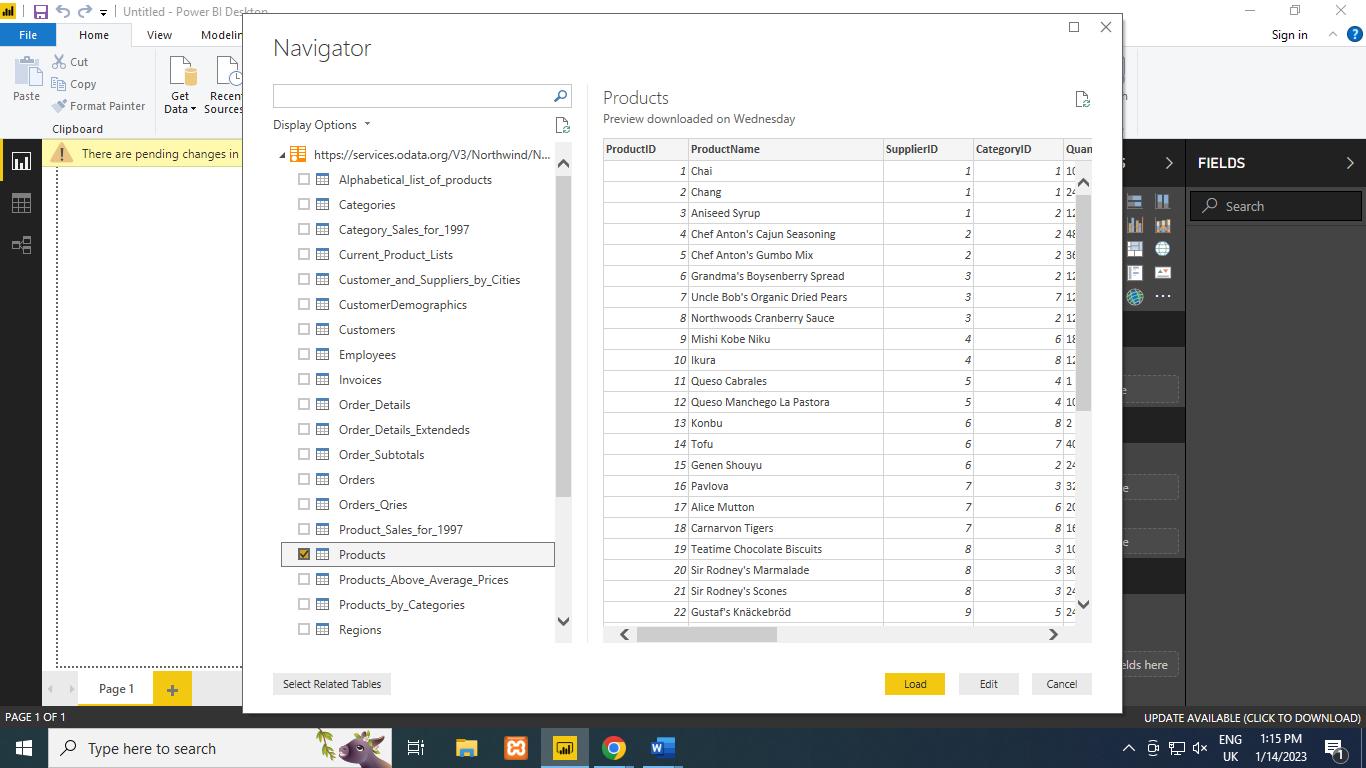
**Step1: Open Power BI .Click on Get Data -> OData Feed**

****

**Step 2: Paste url**

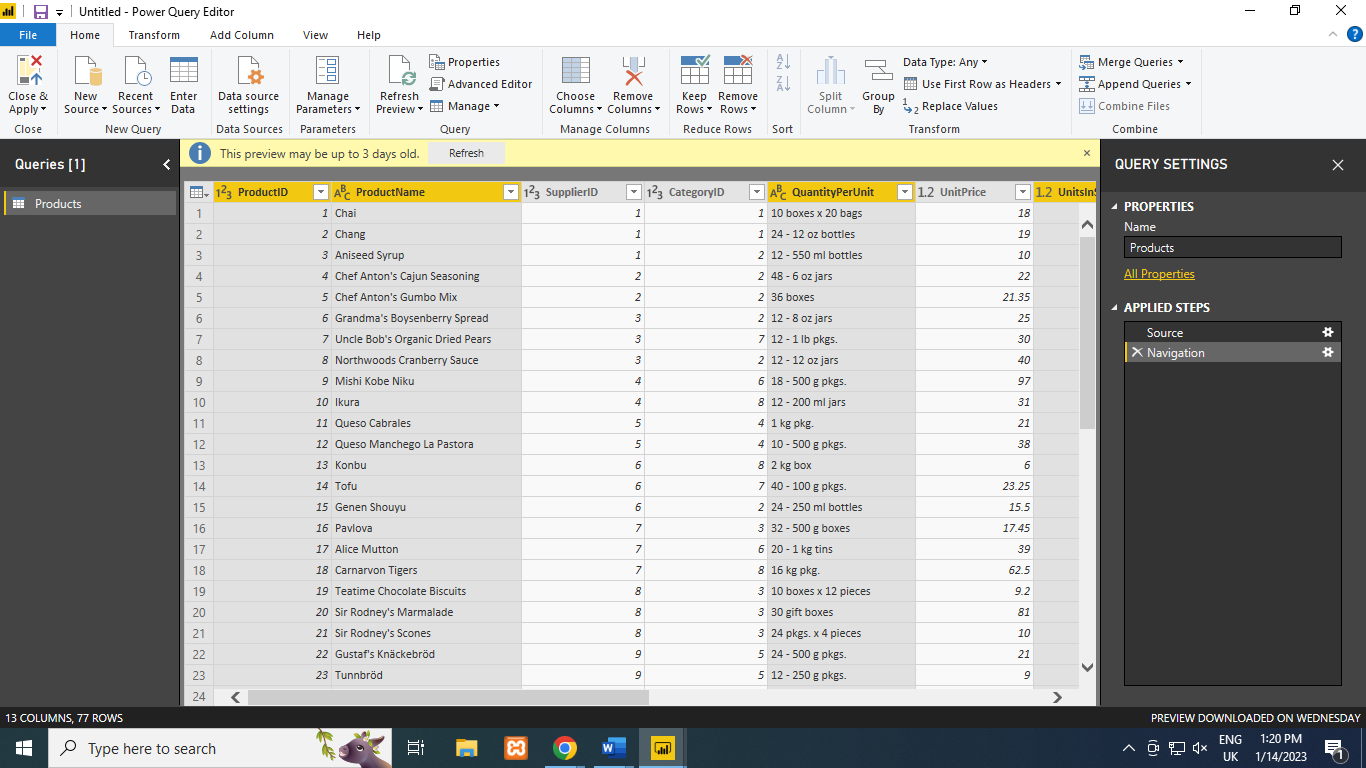
****

**Step 3:Click on Check Box of Products table and then Click on edit**

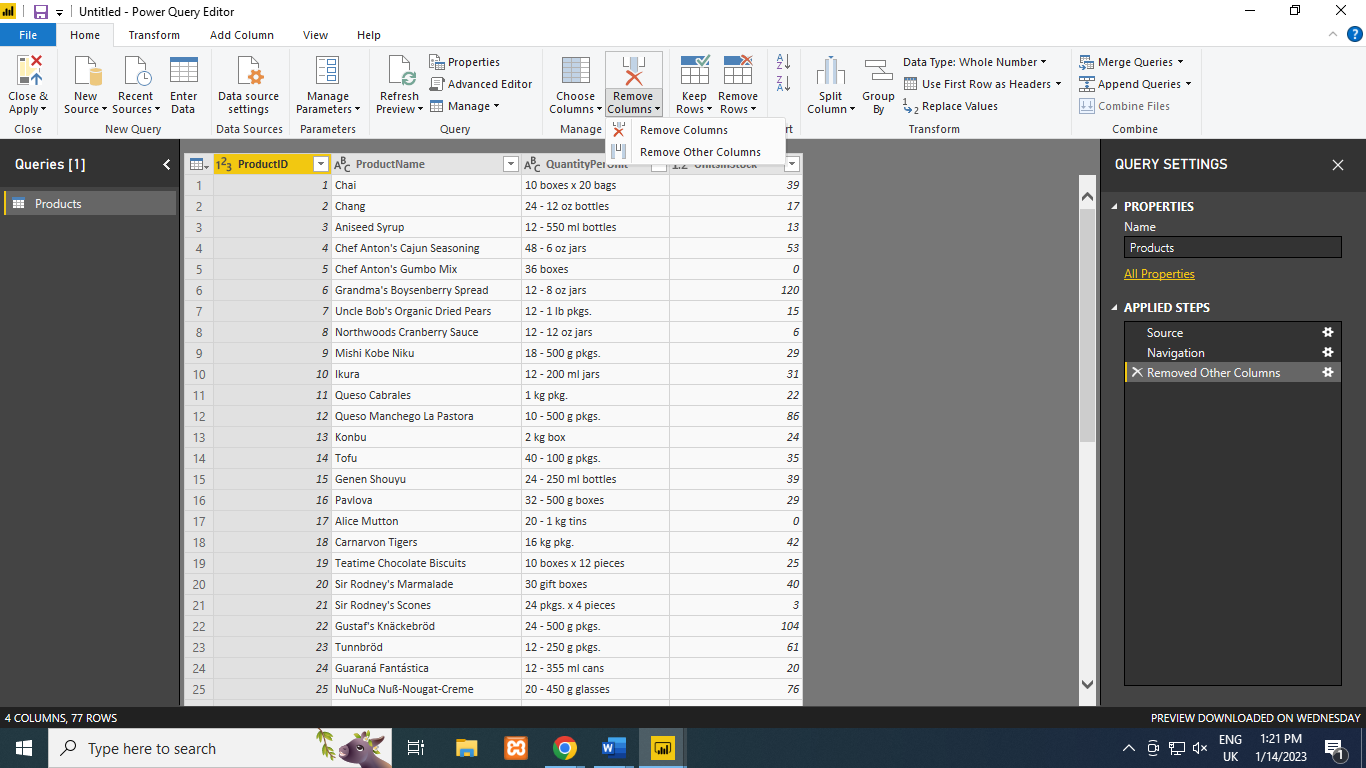
**.**

**Step 4: Remove other columns to display coloums of interest.In Query Editor, Select the ProductId, ProductName, QuantityPerUnit and UnitsInStock columns**

**Select Remove Columns > Remove Other Columns from the ribbon , or right click on a column header and click Remove Other Columns**

****

**After Selecting Remove other Columns only selected four columns are displayed other columns are discarded.**

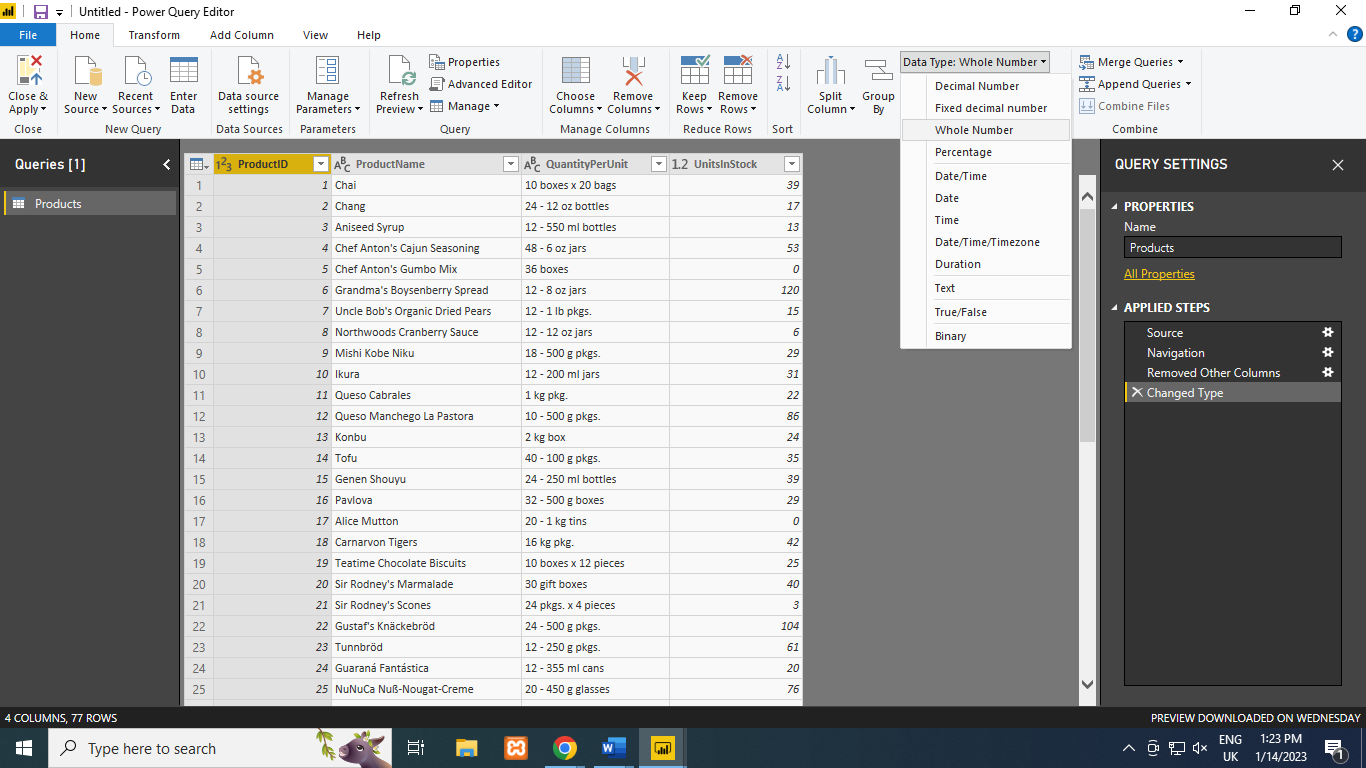
****

**Step 2: Change the data type of the UnitsInStock column.**

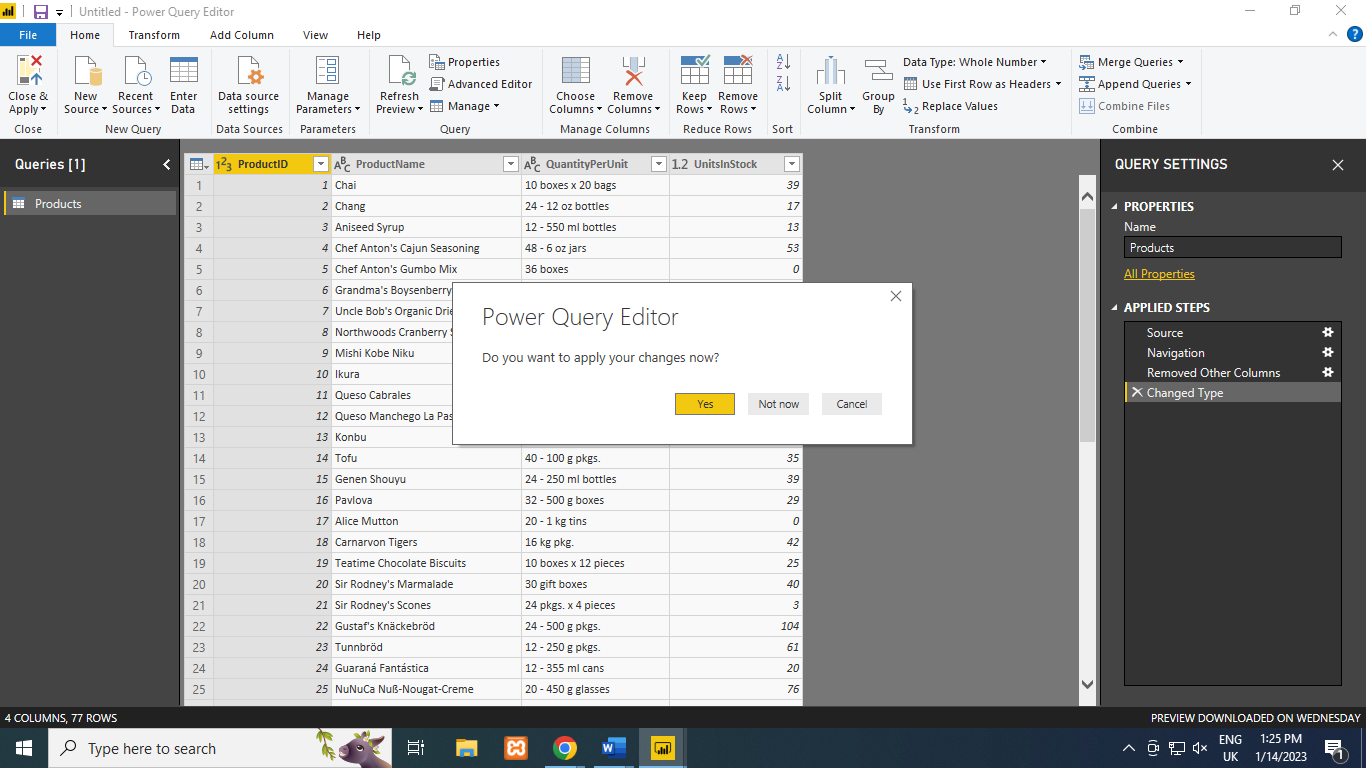
**A)Select the UnitsInStock Column**

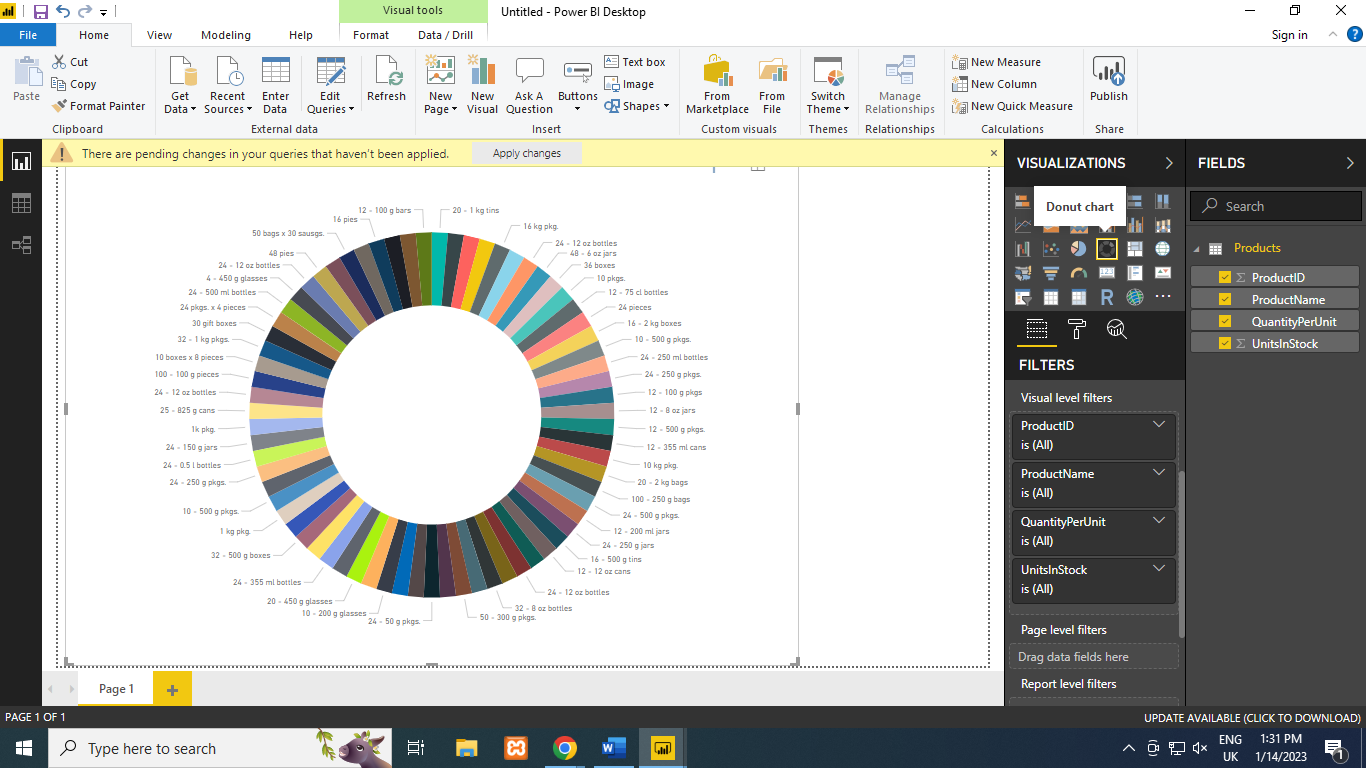
**B) Select the Data Type drop-down button im the Home ribbon .**

**C) Select data Type as Whole Number**

****

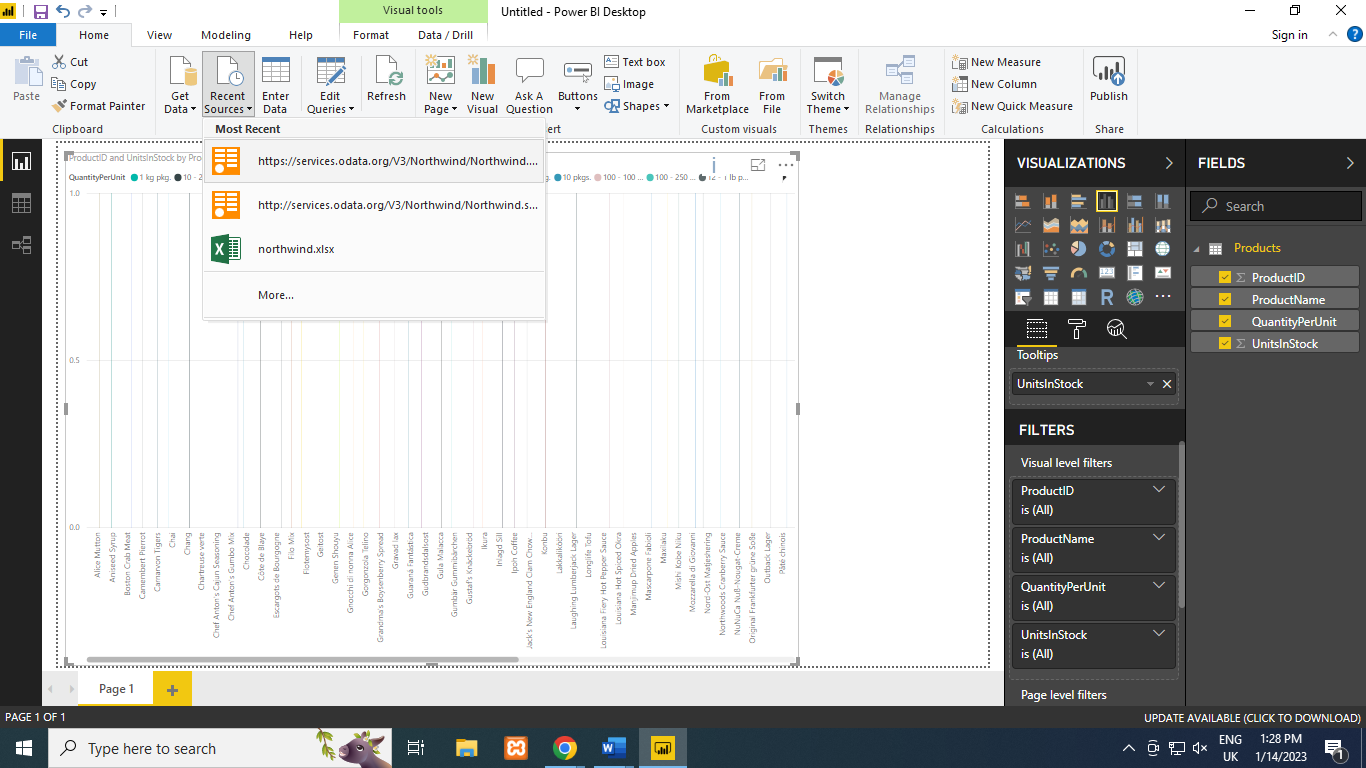
**Close Query editor and Click Yes to save changes.**

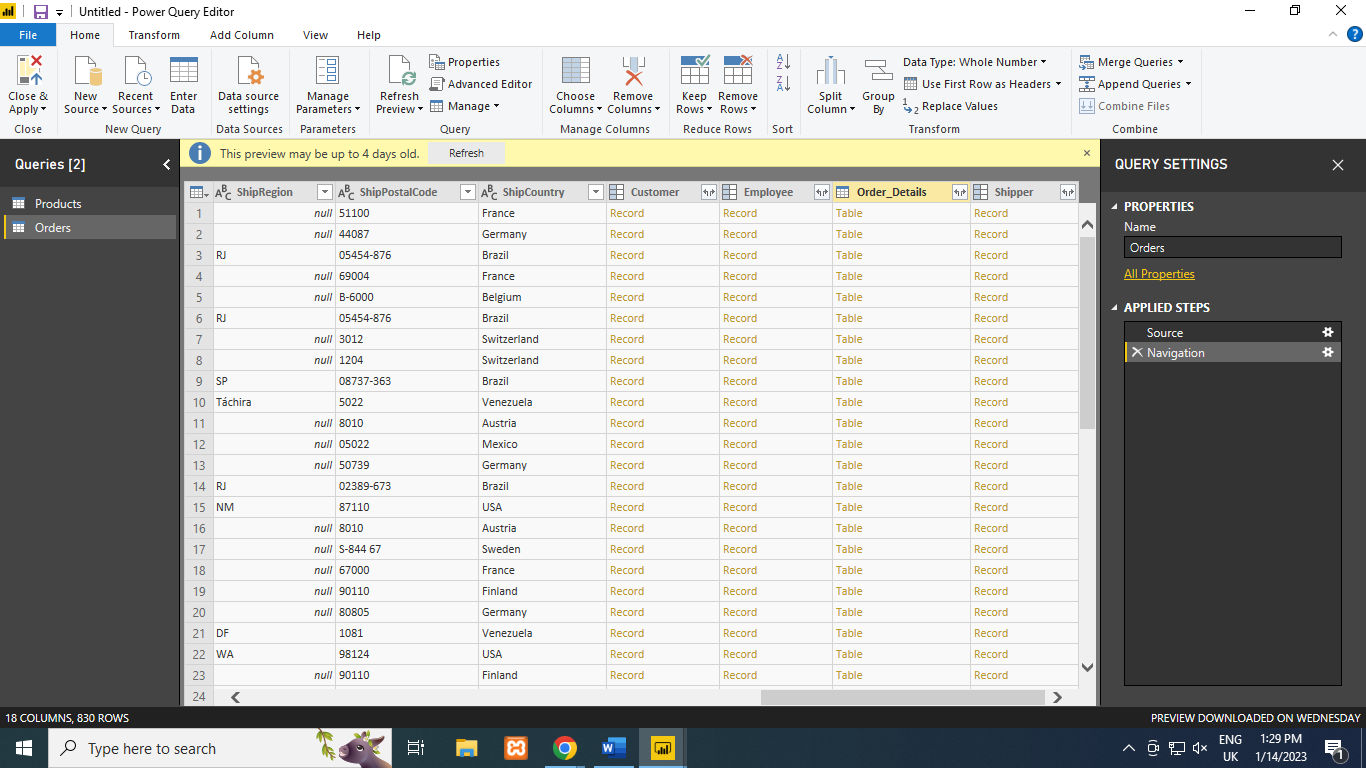
****

**Check all the fields visible on right side . To get representation in Chart form **

**Step 3: Expand Orders Table**

**Click on Recent Sources > Orders table .**

****

****

**Query Editor Window .**

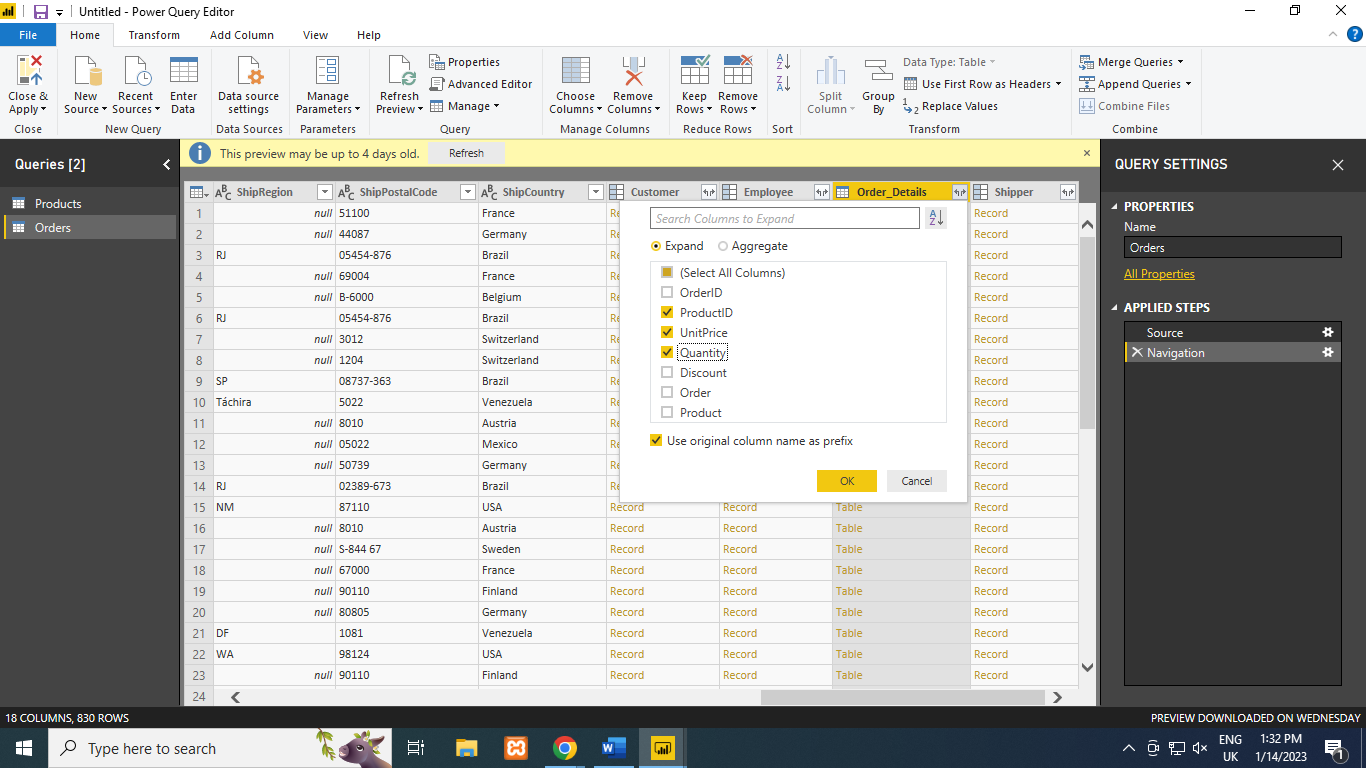
**1. In the Query view , scroll to the Order\_Details column .**

**2. In the Order\_details column, select expand icon.**

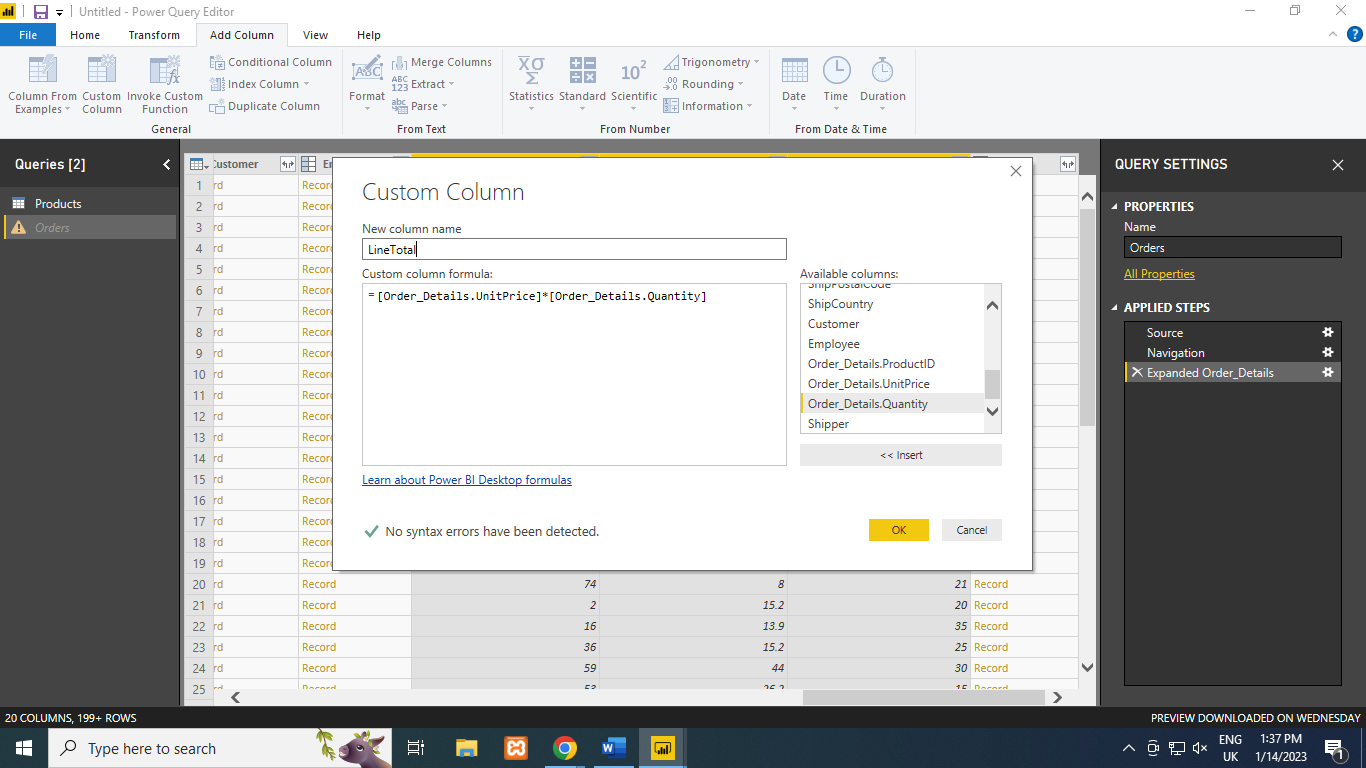
**3.In the Expand Drop down :**

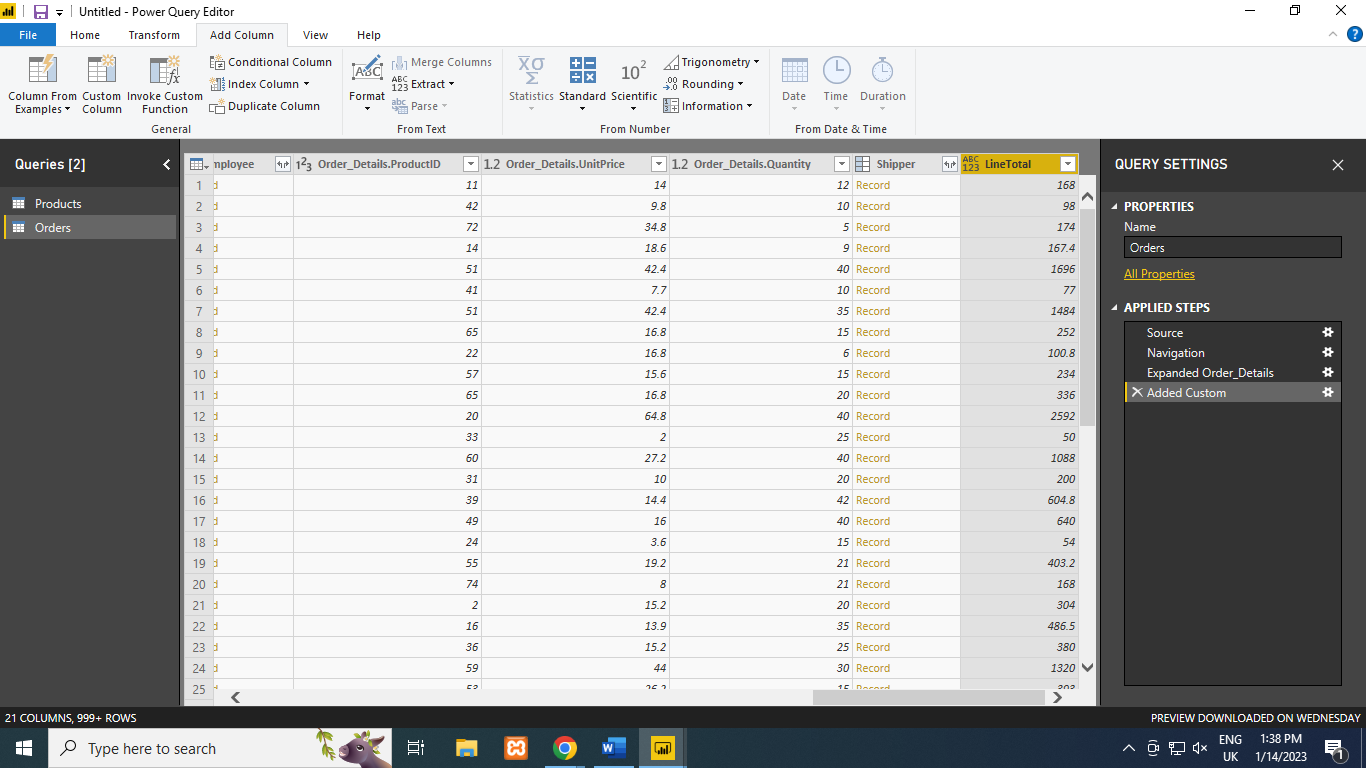
**a)select all columns to clear all columns**

**b)Select ProductID , UnitPrice, and Quantity. Click Ok.**

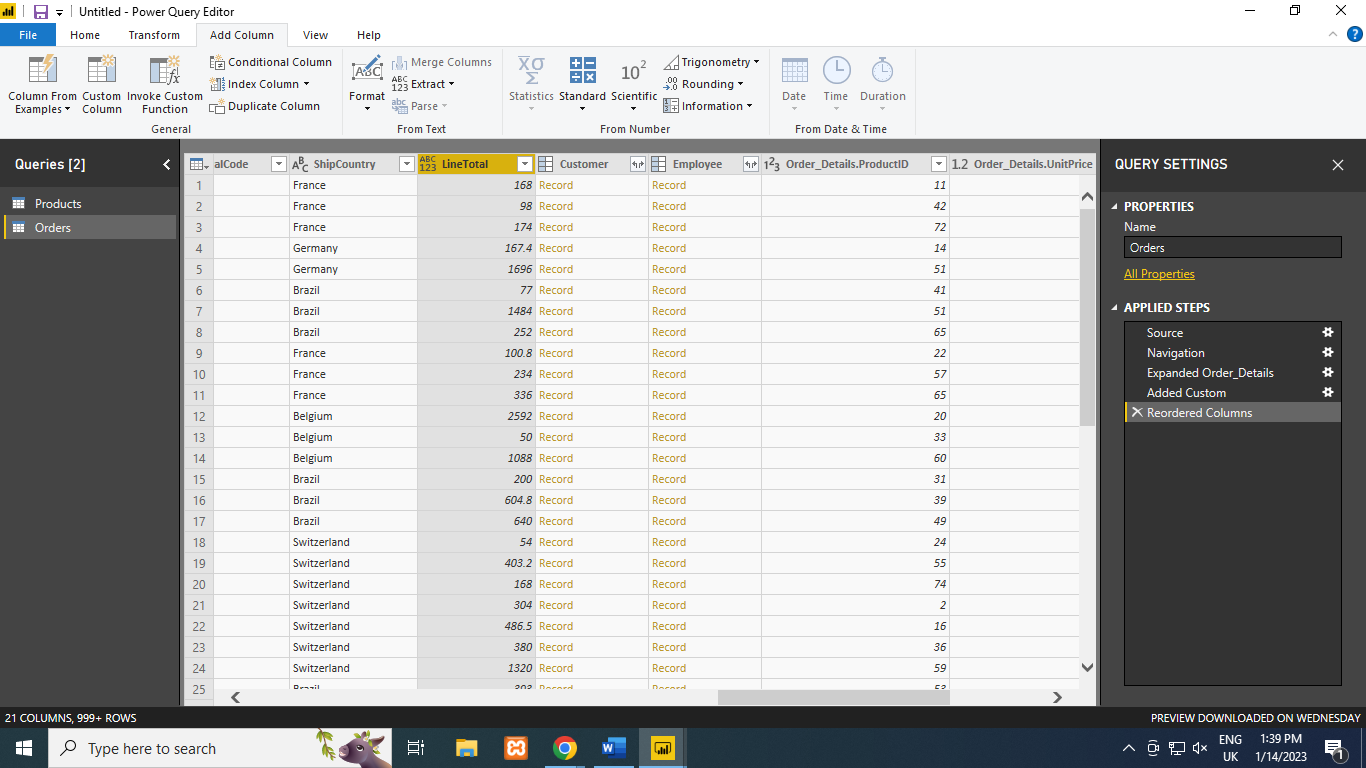
****

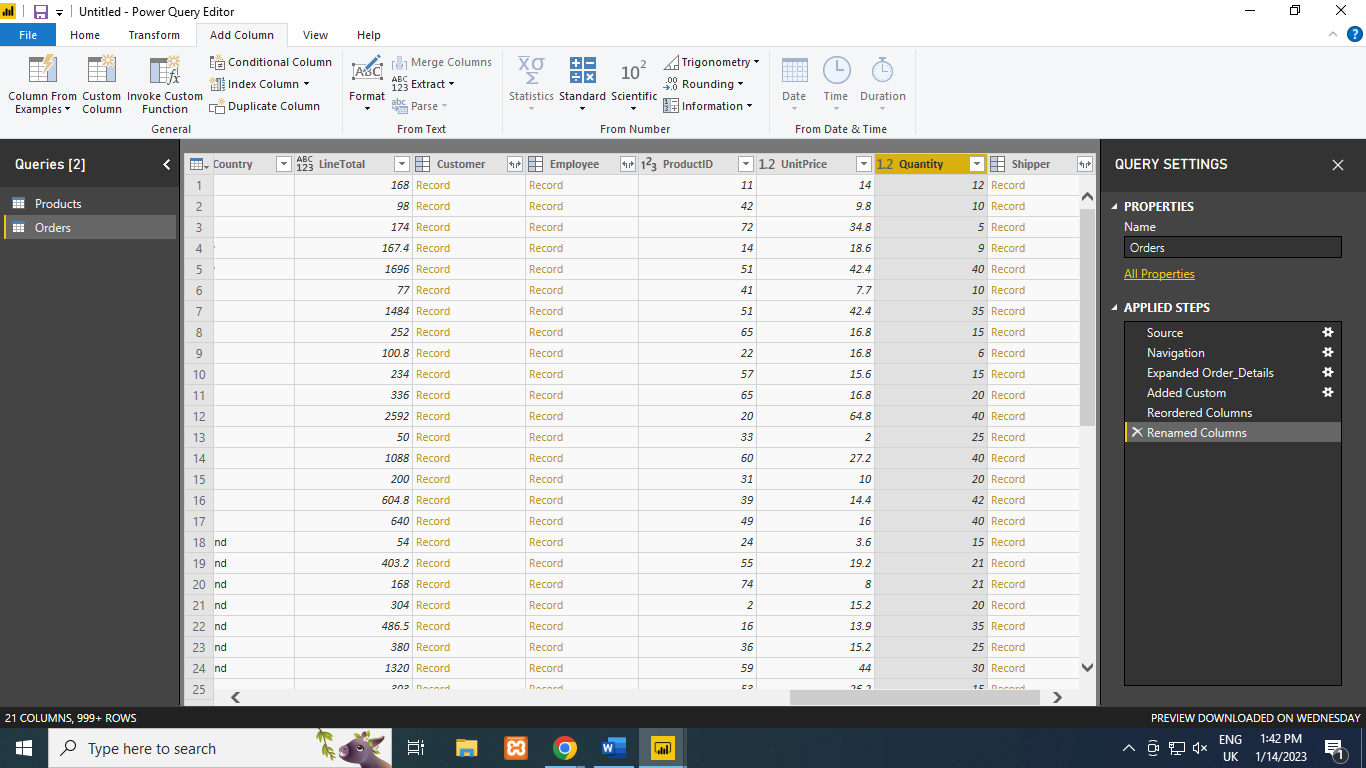
**In add column click add custom culomn > Enter formula > name: Line Total > ok**

****

****

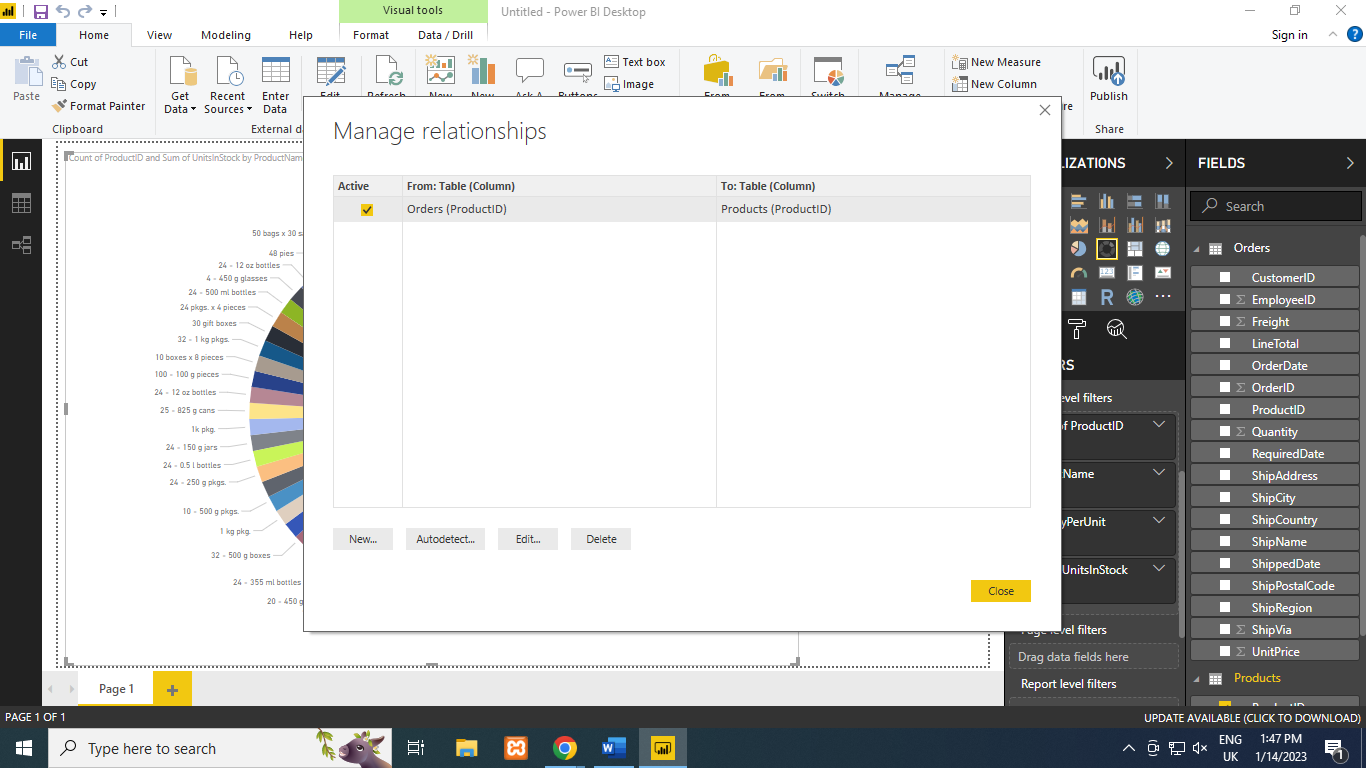
**Step 5: Drag Linetotal after ShipCountry**

****

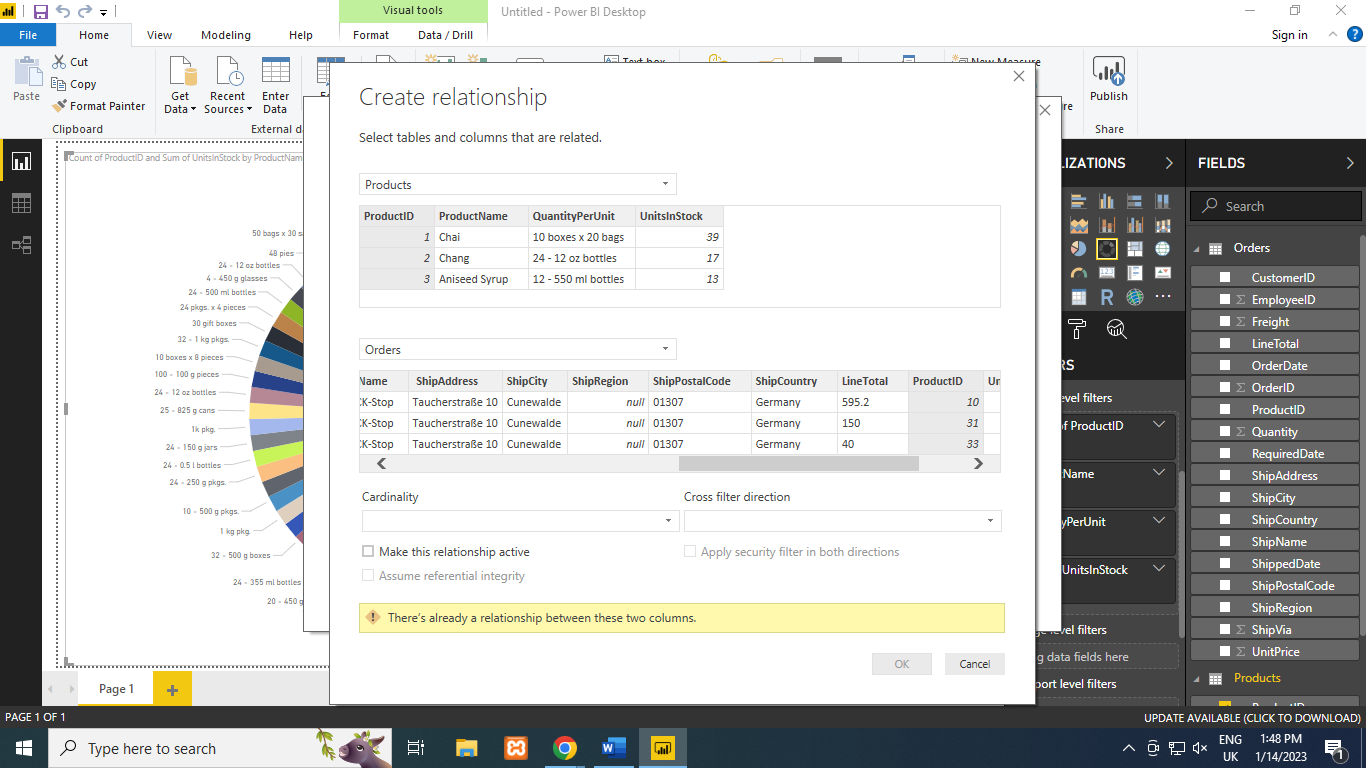
**-> Rename all order.detials by removing prefix Order.Details **

**Step 6: From Home > select Close & Apply**

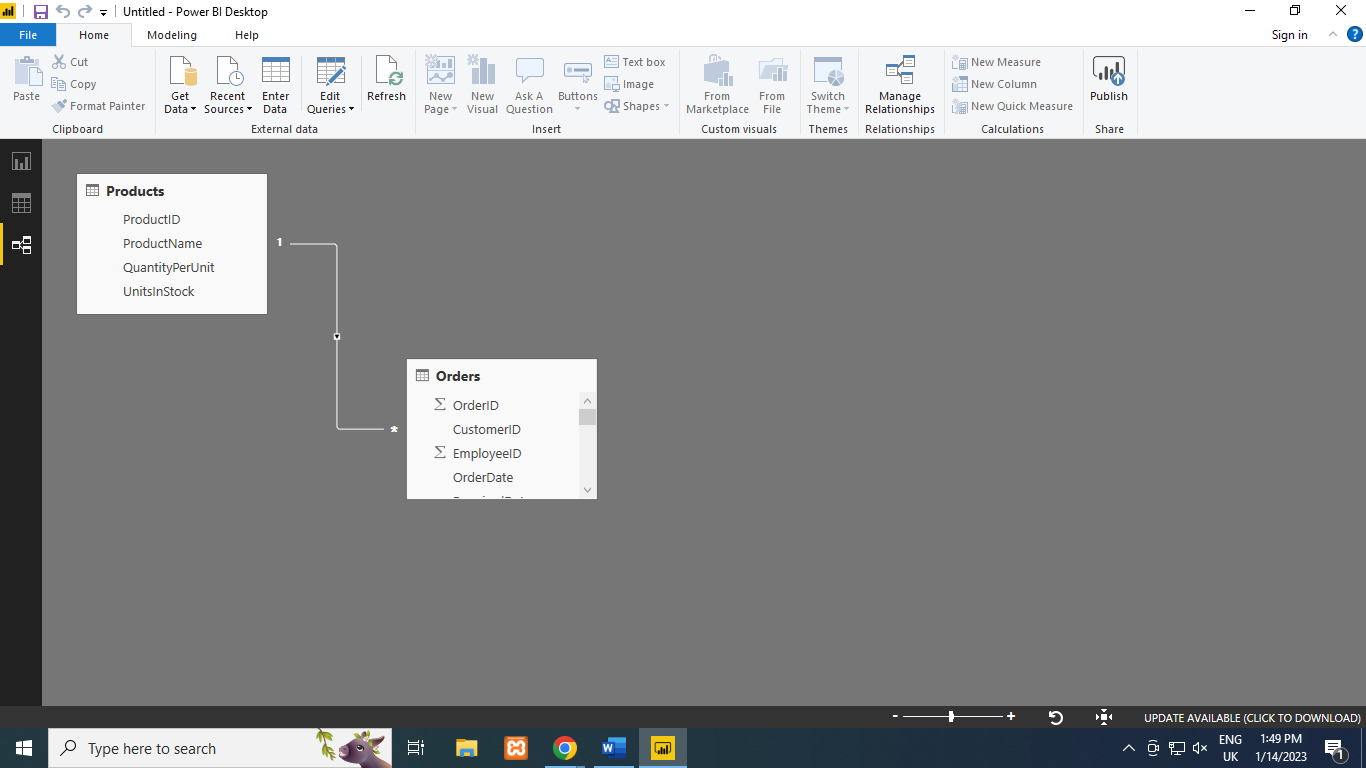
* **Power BI Desktop loads the Data from the Queries.**
* **Once the data is loaded Select The Manage Relationships Button Home ribbon.**
* **Select New**

****

**When we attempt to create we see it already exists -> Select cancel**

****

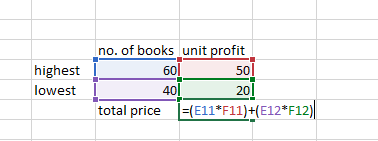
**And Select Relationship view on Power BI Desktop .**

****

**Practical 6 :-Apply the what – if Analysis for data visualization. Design and generate necessary reports based on the data warehouse data.**

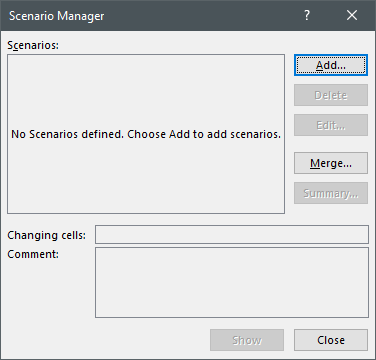
**Q. A book store and have 100 books in storage . You sell a certain % for the highest price of $50 and certain % for the lower price of $20.**



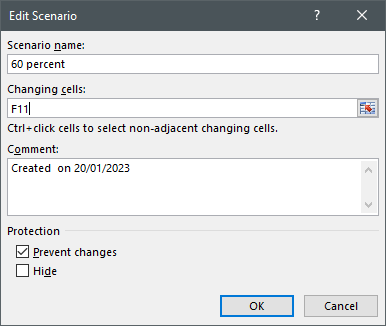


**Steps:**

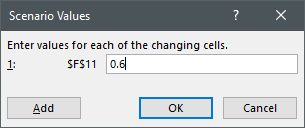
1. **In excel , on the data tab in the data tools group , click on What If Analysis .**
2. **Click on What-If-Analysis & select scenario Manager .**



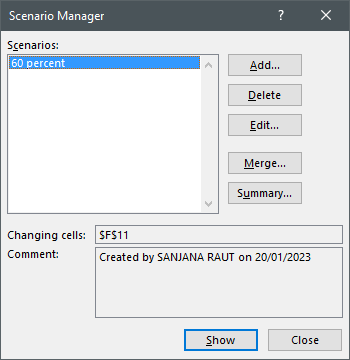
1. **Add a scenario by clicking on add**
2. **Type name 60 Percent , select shell (50) , for the changing cells & click on ok .**

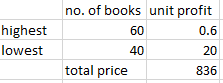


1. **Enter the corresponding value 0.6,click on OK again.**

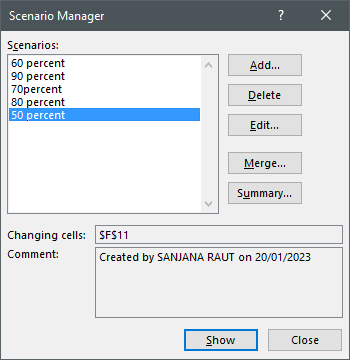


1. **To apply scenario . click on show .**





1. **Next do similar , by adding four more scenarios .**



**Practical 07:- Perform the data classification using classification algorithm.**

**Q. 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 .**

**Steps:- Code in R software .**

1. **Get the data point in the form of a R vector**

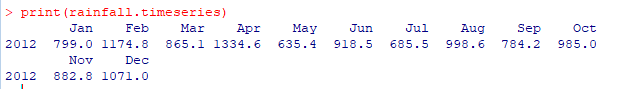
**rainfall <- c(799,1174.8,865.1,1334.6,635.4,918.5,685.5,998.6,784.2,985,882.8,1071)**

1. **Convet it to a time series object.**

**rainfall.timeseries <- ts(rainfall,start=c(2012,1),frequency=12)**

1. **Print the time series data .**

**print(rainfall.timeseries)**

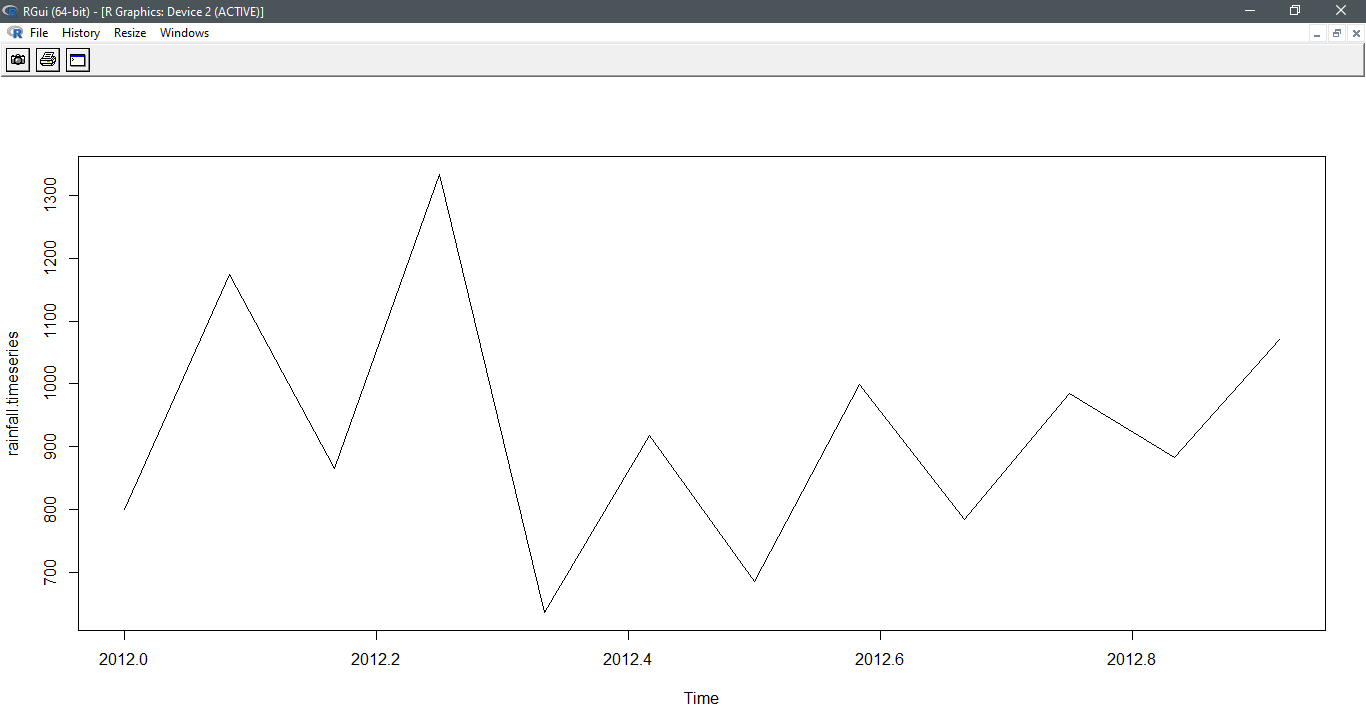


1. **Plot the graph of the time series**

**plot(rainfall.timeseries)**

1. **Give the chart file a name**

**png(file="rainfall.png")**



1. **Save the file**

**dev.off()**

1. **After this again plot to get the chart**

**plot(rainfall.timeseries)**

**Practical 8 :- Perform the data clustering using clustering algorithm**

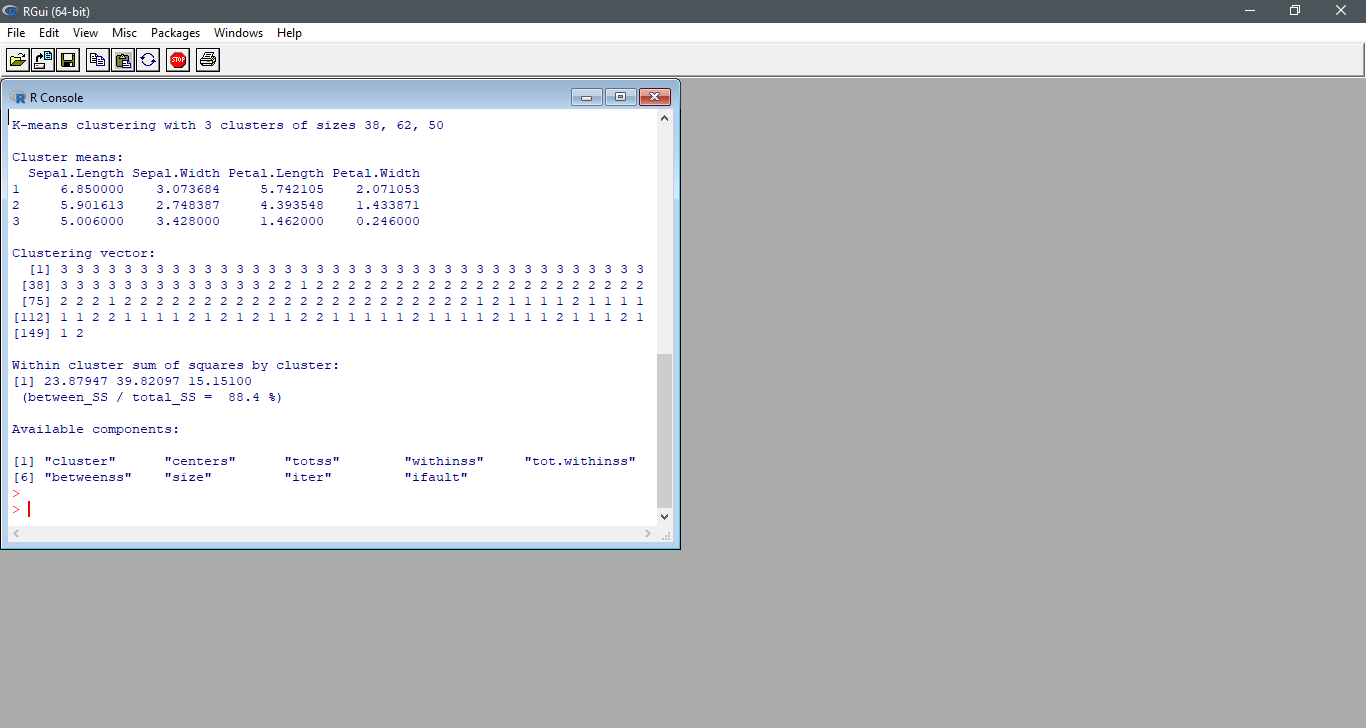
**Steps:**

1. **Apply K means to iris and store result .**

**>newiris <- iris**

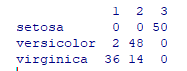
**> newiris $ Species <- NULL**

**> (kc<-kmeans(newiris,3))**



1. **Compare the species label with the clustering result.**

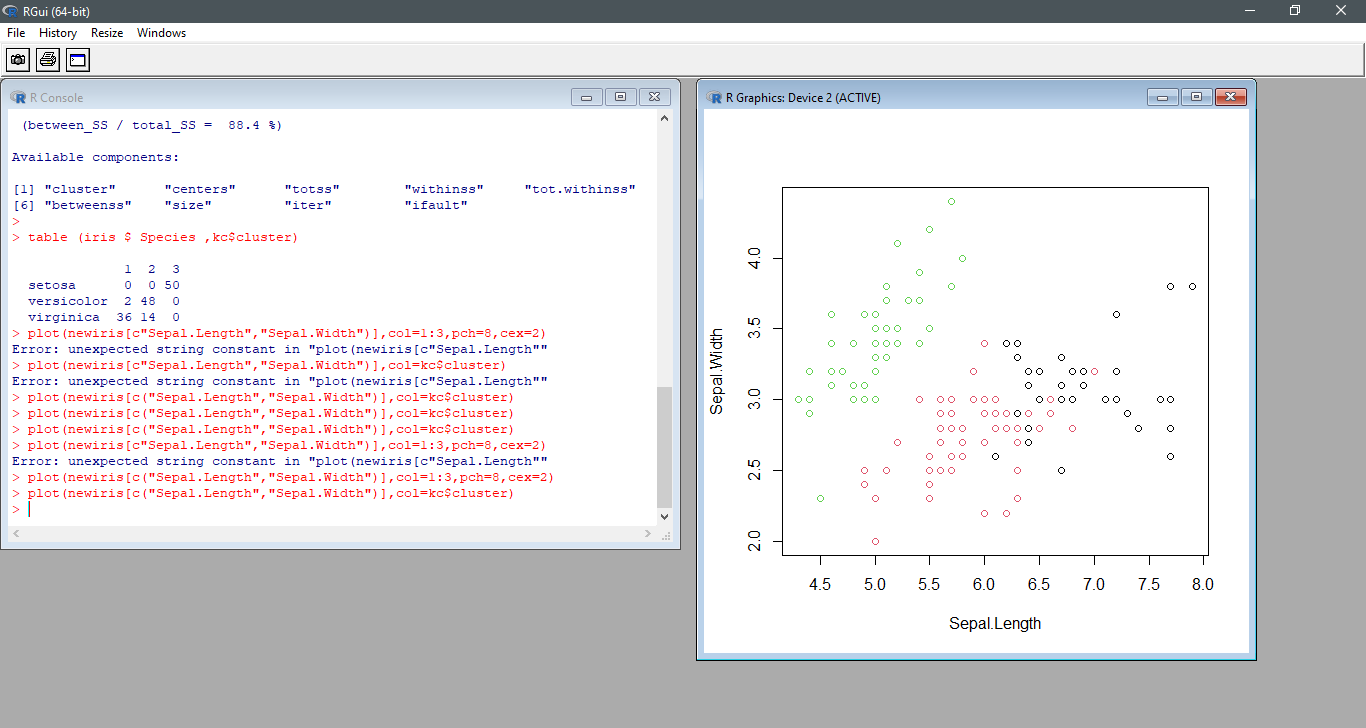
**> table (iris $ Species ,kc$cluster)**



1. **Plot the clusters and the centers**

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

**>plot(newiris[c("Sepal.Length","Sepal.Width")],col=1:3,pch=8,cex=2)**



1. **Save the file**

**> dev.off()**



1. **Again Plot the cluster and centers**

**Practical 09:-Perform the Linear regression on the given data warehouse data.**

**Steps :**

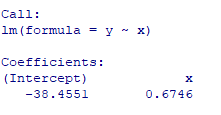
1. **Input data**
2. **Get Relationship Model & get the coefficients.**

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

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

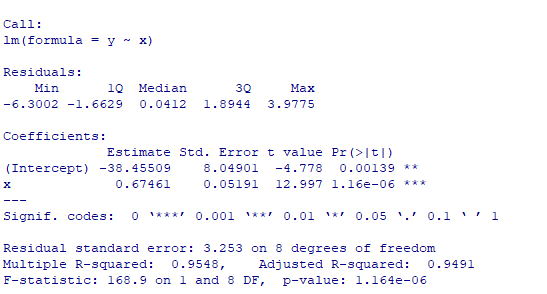
**> relation<- lm(y~x)**

**> print(relation)**



1. **Get the summary of the relationship.**

**> print(summary(relation))**



1. **Predict the weight of new persons .**

> a<-data.frame(x=170)

> result <- predict(relation,a)

> print (result)

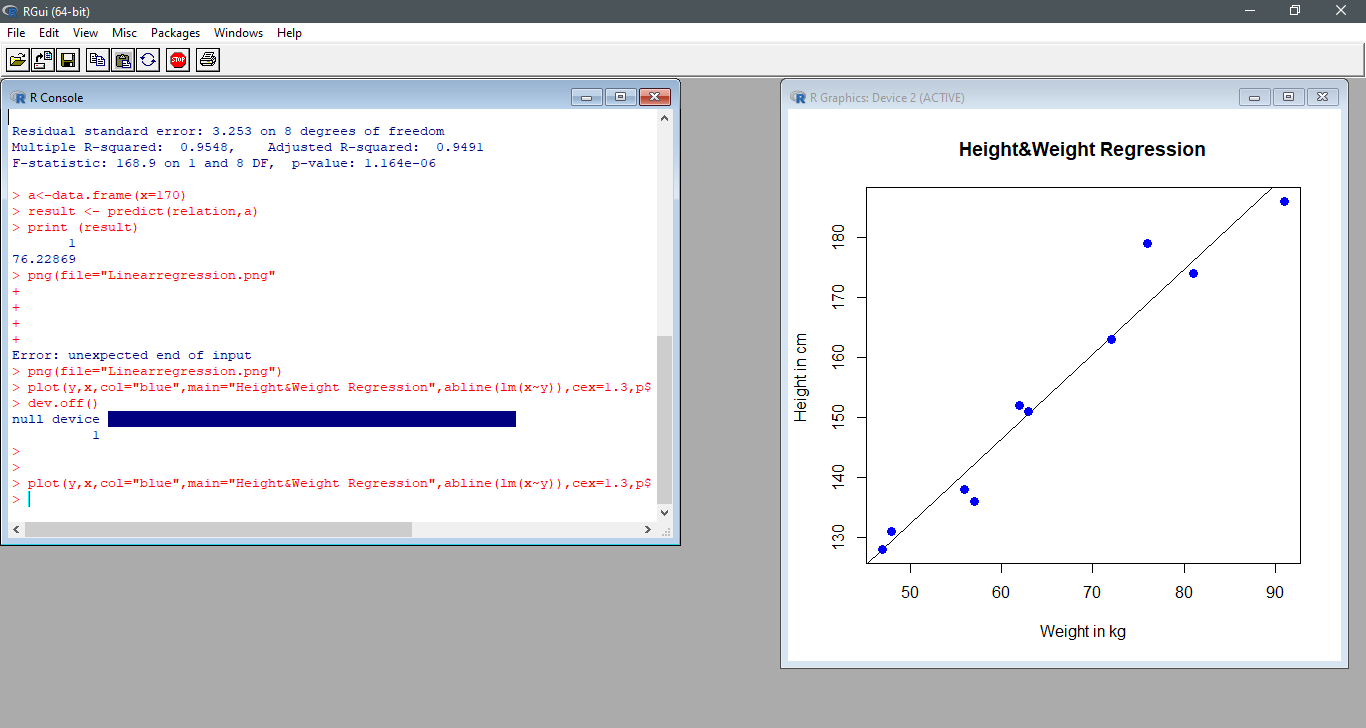


1. **Visualize the Regression Graphically**

> png(file="Linearregression.png")

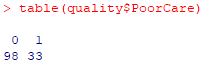
> plot(y,x,col="blue",main="Height&Weight Regression",abline(lm(x~y)),cex=1.3,pch=16,xlab="Weight in kg",ylab="Height in cm ")

> dev.off()

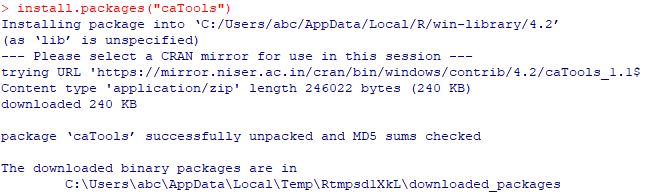


**Practical 10:-Perform the logistic regression on the given data warehouse data.**

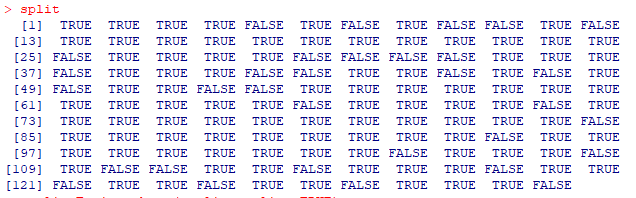
* quality<-read.csv("C:/Users/abc/Desktop/quality.csv")str(quality)
* table(quality$PoorCare)



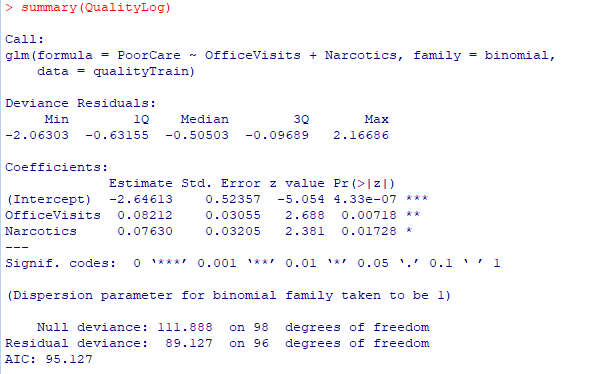
* 98/131
* 
* install.packages("caTools")



* library("caTools")
* set.seed(88)
* split=sample.split(quality$PoorCare,SplitRatio=0.75)
* split



* qualityTrain=subset(quality,split==TRUE)
* qualityTest=subset(quality,split==FALSE)
* nrow(qualityTrain)
* 
* nrow(qualityTest)
* 
* QualityLog=glm(PoorCare ~ OfficeVisits + Narcotics,data=qualityTrain,family=binomial)
* summary(QualityLog)



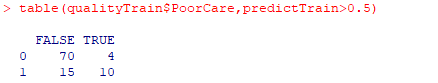
* predictTrain=predict(QualityLog,type="response")
* summary(predictTrain)



* tapply(predictTrain,qualityTrain$PoorCare,mean)



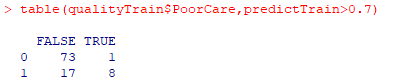
* table(qualityTrain$PoorCare,predictTrain>0.5)



* 70/74



* table(qualityTrain$PoorCare,predictTrain>0.7)



* 8/25



* 73/74



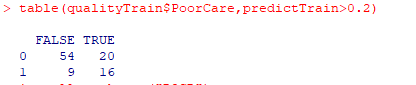
* 16/25



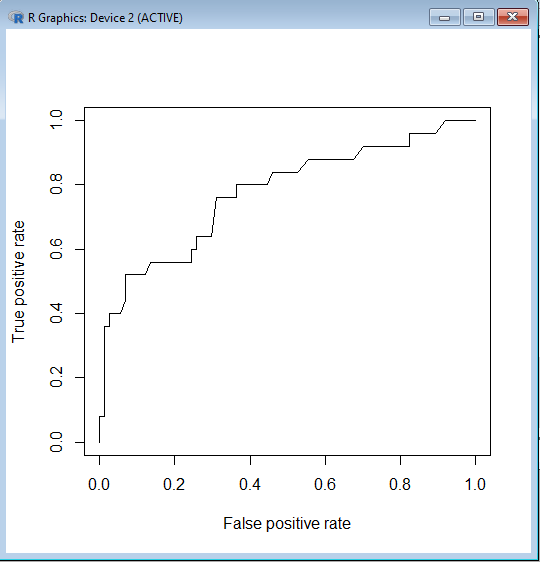
* 54/74



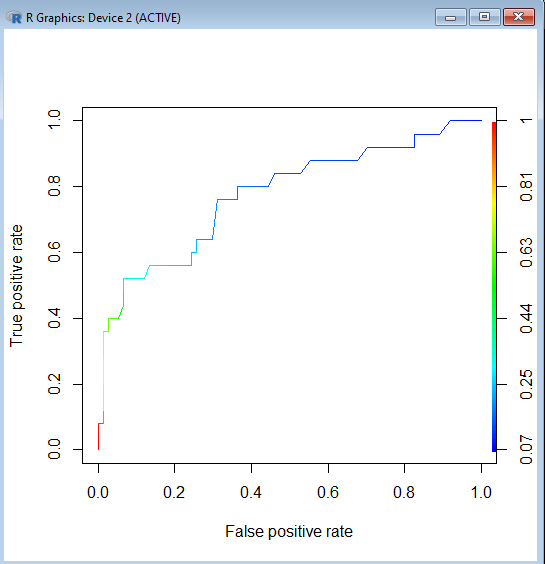
* table(qualityTrain$PoorCare,predictTrain>0.2)



* install.packages("ROCR")
* library(ROCR)
* ROCRpred=prediction(predictTrain,qualityTrain$PoorCare)
* ROCRperf=performance(ROCRpred,"tpr","fpr")
* plot(ROCRperf)



* plot(ROCRperf,colorize=TRUE)



* plot(ROCRperf,colorize=TRUE,print.cutoffs.at=seq(0.1,by=0.1),text.adj=c(-0.2,1.7))

