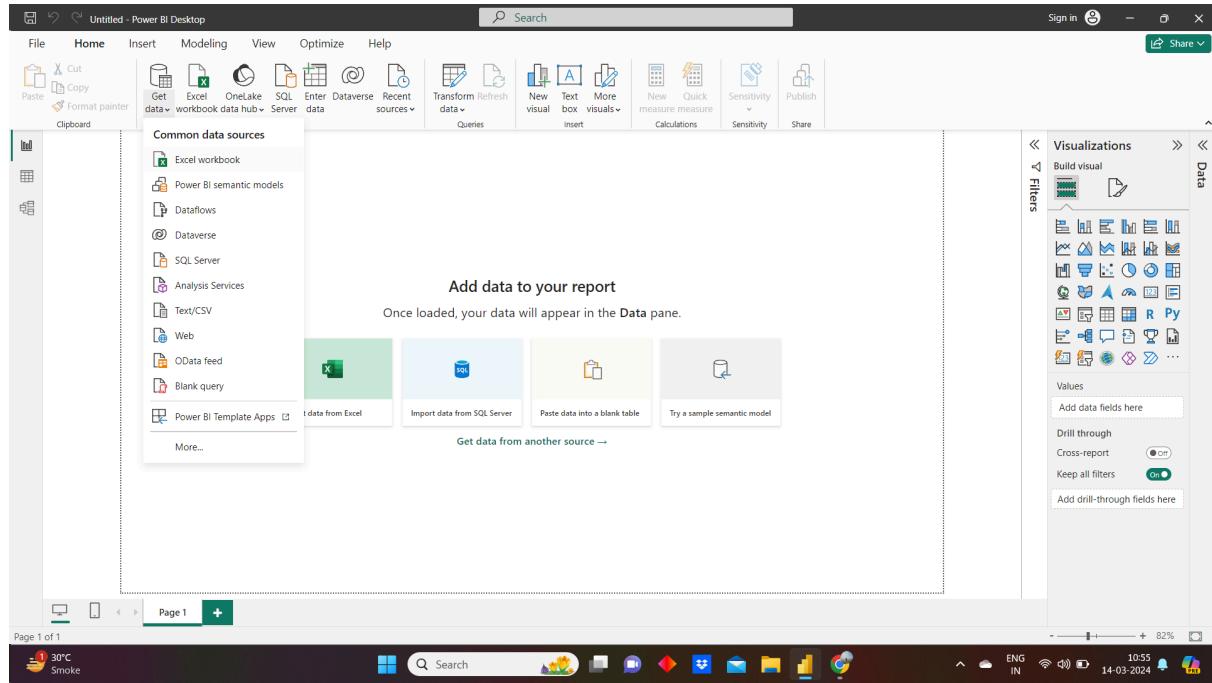


Bi journal Practical 1

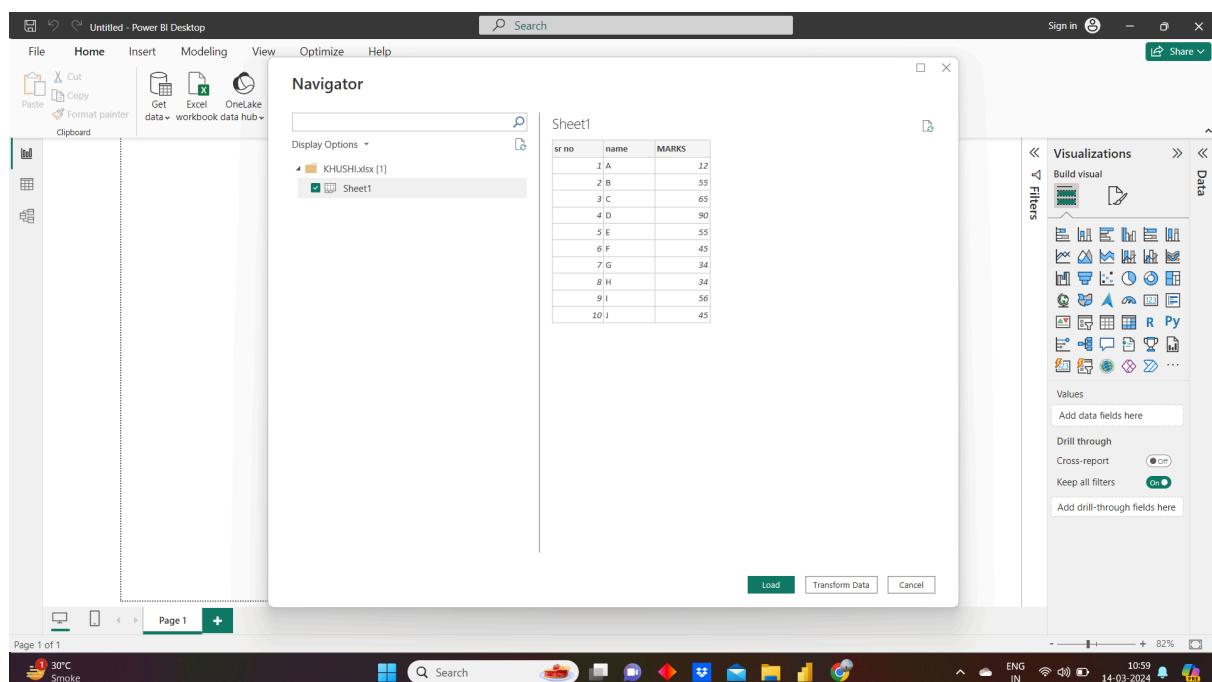
Aim :- Import the legacy data from different sources such as (excel,sql,server , oracle , etc) and load in the target system.

Step 1 :- importing Excel data to power BI

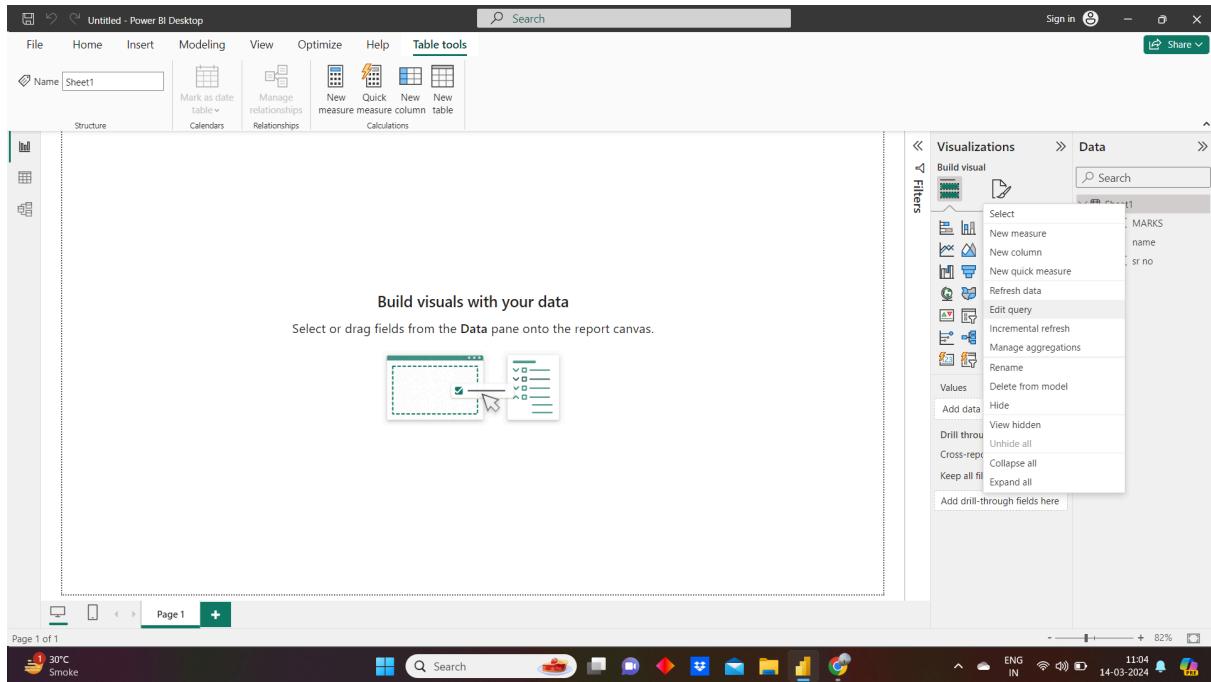
Open power BI -> click on get data -> select Excel



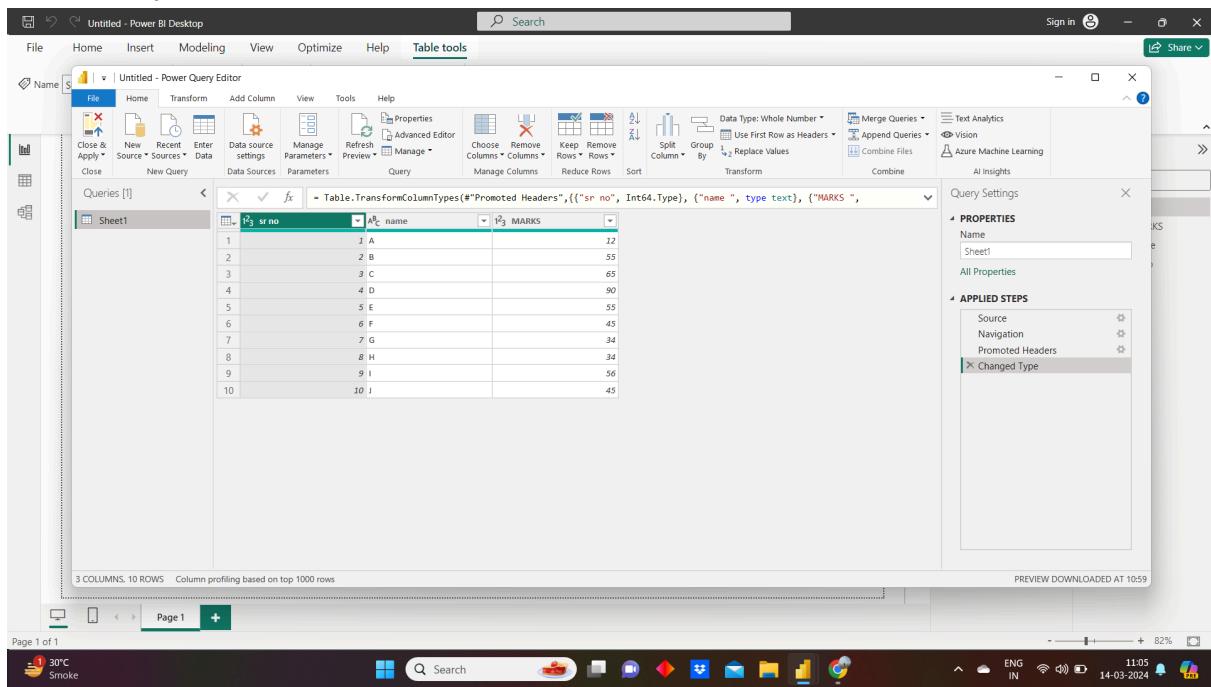
Select required file and click on open, navigator screen appears -> Select file -> load



Click on edit query

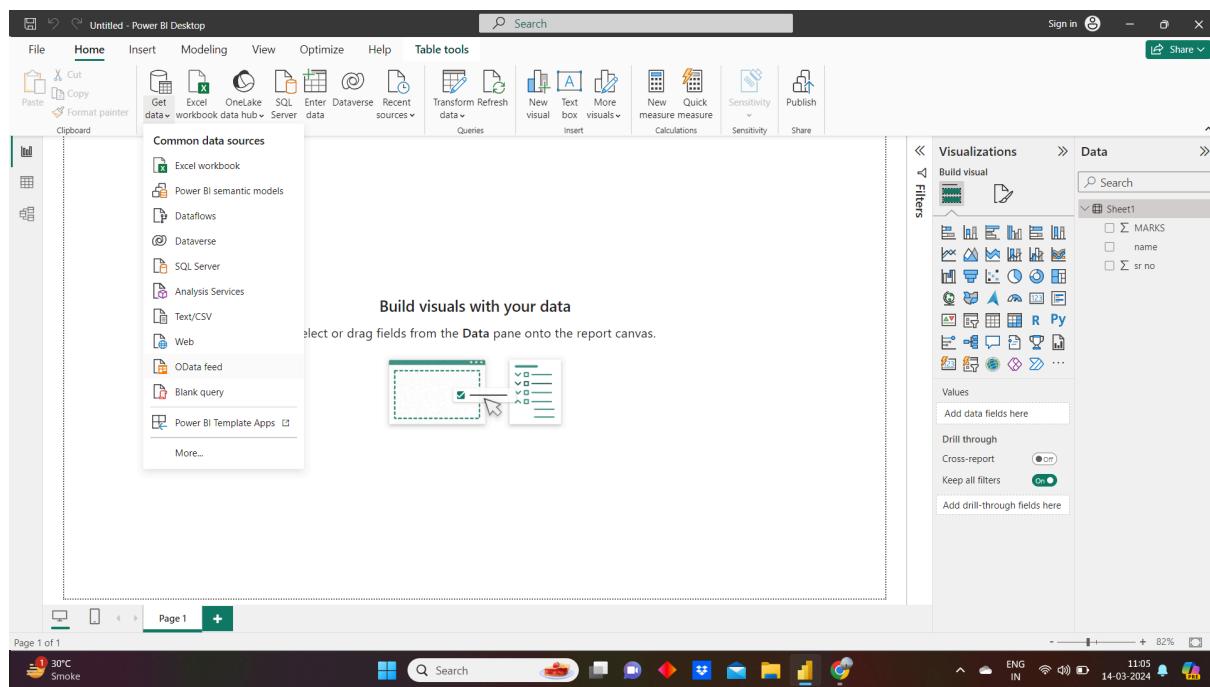


Power query editor appears

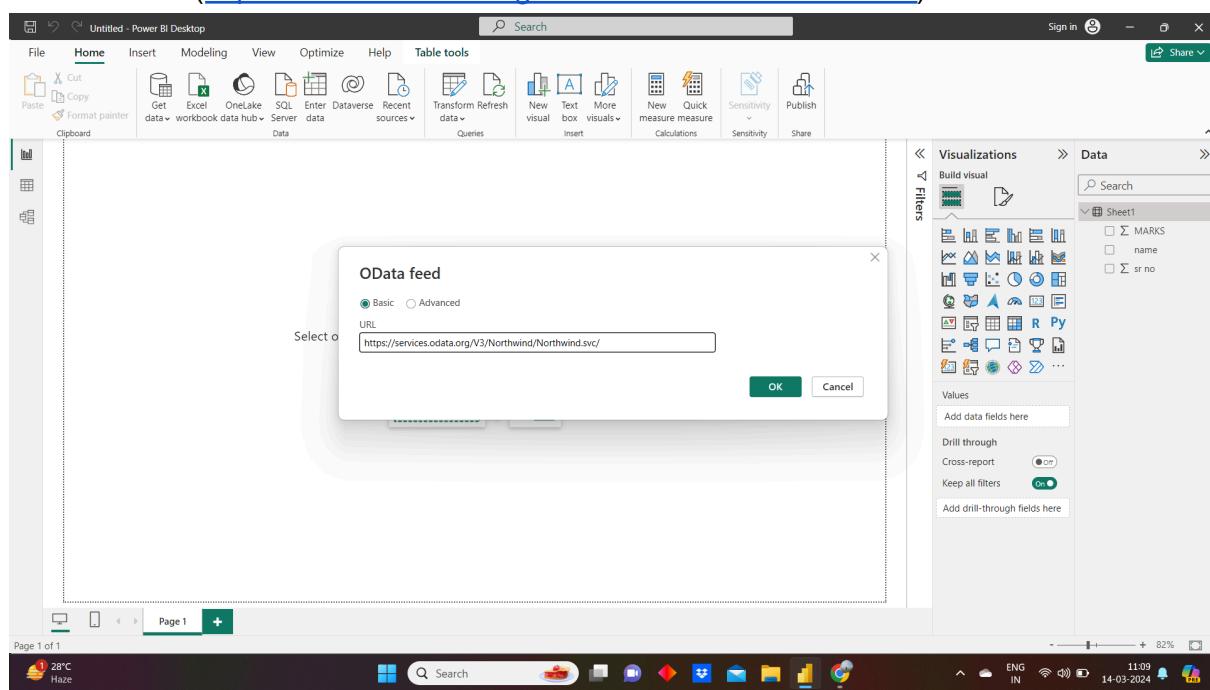


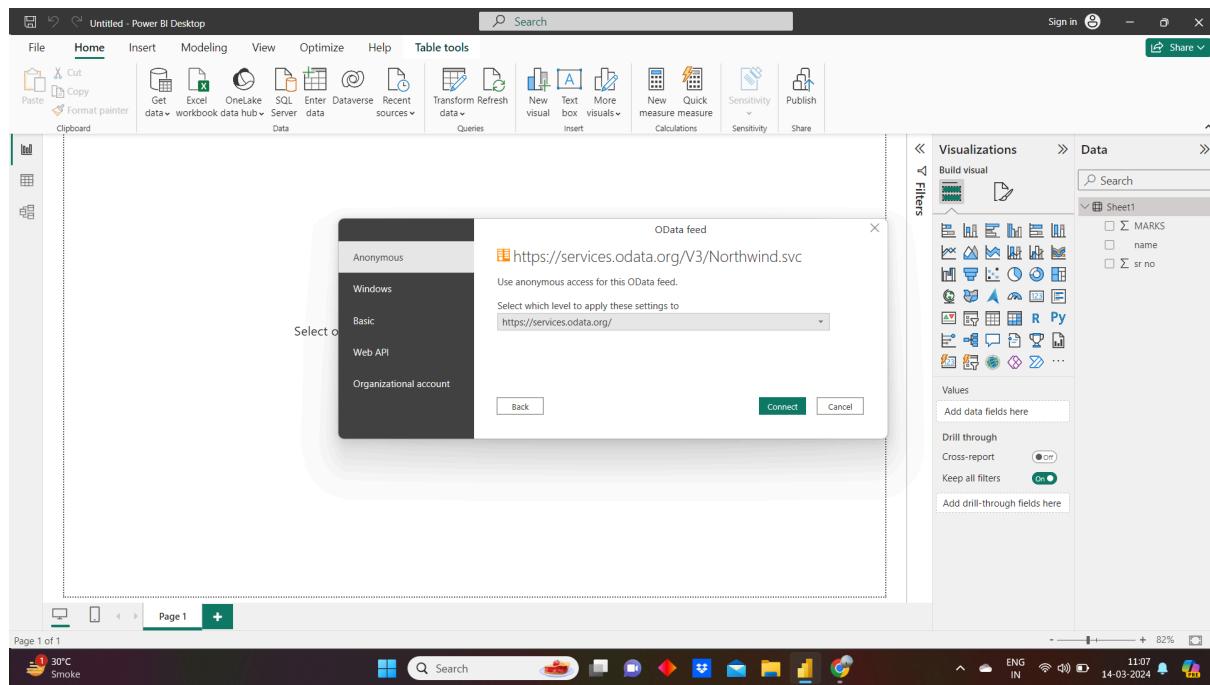
Step 2: Importing ODATA to power BI

Again go to power BI and get data -> OData feed



Paste URL as (<https://services.odata.org/V3/Northwind/Northwind.svc/>) ->click OK ->connect





Select order table -> click load

OrderID	CustomerID	EmployeeID	OrderDate	RequiredDate
10248	VINET	5	04-07-1996 00:00:00	01-08-19%
10249	TOMSP	6	05-07-1996 00:00:00	16-08-19%
10250	HANAR	4	08-07-1996 00:00:00	05-08-19%
10251	VICTE	3	08-07-1996 00:00:00	05-08-19%
10252	SUPRD	4	09-07-1996 00:00:00	06-08-19%
10253	HANAR	3	10-07-1996 00:00:00	24-07-19%
10254	CHOPS	5	11-07-1996 00:00:00	08-08-19%
10255	RICSU	9	12-07-1996 00:00:00	09-08-19%
10256	WELLI	3	15-07-1996 00:00:00	12-08-19%
10257	HILAA	4	16-07-1996 00:00:00	13-08-19%
10258	ERNSH	1	17-07-1996 00:00:00	14-08-19%
10259	CENTC	4	18-07-1996 00:00:00	15-08-19%
10260	OTTIK	4	19-07-1996 00:00:00	16-08-19%
10261	QUEDE	4	19-07-1996 00:00:00	16-08-19%
10262	RATTIC	8	22-07-1996 00:00:00	19-08-19%
10263	ERNSH	9	23-07-1996 00:00:00	20-08-19%
10264	FOLKO	6	24-07-1996 00:00:00	21-08-19%
10265	BLOPL	2	25-07-1996 00:00:00	22-08-19%
10266	WARTH	3	26-07-1996 00:00:00	06-09-19%
10267	FRANK	4	29-07-1996 00:00:00	26-08-19%
10268	GROSR	8	30-07-1996 00:00:00	27-08-19%
10269	WHITC	5	31-07-1996 00:00:00	14-08-19%
10270	WARTH	1	01-08-1996 00:00:00	29-08-19%

OUTPUT

The screenshot shows the Power BI Desktop interface with the 'Table tools' ribbon selected. The main area displays a preview of the 'Orders' table with 830 rows. The table includes columns such as OrderID, CustomerID, EmployeeID, OrderDate, RequiredDate, and ShippedDate. The preview also shows the top 20 rows in detail. On the right side, there are sections for 'PROPERTIES' (Name: Orders) and 'APPLIED STEPS' (Source, Navigation). The bottom status bar shows system information like battery level (28% Haze), network (ENG IN), and date/time (14-03-2024 11:11).

Practical 2

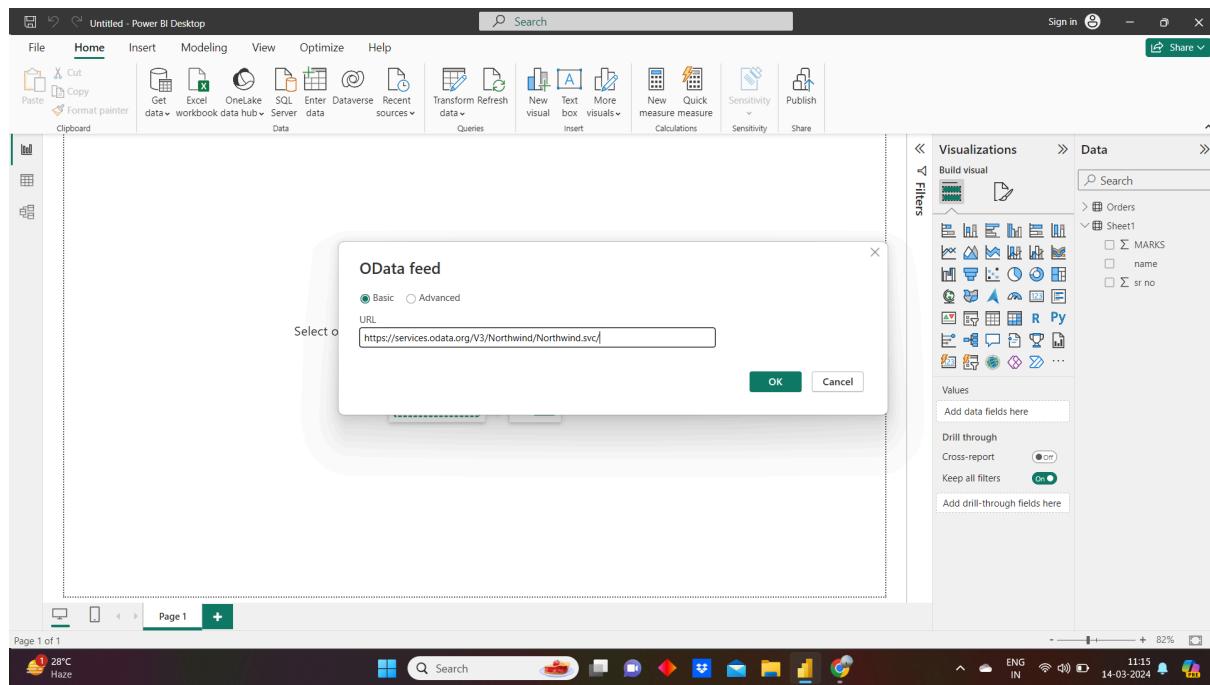
Aim : perform the extraction and transformation and loading process to construct the database in power BI.

Step 1 :ETL process in power BI

Open power bi -> click on getdata -> Odata feed

The screenshot shows the Power BI Desktop interface with the 'Home' ribbon selected. The 'Get data' button is highlighted. The left sidebar shows common data sources like Excel workbook, Power BI Semantic models, Dataflows, and OData feed. The right sidebar shows visualizations and data fields. The bottom status bar shows system information like battery level (28% Haze), network (ENG IN), and date/time (14-03-2024 11:14).

Paste URL as (<https://services.odata.org/V3/Northwind/Northwind.svc/>) ->click OK ->connect



Click on check box of products table and then load

ProductID	ProductName	SupplierID	CategoryID	QuantityPerUnit
1	Chai	1	1	10
2	Chang	1	1	12
3	Aniseed Syrup	1	2	21
4	Chef Anton's Cajun Seasoning	2	2	24
5	Chef Anton's Gumbo Mix	2	2	23
6	Grandma's Boysenberry Spread	3	2	21
7	Uncle Bob's Organic Dried Pears	3	7	71
8	Northwoods Cranberry Sauce	3	2	21
9	Mishi Kobe Niku	4	6	1
10	Iura	4	8	1
11	Queso Cabrales	5	4	41
12	Queso Manchego La Pastora	5	4	41
13	Konbu	6	8	2
14	Tofu	6	7	4
15	Genen Shouyu	6	2	2
16	Pavlova	7	3	3
17	Alice Mutton	7	6	2
18	Carnarvon Tigers	7	8	1
19	Teatime Chocolate Biscuits	8	3	1
20	Sir Rodney's Marmalade	8	3	3
21	Sir Rodney's Scones	8	3	2
22	Gustaf's Kräckebrot	9	5	2
23	Tumbrold	9	5	1

Remove other columns to only display columns of interest in Query editor, select productid , productname, quantityperunit and unitinstock columns.

Screenshot of the Power Query Editor showing the 'Products' table. The table has 13 columns and 77 rows. The columns are: ProductID, ProductName, SupplierID, CategoryID, QuantityPerUnit, UnitPrice, and UnitsInStock. The 'UnitsInStock' column is currently set to 'Whole Number'. The 'CategoryID' column is highlighted.

Change the data type of UnitInStock column
Select the UnitInStock column

Screenshot of the Power Query Editor showing the 'Products' table. The 'UnitInStock' column is selected and a context menu is open. The 'Change Type' option is selected, showing a dropdown menu with options like Decimal Number, Fixed decimal number, Whole Number, Percentage, and others.

Expand the orders table
Once you have loaded a data source you can click on recent sources to select desirable table (orders)

File Home Transform Add Column View Tools Help

Close & Apply New Source Recent Sources Enter Data Data source settings Manage Parameters Refresh Preview Advanced Editor Properties Choose Columns Remove Columns Keep Rows Remove Rows Split Column Group By Sort Data Type: Whole Number Merge Queries Use First Row as Headers Text Analytics Vision Combine Files AI Insights

Queries [3]

Sheet1 Orders Products

KHUSHI.xlsx

More...

	productName	QuantityPerUnit	UnitsInStock
4	10 boxes x 20 bags	39	
5	24 - 12 oz bottles	17	
6	12 - 550 ml bottles	13	
7	48 - 6 oz jars	53	
8	36 boxes	0	
9	12 - 8 oz jars	120	
10	12 - 1 lb pkgs.	15	
11	12 - 12 oz jars	6	
12	18 - 500 g pkgs.	29	
13	12 - 200 ml jars	31	
14	1 kg pkgs.	22	
15	10 - 500 g pkgs.	86	
16	2 kg box	24	
17	40 - 100 g pkgs.	35	
18	24 - 250 ml bottles	39	
19	32 - 500 g boxes	29	
20	20 - 1 kg tins	0	
21	16 kg pkgs.	42	
22	10 boxes x 12 pieces	25	
23	30 gift boxes	40	
24	24 pkgs x 4 pieces	3	
25	24 - 500 g pkgs.	104	
26	12 - 250 g pkgs.	61	
27	12 - 355 ml cans	20	
28	20 - 450 g glasses	76	
29	100 - 250 g bags	15	
30	100 - 100 g pieces	49	
31	25 - 825 g cans	26	

4 COLUMNS, 77 ROWS Column profiling based on top 1000 rows

28°C Haze

Search

ENG IN 11:20 14-03-2024

PREVIEW DOWNLOADED AT 11:09

Query Settings

Properties Name: Products All Properties

Applied Steps Source Navigation Removed Columns Changed Type

After selecting the URL ,Navigator window will appear from which you can select orders table -> click on load

File Home Transform Add Column View Tools Help

CLOSE & APPLY NEW SOURCE * SOURCES NEW QUERY
CLOSE
Data sources Data Sources Parameters
Queries [3] Sheet1 Orders Products
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

ProductID
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28

Alphabetical_list_of_products
Categories
Category_Sales_for_1997
Current_Product_Lists
Customer_and_Suppliers_by_Cities
CustomerDemographics
Customers
Employees
Invoices
Order_Details
Order_Details_Extendeds
Order_Subtotals
Orders
Orders_Qries
Product_Sales_for_1997
Products
Products_Above_Average_Prices
Products_by_Categories
Regions

Navigator

Display Options

Orders

OrderID	CustomerID	EmployeeID	OrderDate	RequiredDate
10248	VINET	5	04-07-1996 00:00:00	01-08-1995
10249	TOMSP	6	05-07-1996 00:00:00	16-08-1995
10250	HANAR	4	08-07-1996 00:00:00	05-08-1995
10251	VICTE	3	08-07-1996 00:00:00	05-08-1995
10252	SUPRD	4	09-07-1996 00:00:00	06-08-1995
10253	HANAR	3	10-07-1996 00:00:00	24-07-1995
10254	CHOPS	5	11-07-1996 00:00:00	08-08-1995
10255	RICSU	9	12-07-1996 00:00:00	09-08-1995
10256	WELLU	3	15-07-1996 00:00:00	12-08-1995
10257	HILAA	4	16-07-1996 00:00:00	13-08-1995
10258	ERNSH	1	17-07-1996 00:00:00	14-08-1995
10259	CENTC	4	18-07-1996 00:00:00	15-08-1995
10260	OTTIK	4	19-07-1996 00:00:00	16-08-1995
10261	QUEDE	4	19-07-1996 00:00:00	16-08-1995
10262	RATIC	8	22-07-1996 00:00:00	19-08-1995
10263	ERNSH	9	23-07-1996 00:00:00	20-08-1995
10264	FOLKO	6	24-07-1996 00:00:00	21-08-1995
10265	BLONP	2	25-07-1996 00:00:00	22-08-1995
10266	WARTH	3	26-07-1996 00:00:00	06-09-1995
10267	FRANK	4	29-07-1996 00:00:00	26-08-1995
10268	GROSR	8	30-07-1996 00:00:00	27-08-1995
10269	WHITC	5	31-07-1996 00:00:00	14-08-1995
10270	WARTH	1	01-08-1996 00:00:00	29-08-1995

OK Cancel

4 COLUMNS, 77 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 11:09

28°C Haze

Search

ENG IN 11:20 14-03-2024

Query editor window will appear

In query view scroll to Order Details column

In the Orders Details column, select the expand icon.

In the expand drop-down:

Select (select all columns) to clear all columns

Select ProductID, UnitPrice and quantity

Select File
Click OK

Screenshot of Power Query Editor showing the 'Transform' tab selected. A context menu is open over the 'Customer' column, specifically for the 'ShipCountry' column. The menu includes options like 'Expand', 'Aggregate', and checkboxes for columns: OrderID, ProductID, UnitPrice, Quantity, Discount, Order, and Product. The 'Quantity' checkbox is checked. The 'OK' button is visible at the bottom right of the menu.

Screenshot of Power Query Editor showing the 'Transform' tab selected. A context menu is open over the 'Customer' column, specifically for the 'ShipCountry' column. The menu includes options like 'Expand', 'Aggregate', and checkboxes for columns: OrderID, ProductID, UnitPrice, Quantity, Discount, Order, and Product. The 'Quantity' checkbox is checked. The 'OK' button is visible at the bottom right of the menu.

After clicking on OK following screen appears with combined columns

Queries [3]

Sheet1

Orders

Products

Order_Details.ProductID

Order_Details.UnitPrice

Order_Details.Quantity

Shippers

20 COLUMNS, 199+ ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 11:09

Calculate the lines total for each Order_Details row

In the add column ribbon tab, click add custom column

In the custom column dialog box, in the custom column formula text box, enter [order_details.unitprice]*[order_details.quantity] by selecting from available columns and click on insert for each column

In the new column name , enter LineTotal

Click OK

File Home Transform Add Column View Tools Help

Column From Examples Custom Function General

Conditional Column Index Column Duplicate Column

Merge Columns ABC Extract - abc Parse -

Format From Text From Number From Date & Time

Trigonometry Statistics Standard Scientific Rounding - information -

Date Time Duration -

Text Analytics Vision Azure Machine Learning AI Insights

Queries [3]

Sheet1

Orders

Products

Order_Details.ProductID

Order_Details.UnitPrice

Order_Details.Quantity

Shippers

20 COLUMNS, 199+ ROWS Column profiling based on top 1000 rows

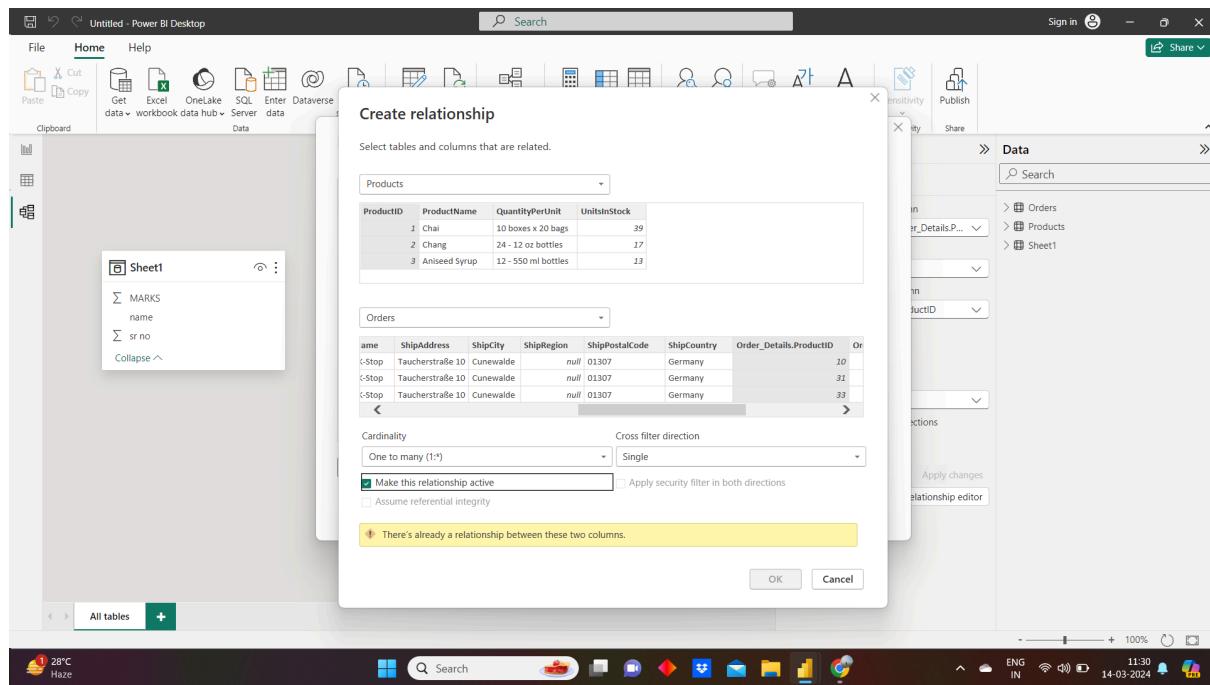
PREVIEW DOWNLOADED AT 11:09

The screenshot shows the Power Query Editor interface. A 'Custom Column' dialog box is open, prompting for a new column name ('LineTotal') and a formula ('= [Order_Details.UnitPrice]*[Order_Details.Quantity]'). The formula bar also shows the full formula: '= [Order_Details.UnitPrice]*[Order_Details.Quantity]'. Below the formula, there's a list of 'Available columns' including ShipCountry, Customer, Employee, Order_Details.ProductID, Order_Details.UnitPrice, Order_Details.Quantity, and Shipper. The main table view shows data from the 'Orders' table, including columns like OrderID, CustomerID, EmployeeID, and Order_Details. The status bar at the bottom indicates '20 COLUMNS, 199+ ROWS' and 'Column profiling based on top 1000 rows'.

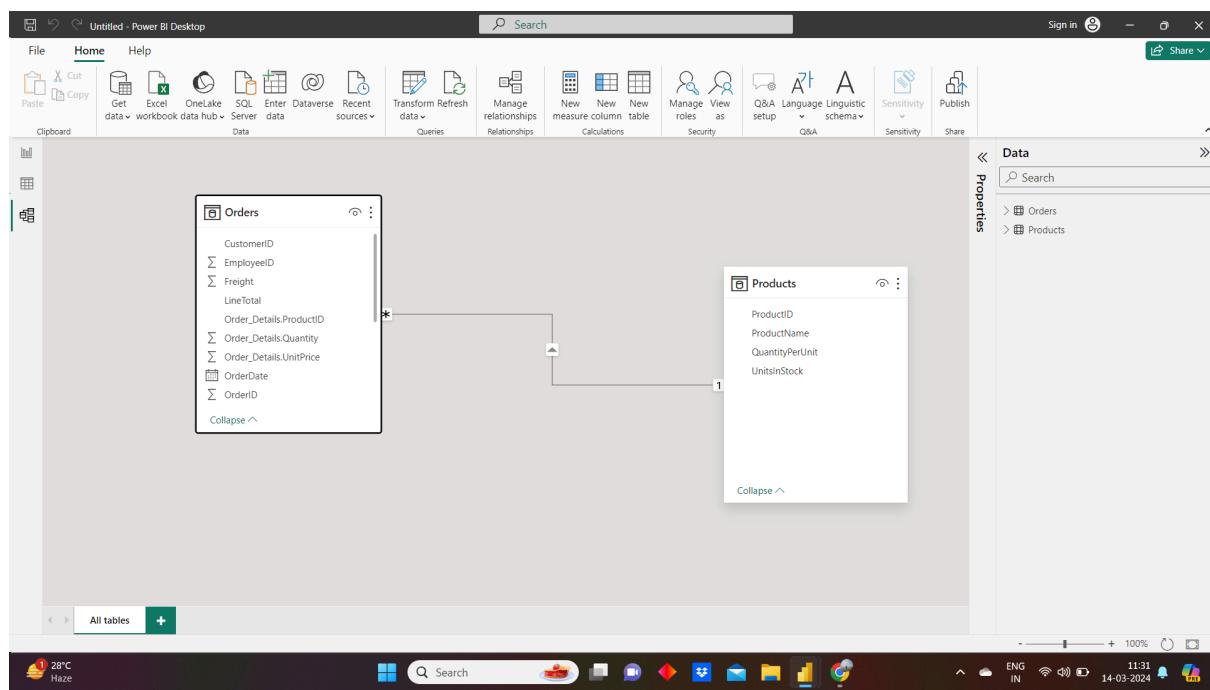
Once the data is loaded, select the manage relationship button home ribbon from untitled -> power bi desktop window
 Click on model -> manage relationships

The screenshot shows the Power BI Desktop interface with the 'Home' ribbon selected. In the center, two tables are listed: 'Orders' and 'Products'. The 'Orders' table contains columns like CustomerID, EmployeeID, LineTotal, Order_Details.ProductID, Order_Details.Quantity, Order_Details.UnitPrice, and OrderDate. The 'Products' table contains columns like ProductID, ProductName, QuantityPerUnit, and UnitsInStock. On the right side, the 'Relationships' pane is open, showing a relationship between 'Orders' and 'Products' based on 'Order_Details.ProductID' and 'ProductID'. The cardinality is set to 'Many to one (*:1)'. The 'Make this relationship active' toggle is turned on. The status bar at the bottom indicates '11:27 14-03-2024'.

Click on new ->create relationship -> ok

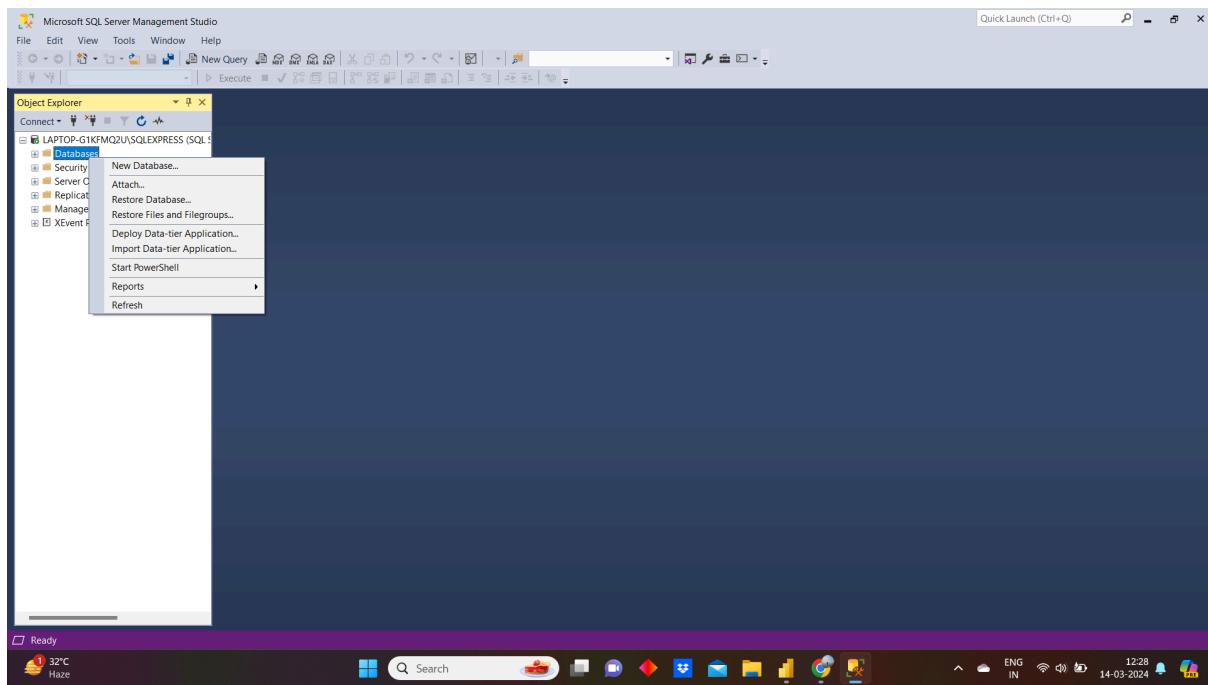


Output

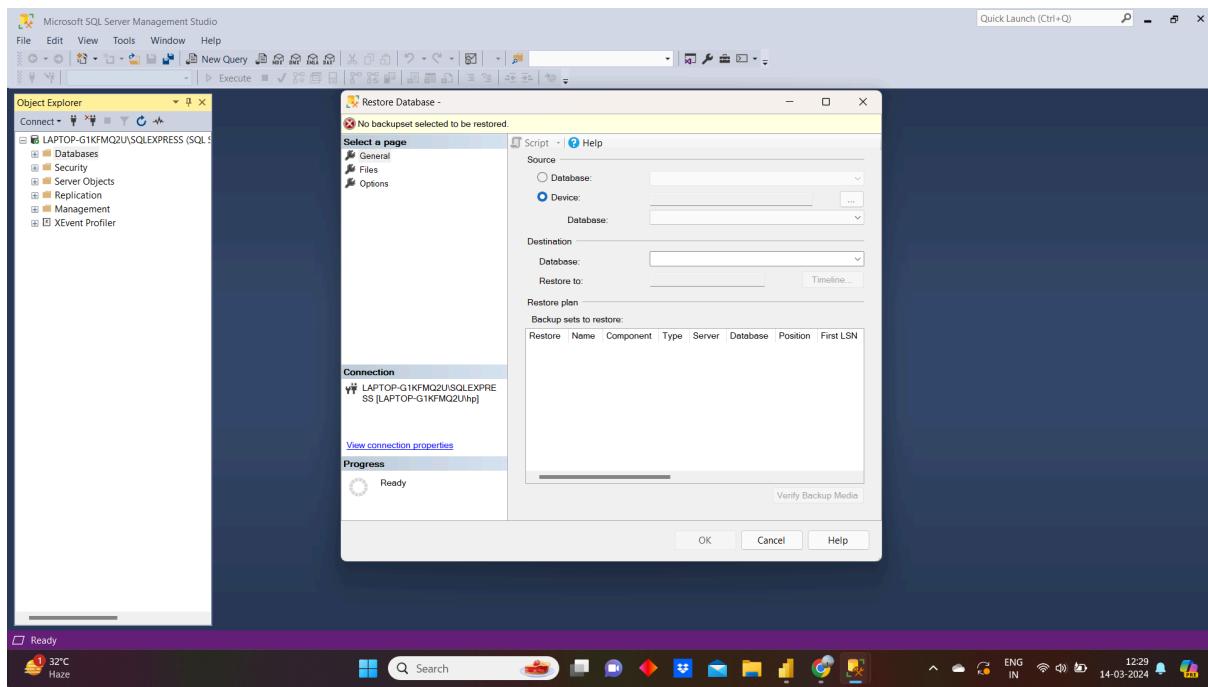


2B) ETL PROCESS IN SQL SERVER

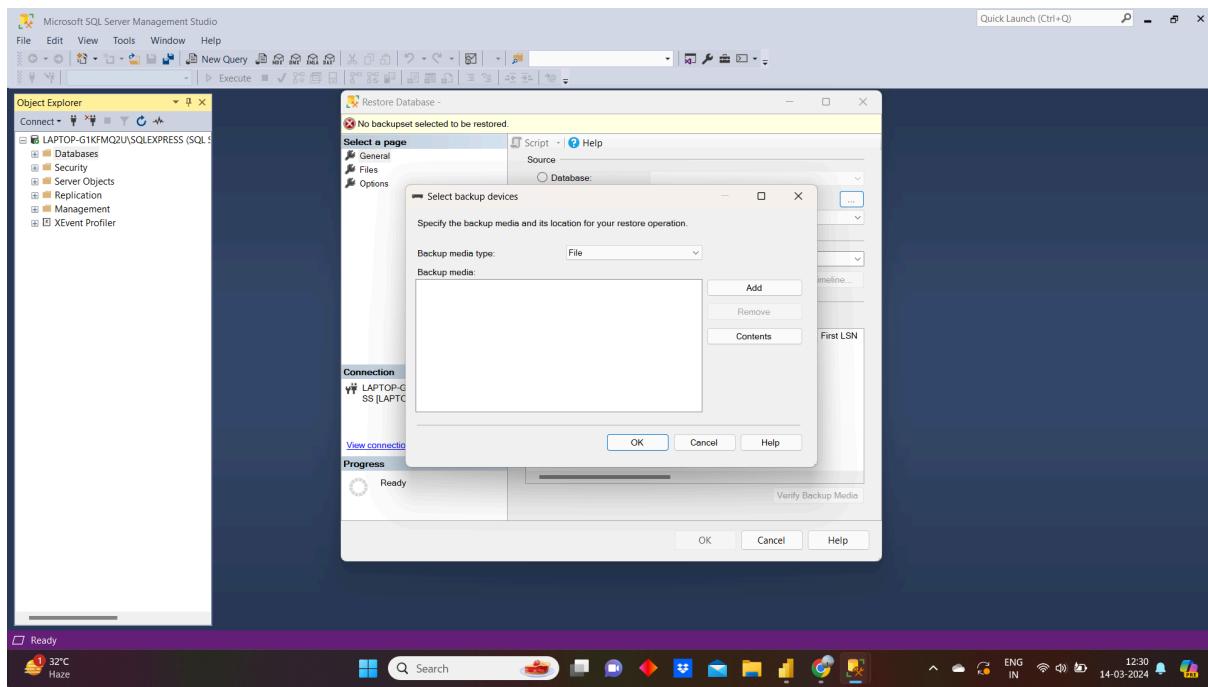
1. Open sql server management studio to restore backup file
2. Right click on databases -> restore database



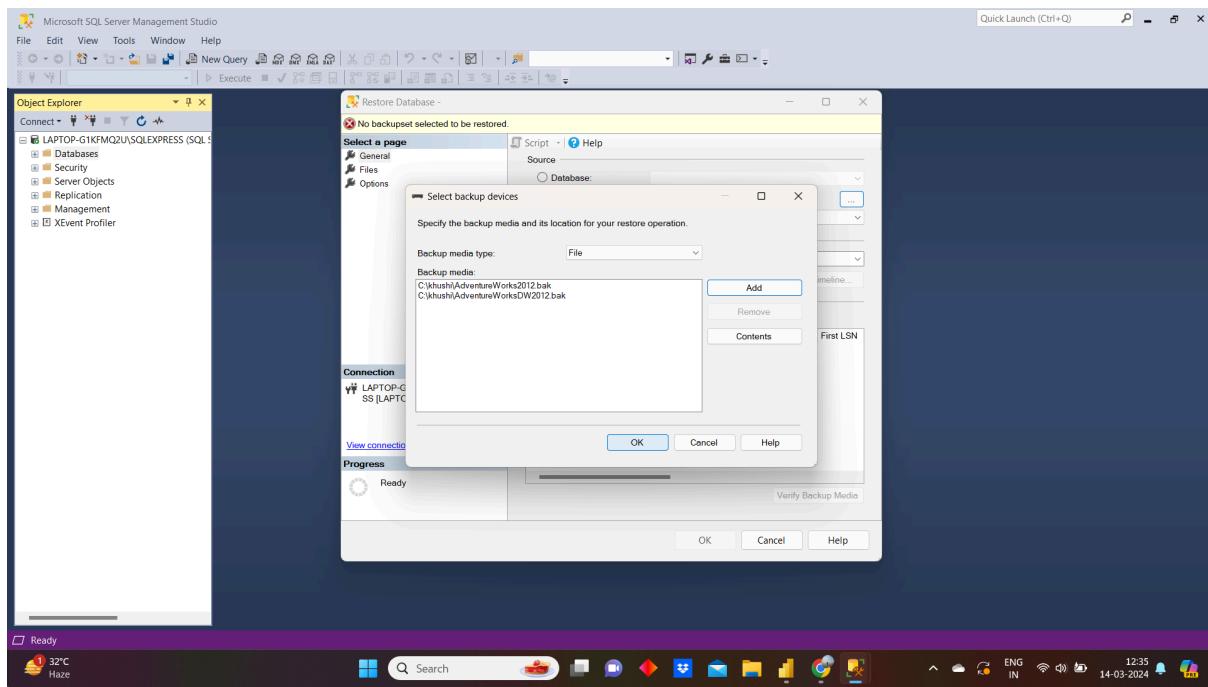
3. Select device -> click on ... icon towards end of device box

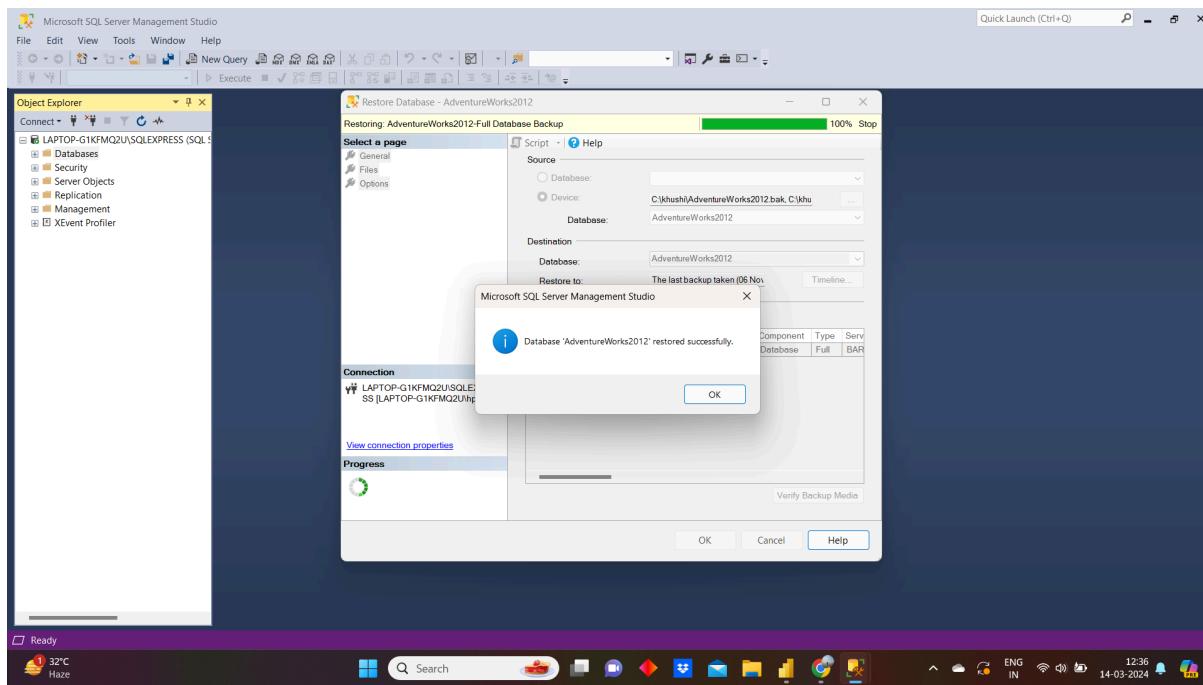


4. Click on Add -> select on folder where you have downloaded your backup files -> Click ok

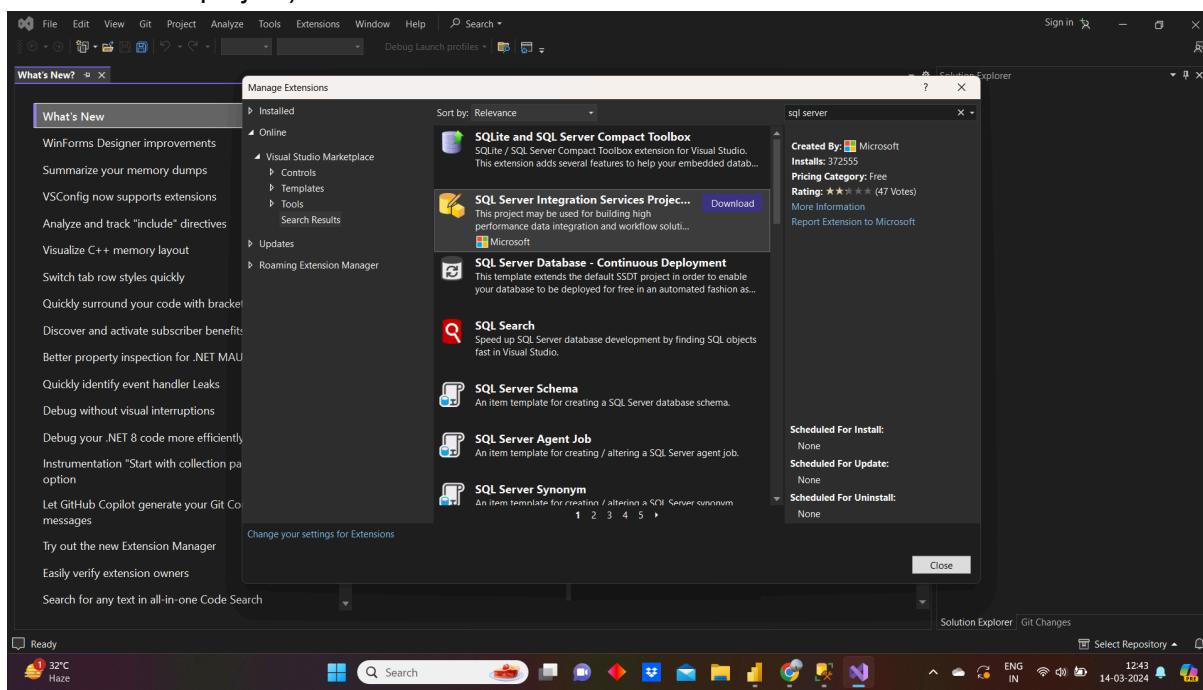


Click ok

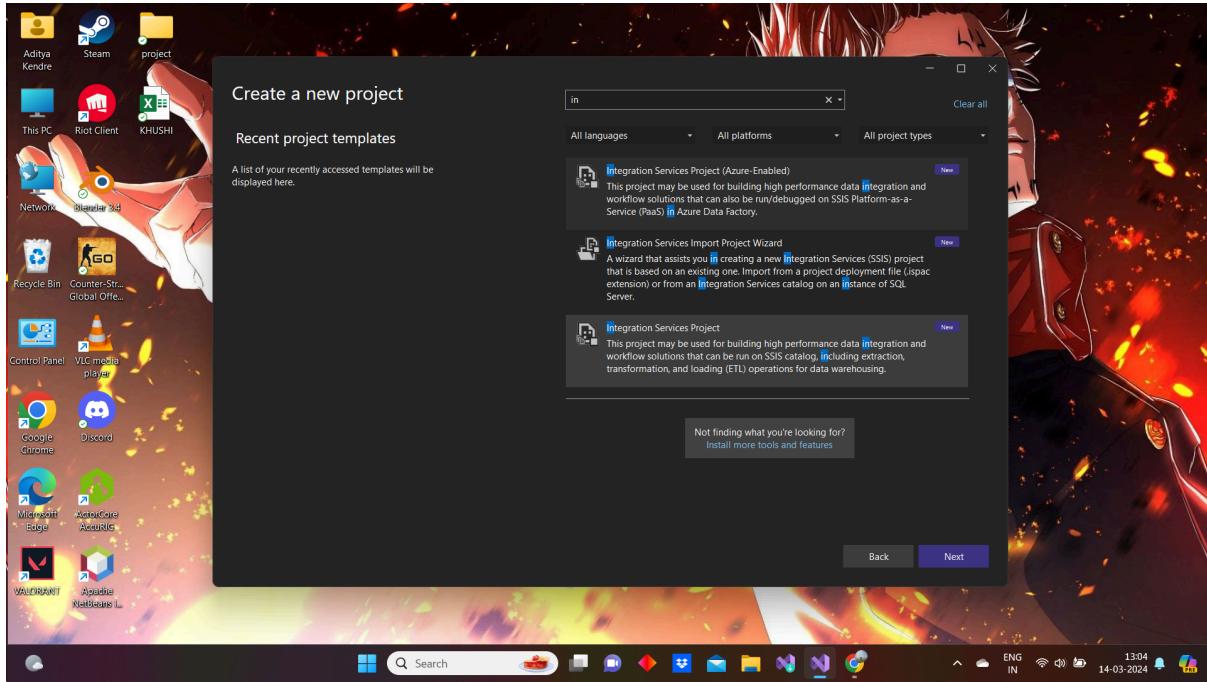




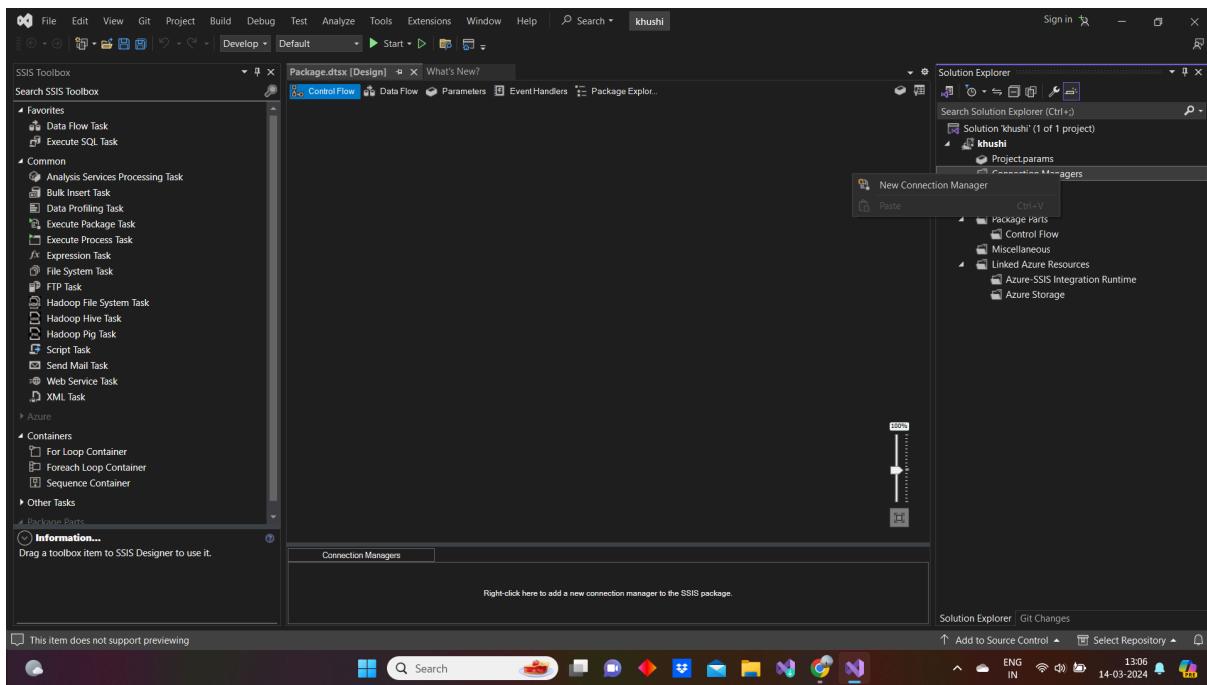
5. Open visual studio 2022 -> download & install extension (sql server integration service project)



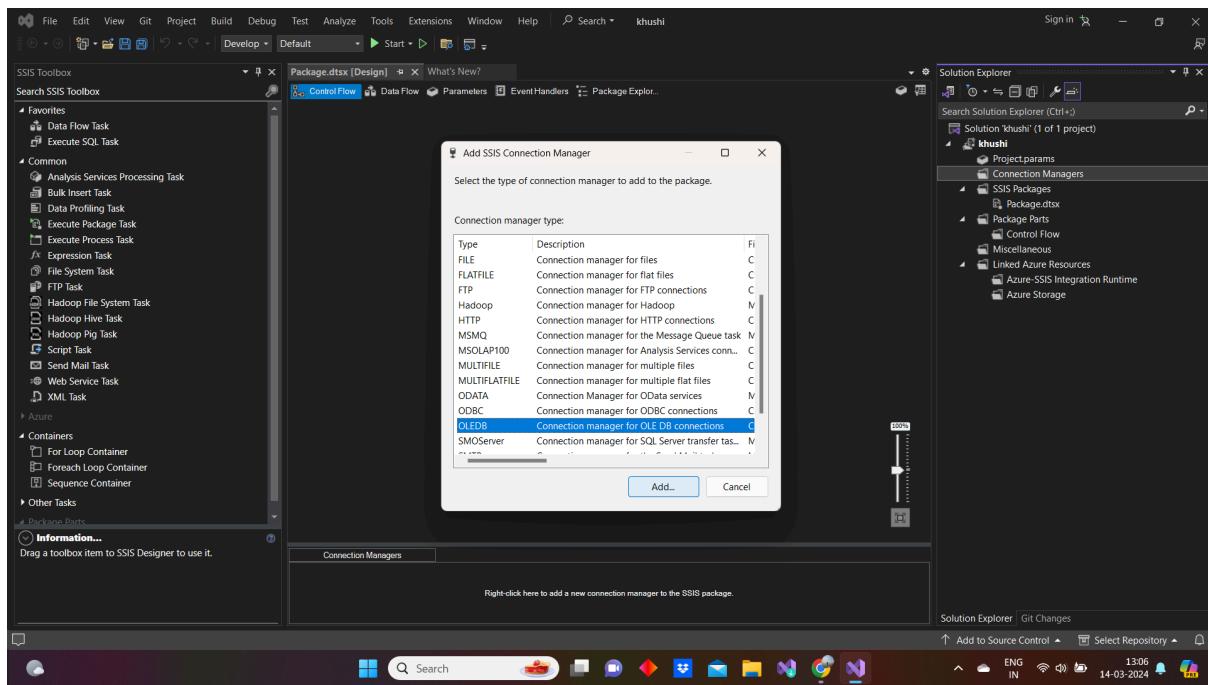
6. Open visual studio -> create new project -> integration service project & project name



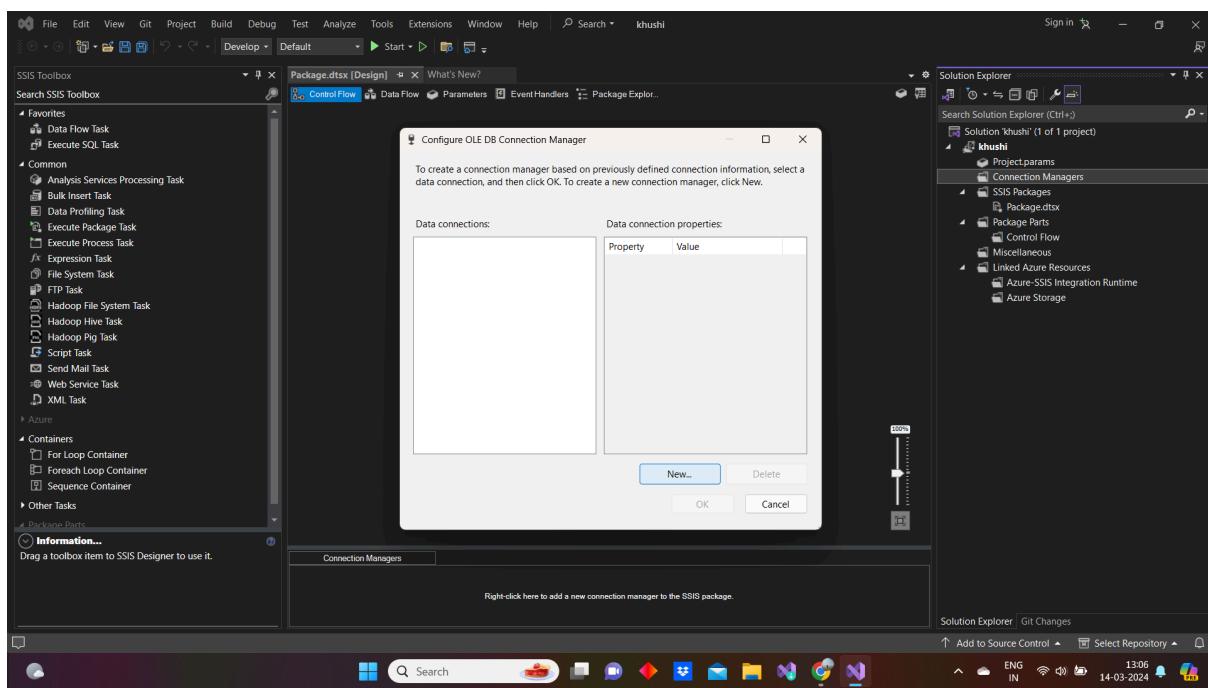
7. Right click on connection manager in solution explorer and click on new manager.

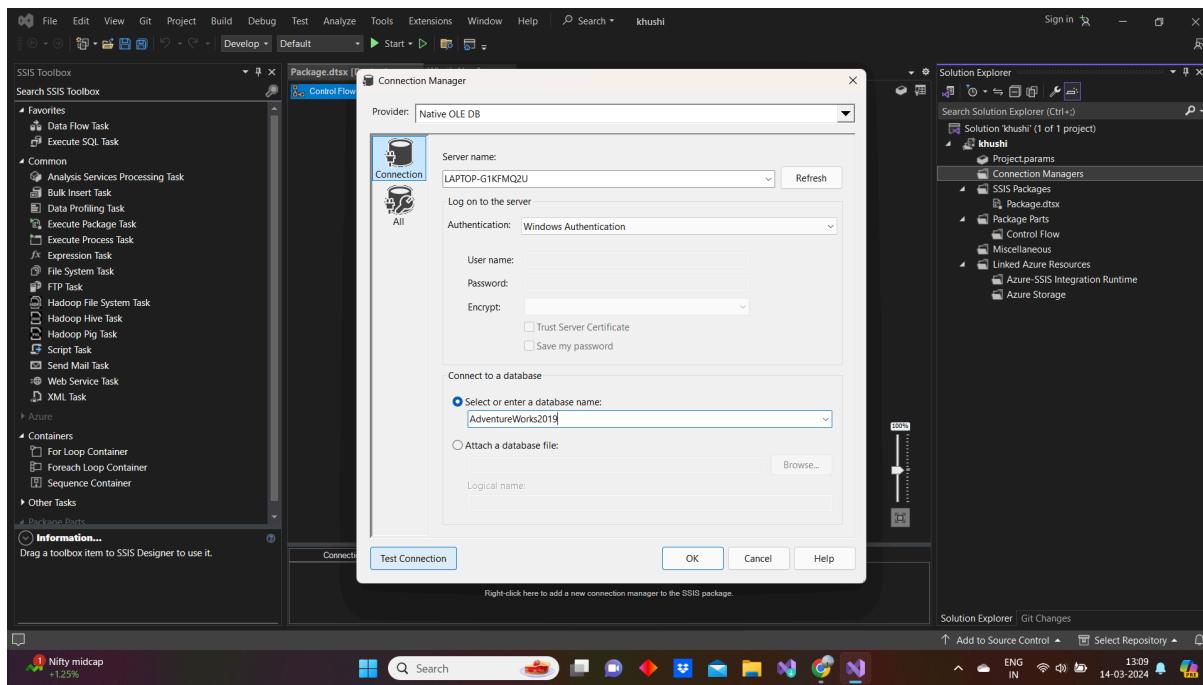


8. Configure LOEB connection manager window appears -> click on add

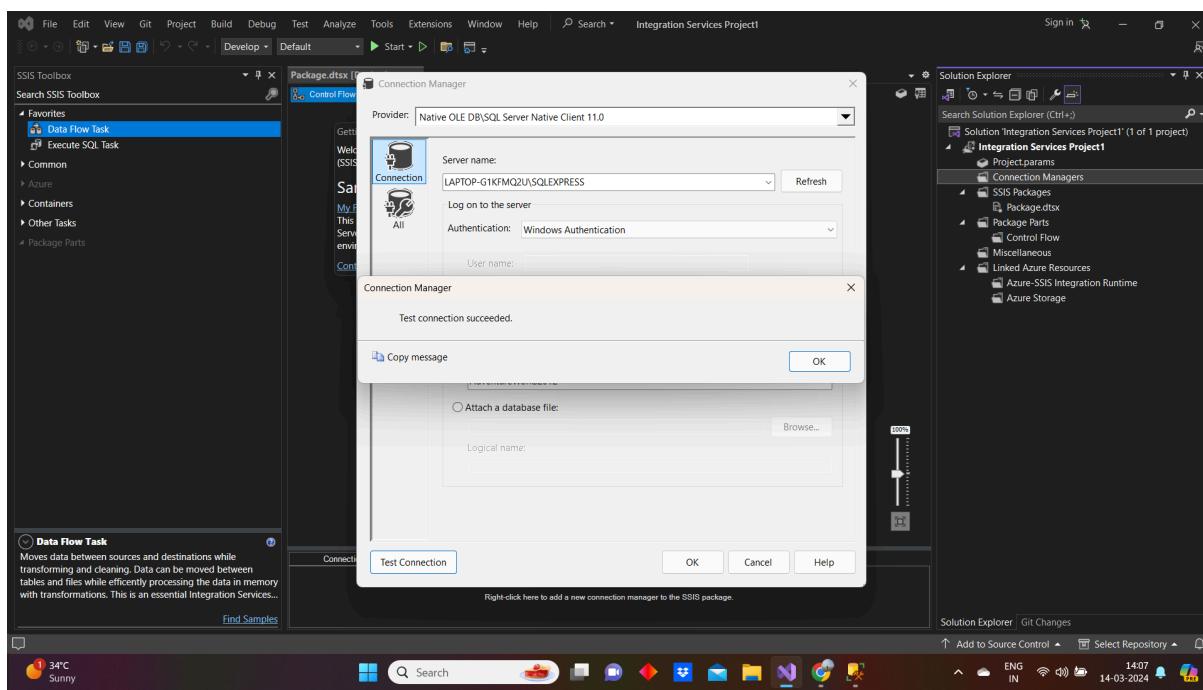


Click on new

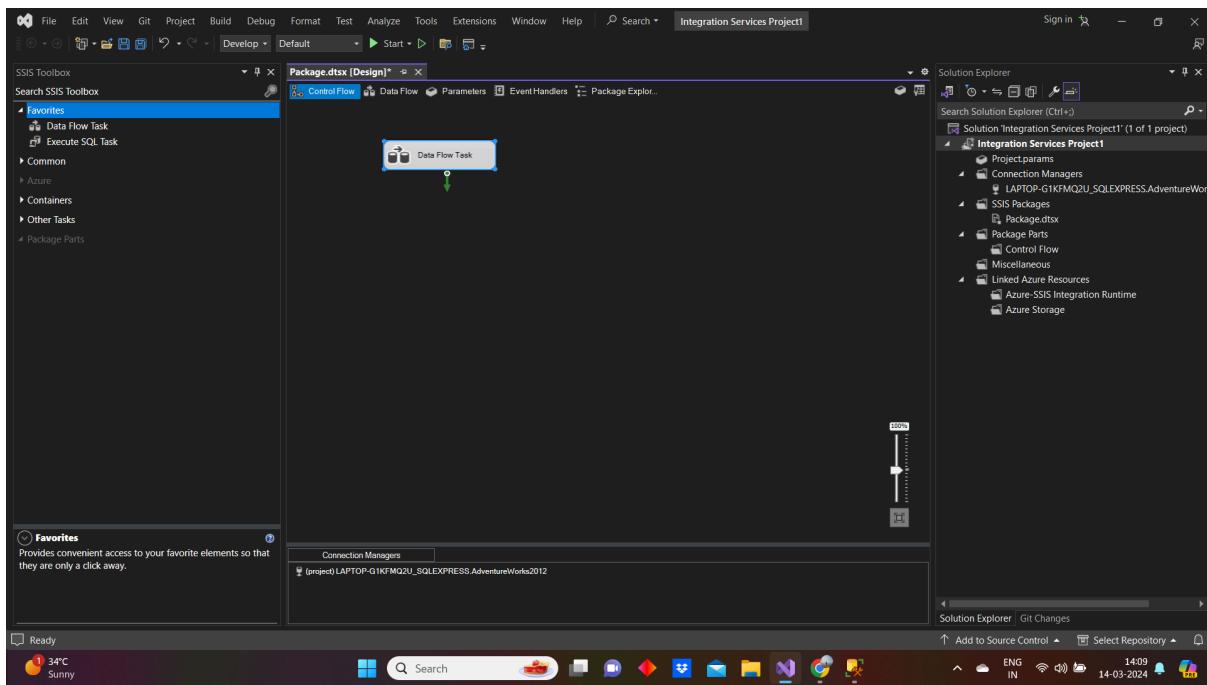




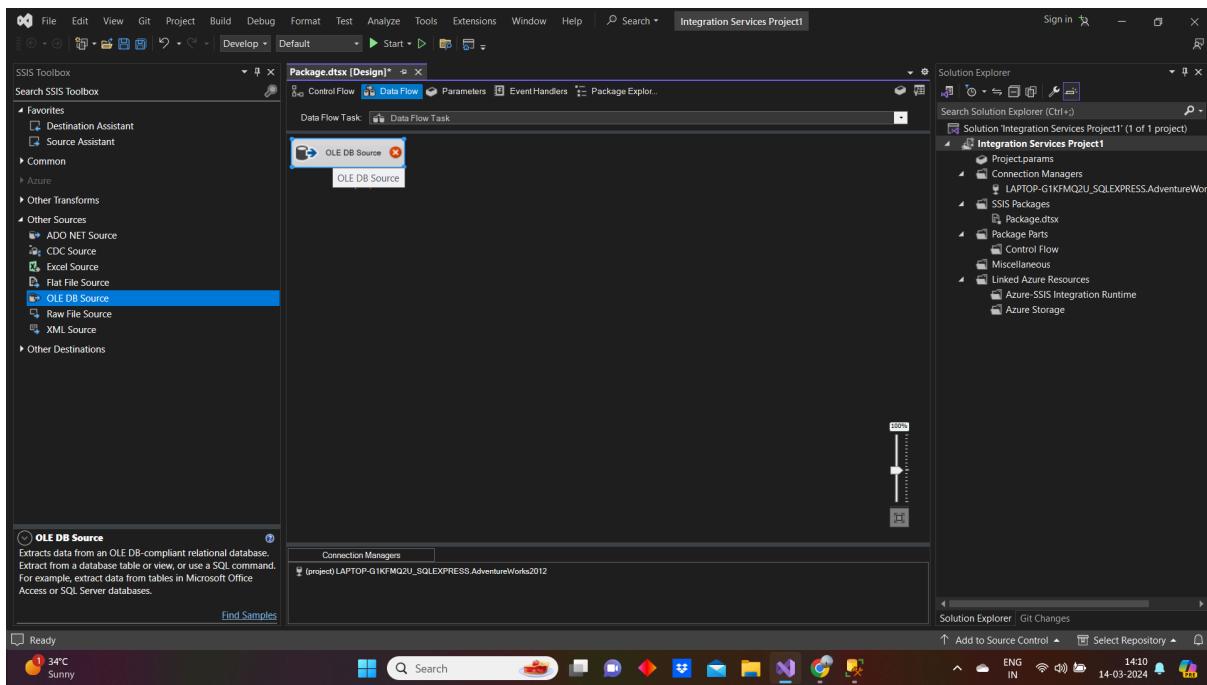
9. Select server name from the drop down and database name and click on test connection.



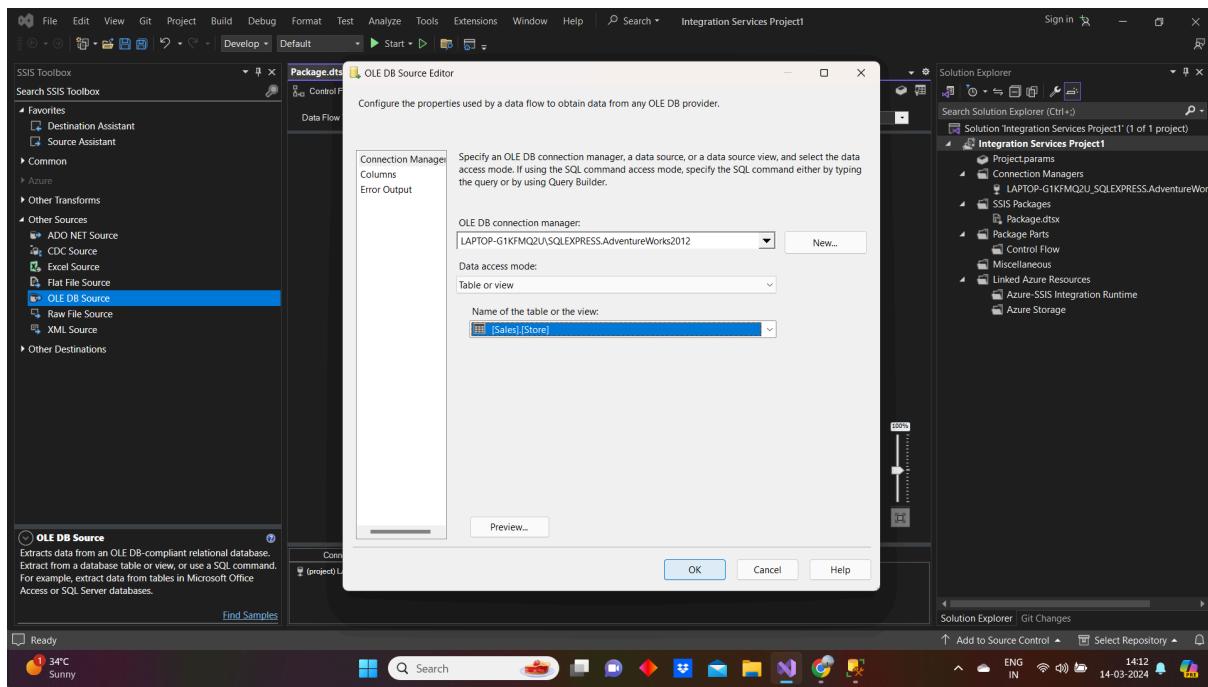
10. Drag and drop data flow task in control flow tab



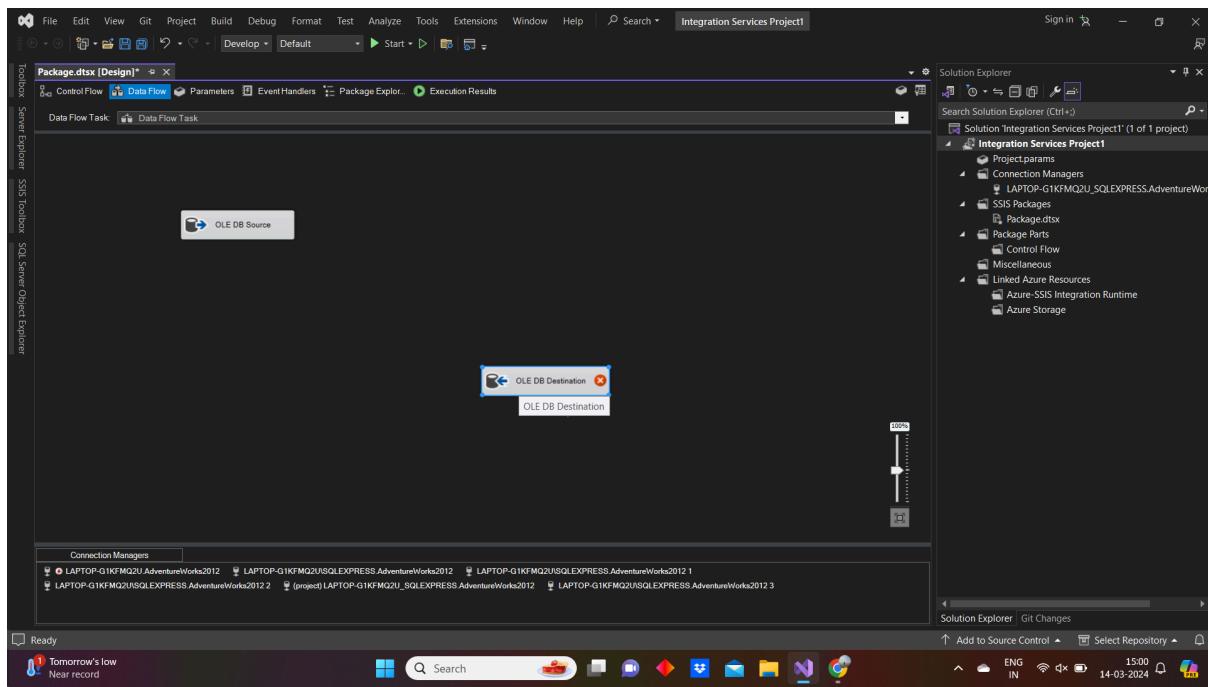
11. Drag OLE DB source from other sources and drop into data flow tab



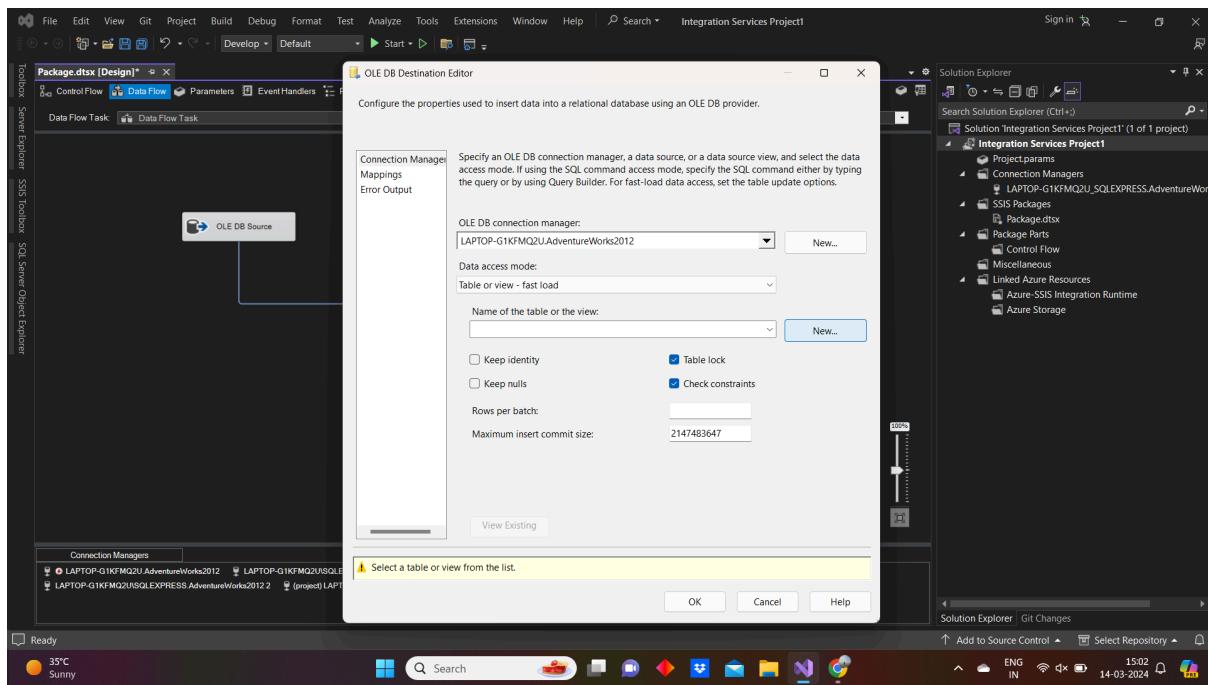
12. Double click on OLE DB source -> OLE DB source editor appears-> click on new to add connection manager . select [sales].[store] table from drop down -> ok



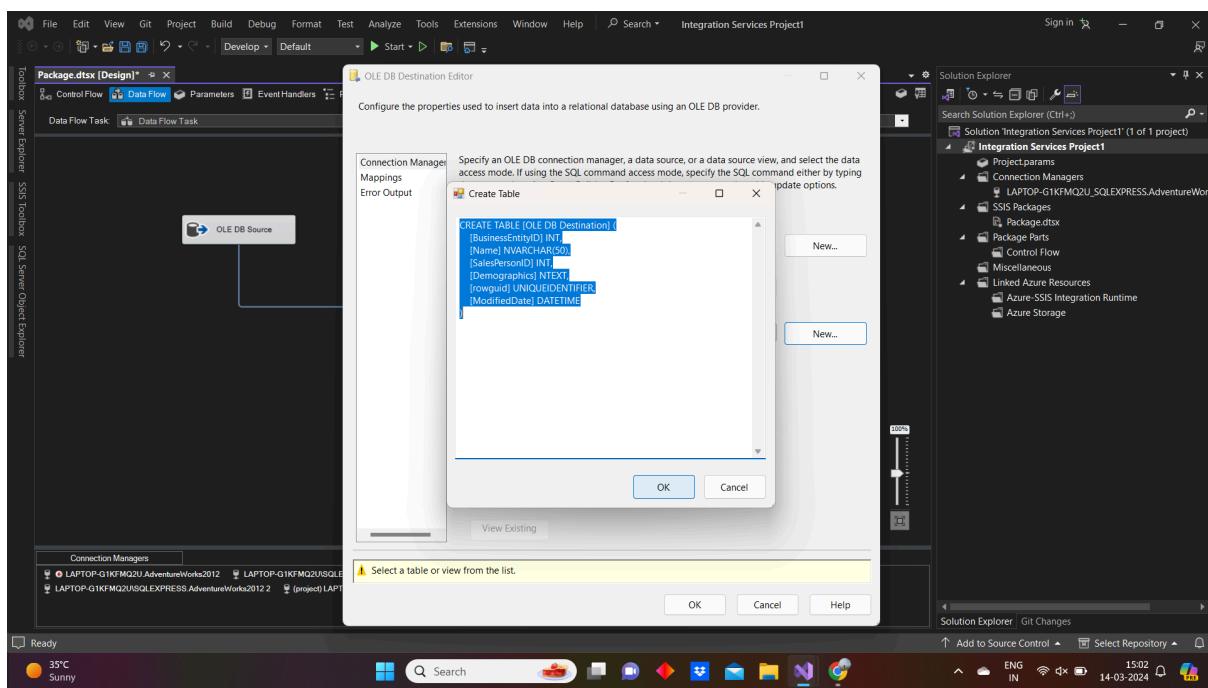
13. Drag OLE DB destination from other destination in data flow tab and connect both

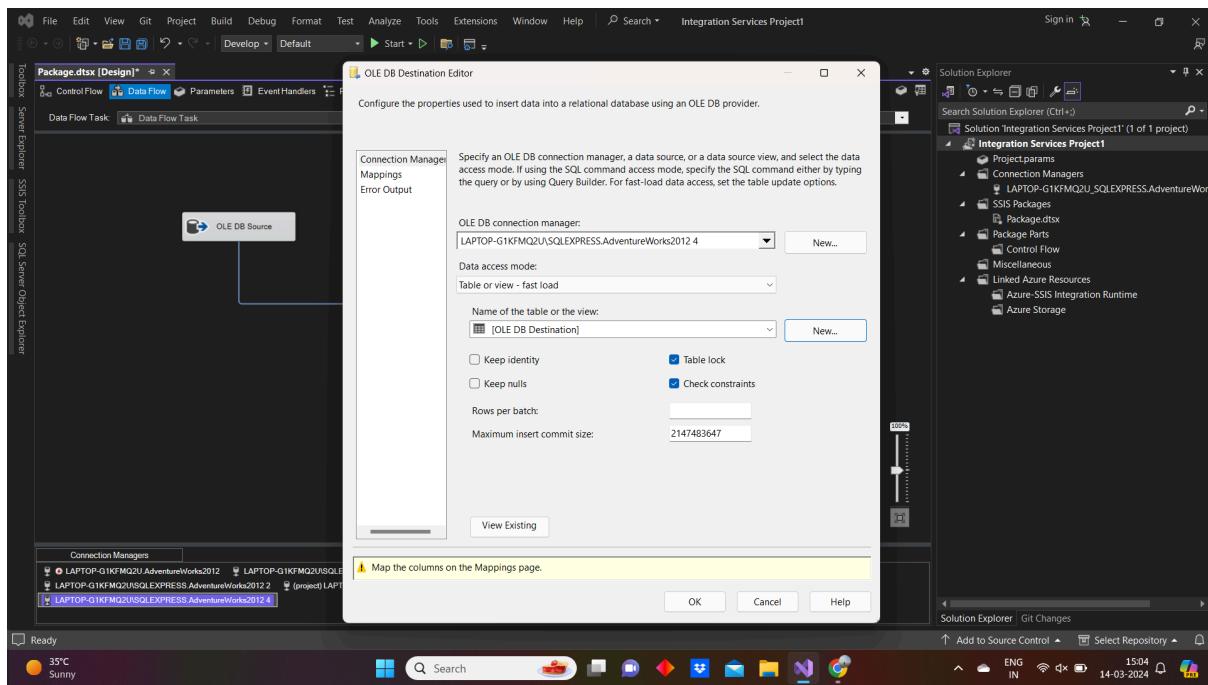


14. Double click on OLE DB destination click on new to run the query to get [OLEDB destination] in name of table or the view.

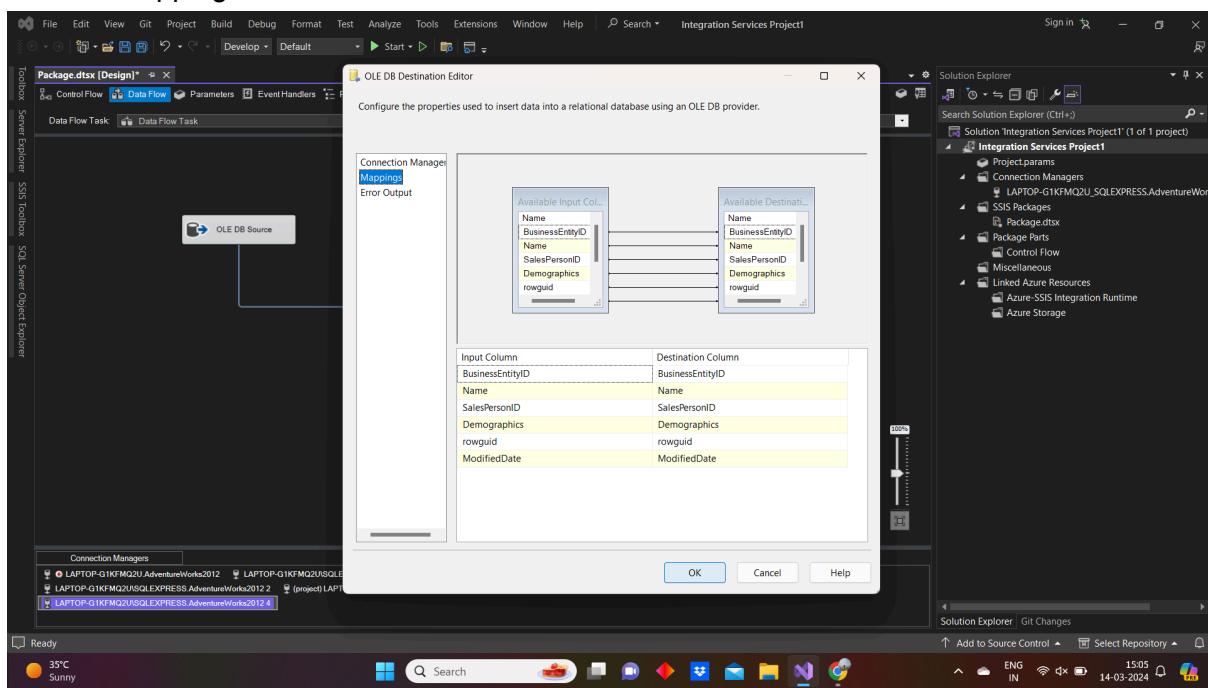


Click ok

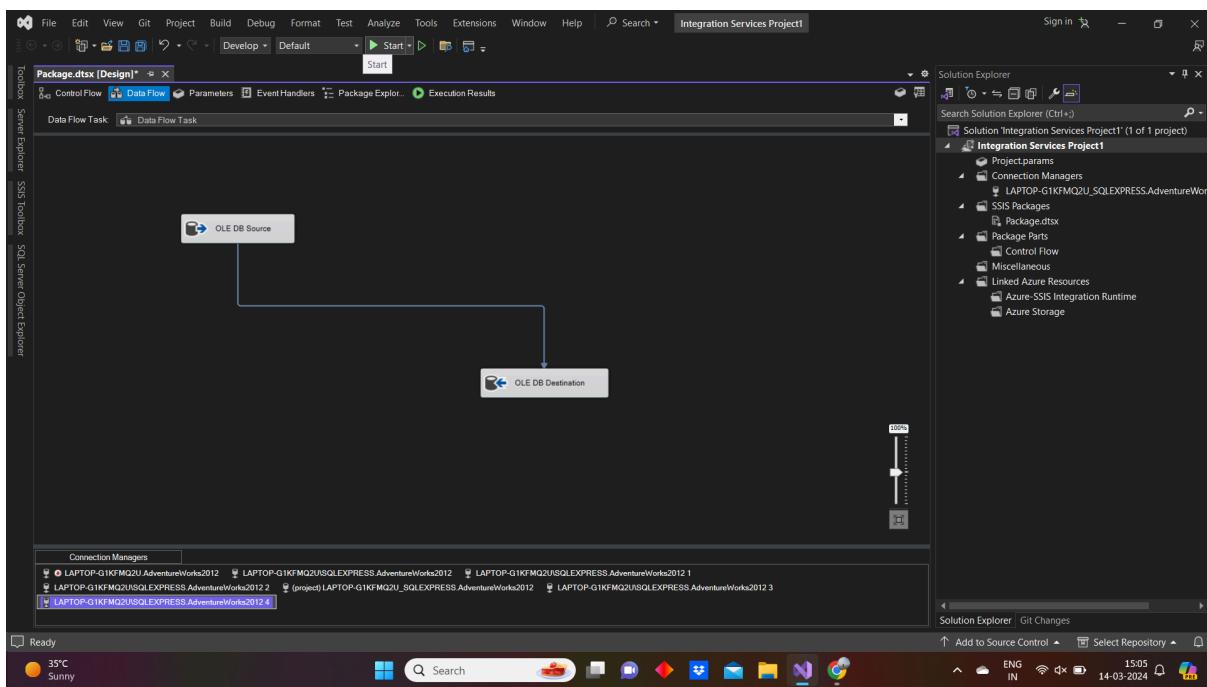




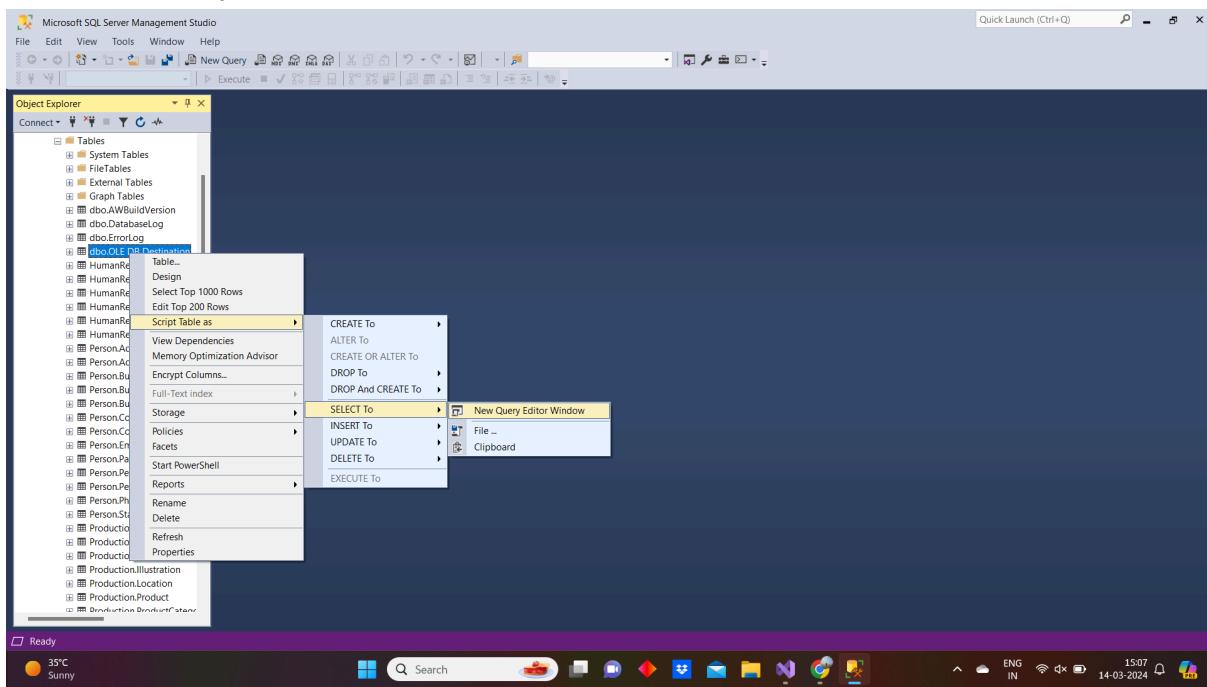
Click on mappings



15. Click on start



16. Go to SQL Server management studio in database tab -> adventureworks2019
->tables-> right click on [dbo].[OLE DB Destination] -> script table as ->select to ->
new query editor window



17. Execute following query to get output.

SQLQuery1.sql - LAPTOP-G1KFMQ2U\SQLEXPRESS.AdventureWorks2012 (LAPTOP-G1KFMQ2U\hp (61)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

AdventureWorks2012 Execute New Query Find Replace All Open Close Save All Refresh

Object Explorer

SQLQuery1.sql - L:\G1KFMQ2U\hp (61) GO

```
USE [AdventureWorks2012]
GO
SELECT [BusinessEntityID]
      ,[Name]
      ,[SalesPersonID]
      ,[Demographics]
      ,[Rowguid]
      ,[ModifiedDate]
  FROM [dbo].[OLE DB Destination]
GO
```

Result Messages

BusinessEntityID	Name	SalesPersonID	Demographics	rowguid	ModifiedDate
1 292	AutoDoor Bike Store	279	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	A2517E3-946D-4EBE-B9D9-7473432054	2014-09-12 11:15:07.497
2 294	Professional Sales and Service	276	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	00C0450B-C001-4A10-0075-2003080207184	2014-09-12 11:15:07.497
3 296	Riders Company	277	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	337C3880-1394-41A4-A08A-0542D296E4A9	2014-09-12 11:15:07.497
4 298	The Bike Mechanics	275	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	7894F279-F0C2-4116-BD75-213FD912023	2014-09-12 11:15:07.497
5 300	Nationwide Supply	286	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	C3FC705-A8C4-4F3A-9550-E2524487864B	2014-09-12 11:15:07.497
6 302	Area Bike Accessories	281	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	368BE8D0-30E5-4F8B-9A86-71FD40C5874E	2014-09-12 11:15:07.497
7 304	Bicycle Accessories and Kits	283	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	35F40E36-5105-4905-89E6-27E23119915	2014-09-12 11:15:07.497
8 306	Clamps & Brackets Co.	275	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	64D0BFC-D060-405C-8C60-0067F7E7C7DF	2014-09-12 11:15:07.497
9 308	Valley Bicycle Specialists	277	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	5938980C-6524-466B-B440-4E1711793333	2014-09-12 11:15:07.497
10 310	New Bikes Company	279	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	47C11252-2375-4181-9795-5A0381556	2014-09-12 11:15:07.497
11 312	Vinyl and Plastic Goods Corporation	282	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	0C951125-2375-4181-9795-5A0381556	2014-09-12 11:15:07.497
12 314	Tops of the Line Bikes	288	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	E390E337-4800-4B21-86C2-958B195340	2014-09-12 11:15:07.497
13 316	Fun Tops and Bikes	281	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	60CCD941-4192-40C7-99A4-5AB5A5E095	2014-09-12 11:15:07.497
14 318	Great Bikes	283	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	959FEC10-5E00-4175-8045-E10E30B4340	2014-09-12 11:15:07.497
15 320	Metropolitan Sales and Rental	275	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	00284FE2-5047-4077-8848-B5977A73E3E0	2014-09-12 11:15:07.497
16 322	Irregulars Outlet	288	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	CDE66279-83D8-4340-A83C-686E19514AC4	2014-09-12 11:15:07.497
17 324	United Tops Online	283	<StoreSurvey xmlns='http://schemas.microsoft.com/2006/06/01-identity/141-0075-2003080207184'	8A10D45A-0057-4E6C-8000-379ECE17CDAA	2014-09-12 11:15:07.497

LAPTOP-C1KFMQ2U\SQLEXPRESS ... | LAPTOP-G1KFMQ2U\hp (61) | AdventureWorks2012 | 00:00:00 | 701 rows

Ready 35°C Sunny

Ln 15 Col 1 Ch 1 INS ENG IN 15:08 14-03-2024

Practical 3

Practical 4 : Apply the What-if analysis for data visualization.

- Open excel sheet > enter data as shown below:

Book1 - Excel

File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do

Clipboard Font Alignment Number

SUM : X ✓ fx =A2*B2

A	B	C	D	E	F	G
Total Books	%Sold for Highest price			No. of books	Unit profit	
100	60.00%		Highest	=A2*B2	50	
			Lowest		20	
				Total price		
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						

Book1 - Excel

The screenshot shows the Microsoft Excel ribbon with the 'Home' tab selected. The formula bar displays the formula $=A2*(1-B2)$. A table is visible below the ribbon, containing the following data:

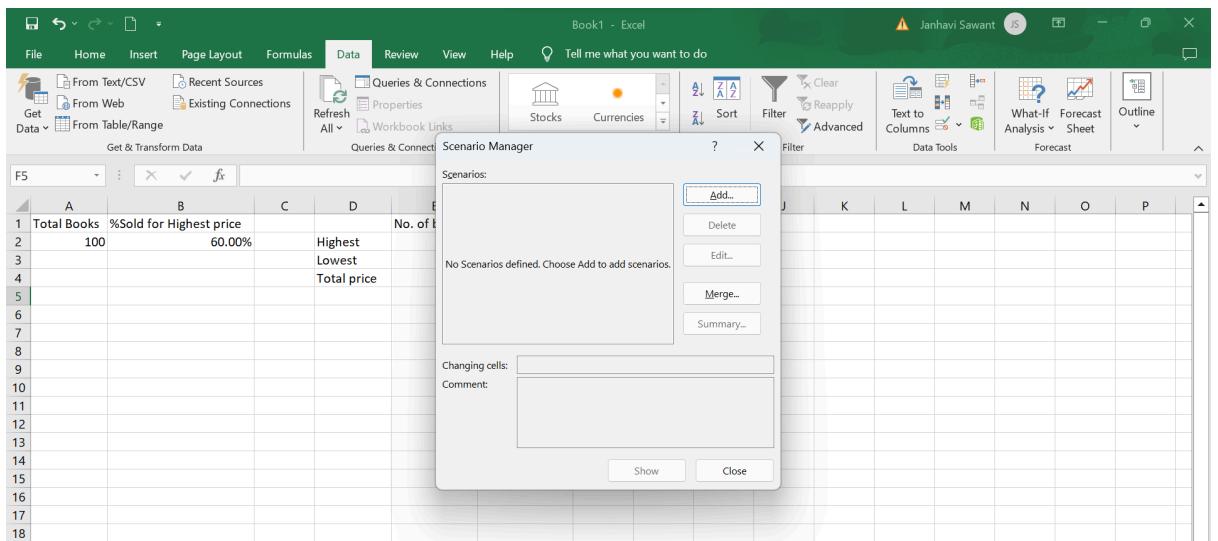
	A	B	C	D	E	F	G
1	Total Books	%Sold for Highest price			No. of books	Unit profit	
2	100	60.00%		Highest	60	50	
3				Lowest	$=A2*(1-B2)$	20	
4				Total price			
5							
6							
7							
8							
9							

Book1 - Excel

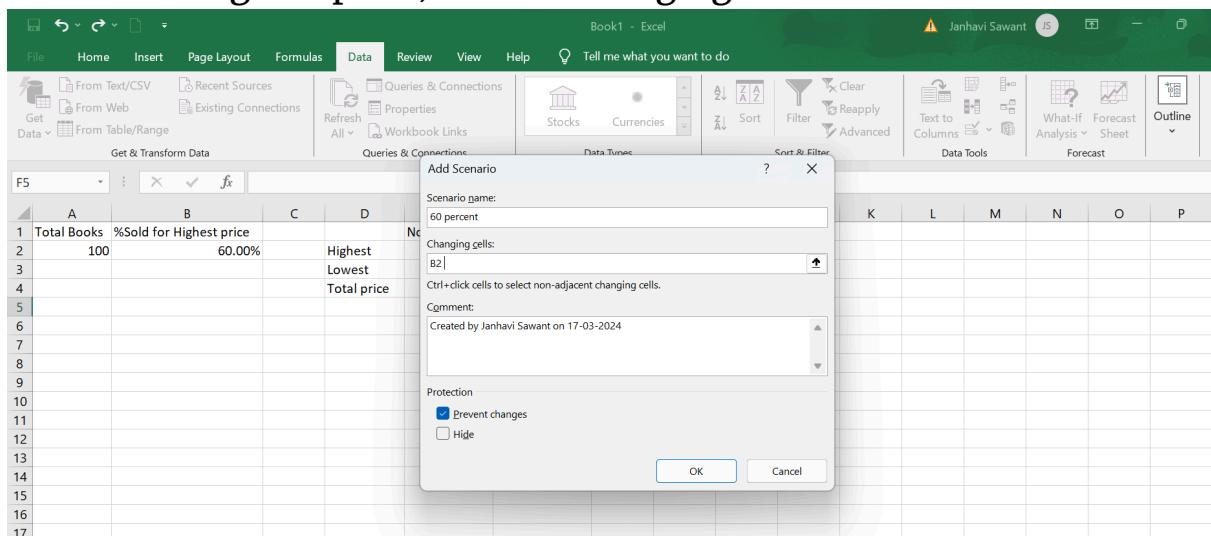
The screenshot shows the Microsoft Excel ribbon with the 'Home' tab selected. The formula bar displays the formula $=E2*F2+E3*F3$. A table is visible below the ribbon, containing the following data:

	A	B	C	D	E	F	G	H
1	Total Books	%Sold for Highest price			No. of books	Unit profit		
2	100	60.00%		Highest	60	50		
3				Lowest	40	20		
4				Total price	$=E2*F2+E3*F3$			
5								
6								
7								
8								
9								
10								
11								
12								

- Click on Data Tab > click on What -if analysis and select Scenario Manager.



- Add a scenario by clicking on add.
- Type a name (60 percent) , select cell \$B\$2(%sold for the highest price) for the changing cells and click on ok .



- Enter the corresponding value 0.6 and click on ok.

The screenshot shows the Microsoft Excel ribbon with the 'Data' tab selected. Below the ribbon, there are several groups of buttons: 'Get & Transform Data', 'Queries & Connections', 'Data Types', 'Sort & Filter', 'Data Tools', and 'Forecast'. The main worksheet area displays a table with columns for Total Books, %Sold for Highest price, Highest, Lowest, and Total price. The 'Highest' column contains the formula =MAX(B2:B4). The 'Lowest' column contains the formula =MIN(B2:B4). The 'Total price' column contains the formula =SUM(B2:B4). A 'Scenario Values' dialog box is overlaid on the worksheet, prompting for values for cells \$B\$2 and \$B\$3.

- Add 4 other scenarios (70% , 80%, 90% & 100%) .

The screenshot shows the Microsoft Excel interface with the 'Data' tab selected in the ribbon. The main area displays a table with columns for Total Books, %Sold for Highest price, No. of books, and Unit profit. A 'Scenario Values' dialog box is overlaid on the screen, prompting for values for changing cells. The dialog includes fields for '1:' and '\$B\$2 0.7', and buttons for 'Add', 'OK', and 'Cancel'.

A	B	C	D	E	F	G	H	I	J	K	L
Total Books	%Sold for Highest price			No. of books	Unit profit						
100	60.00%		Highest	60	50						
			Lowest								
			Total price								

Scenario Values
Enter values for each of the changing cells.
1: \$B\$2 0.7

Add OK Cancel

Book1 - Excel

Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do

From Text/CSV From Web From Table/Range Get & Transform Data

Recent Sources Existing Connections Refresh All Workbook Links

Queries & Connections Properties

Stocks Currencies

A Z A Z Sort Filter Reapply Advanced

A	B	C	D	E	F	G	H	I	J	K
Total Books	%Sold for Highest price			No. of books	Unit profit					
100	60.00%		Highest	60	50					
			Lowest							
			Total price							

Scenario Values

Enter values for each of the changing cells.

1: \$B\$2 0.8

Add OK Cancel

Book1 - Excel

File Home Insert Page Layout Formulas Data Review View Help Tell me what you want to do

From Text/CSV From Web From Table/Range Get & Transform Data

Recent Sources Existing Connections Refresh All Workbook Links

Queries & Connections Properties

Stocks Currencies

A Z A Z Sort Filter Reapply Advanced

Text Col

A	B	C	D	E	F	G	H	I	J	K	L
1 Total Books	%Sold for Highest price			No. of books	Unit profit						
2 100	60.00%		Highest	60	50						
3			Lowest								
4			Total price								
5											
6											
7											
8											
9											
10											
11											
12											
13											

Scenario Values

Enter values for each of the changing cells.

1: \$B\$2 0.9

Add OK Cancel

The screenshot shows a Microsoft Excel spreadsheet titled "Book1 - Excel". The Data tab is selected. A scenario manager dialog box is open over the spreadsheet. The dialog box has the following fields:

- Scenarios:** A list box containing "60 percent", "70 percent", "80 percent", "90 percent", and "100 percent".
- Changing cells:** A text input field containing "\$B\$2".
- Comment:** A text area containing "Created by Janhavi Sawant on 17-03-2024".
- Buttons:** "Add...", "Delete", "Edit...", "Merge...", "Summary...", "Show", and "Close".

The spreadsheet contains the following data:

	A	B	C	D	E	F	G	H	I	J
1	Total Books	%Sold for Highest price			No. of books	Unit profit				
2	100	60.00%		Highest	60	50				
3				Lowest						
4				Total price						
5										
6										
7										
8										
9										
10										
11										
12										
13										

- Click on summary.

A

B

C

D

E

F

G

H

I

J

Total Books %Sold for Highest price No. of books Up profit Scenario Summary ? X

100 60.00% Highest 60 Scenario Summary

3 Lowest 40

4 Total price

A B C D E F G H I J

Scenario Summary						
	Current Values:	60 percent	70 percent	80 percent	90 percent	100 percent
Changing Cells:	\$B\$2	60.00%	60.00%	70.00%	80.00%	90.00%
Result Cells:	\$F\$4	3800	3800	4100	4400	4700
Notes: Current Values column represents values of changing cells at time Scenario Summary Report was created. Changing cells for each scenario are highlighted in gray.						

Practical 5.

Perform the Data classification using classification Algorithm.

- Install R and RStudio and open it.

Code :-

```
rainfalls <- c(799,1174,865,1334,635,918,686,998,784,985,882,1071)
rainfalls.timeseries <- ts(rainfalls,start=c(2022,1),frequency=12)
print(rainfalls.timeseries)
png(file="rainfall")
plot(rainfalls.timeseries)
dev.off()
plot(rainfalls.timeseries)
```

Output

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Console Terminal Background Jobs
R 4.3.3 - ~\r
R version 4.3.3 (2024-02-29 ucrt) -- "Angel Food Cake"
Copyright (C) 2024 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.

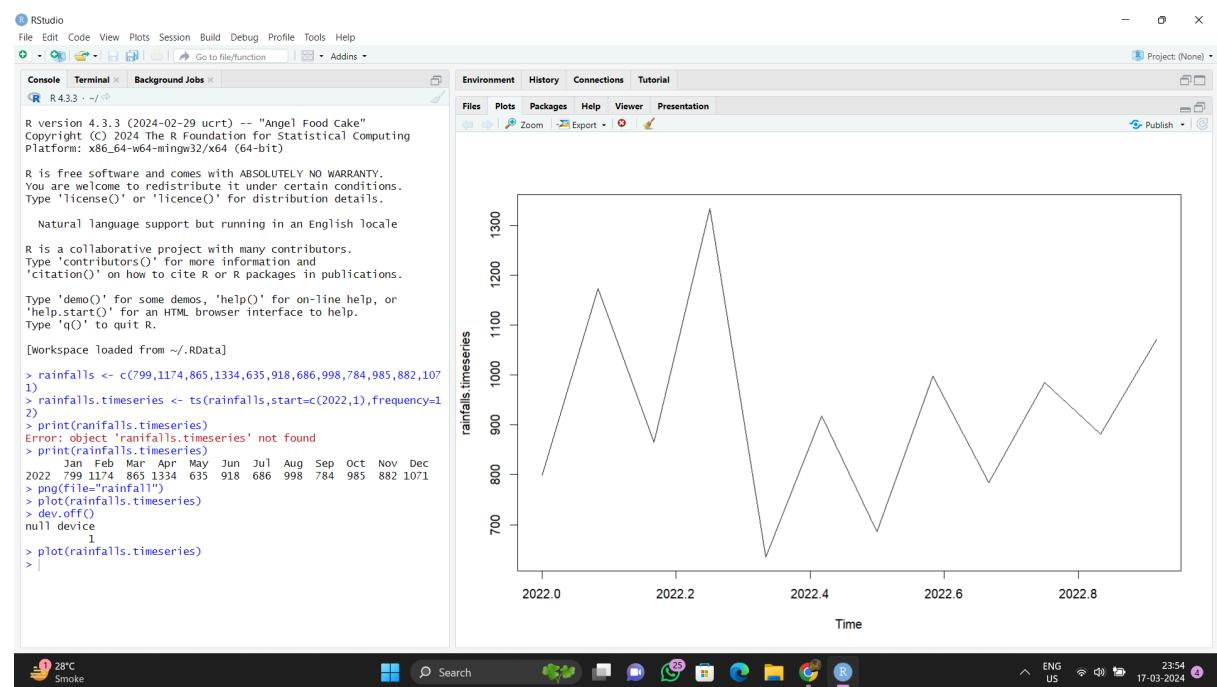
Natural language support but running in an English locale

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.

[workspace loaded from ~/.RData]

> rainfalls <- c(799,1174,865,1334,635,918,686,998,784,985,882,1071)
> rainfalls.timeseries <- ts(rainfalls,start=c(2022,1),frequency=12)
> print(rainfalls.timeseries)
Error: object 'rainfalls.timeseries' not found
> print(rainfalls.timeseries)
  Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec
2022 799 1174 865 1334 635 918 686 998 784 985 882 1071
> png(file="rainfall")
> plot(rainfalls.timeseries)
> dev.off()
null device
1
> plot(rainfalls.timeseries)
> |
```



Practical 6

K-Means clustering using R

Code :-

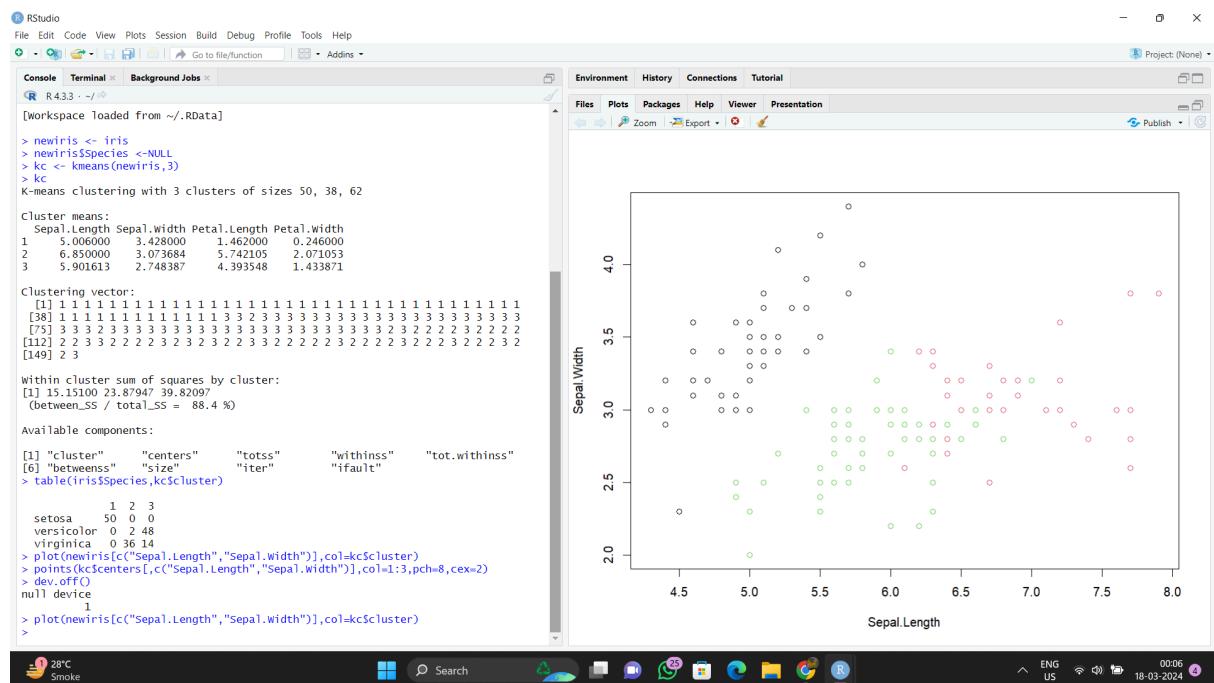
```
newiris <- iris
newiris$Species <- NULL
```

```

kc <- kmeans(newiris,3)
kc
table(iris$Species,kc$cluster)
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)
dev.off()
plot(newiris[c("Sepal.Length","Sepal.Width")],col=kc$cluster)

```

Output



Practical 7

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)
print(relation)

```

```

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)

```

```

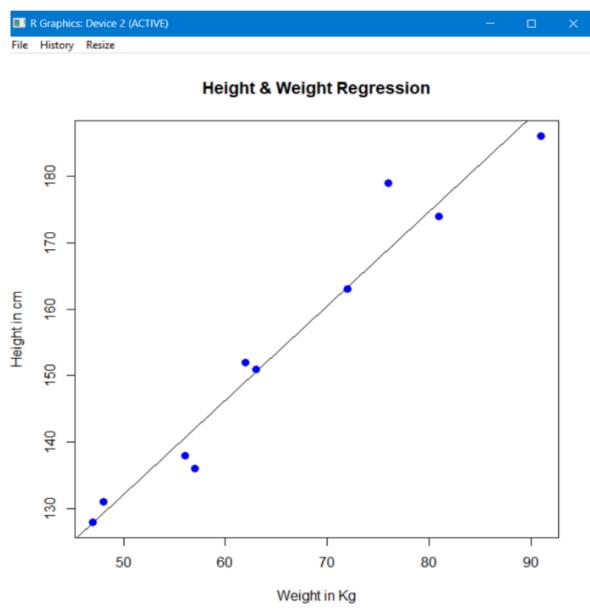
print(summary(relation))

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)
a <-data.frame(x=170)
result<-predict(relation,a)
print(result)

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="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()
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")

```

Output



Practical 8

Perform logistic regression on the given data warehouse data

Code :-

```
quality <- read.csv("C/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)
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)
table(qualityTrain$PoorCare,predictTrain>0.7)
8/25
73/74
table(qualityTrain$PoorCare,predictTrain>0.2)
16/25
54/74
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.ad
j=c(-0.2,1.7))
```

Output