

Business Intelligence  
and  
Big data analysis  
(Business Intelligence)

Journal

By

**Amal Sathyan**

**Roll no: 26**

**MSC CS SEM-II**



*Thakur Educational Trust's (Regd.)*

**THAKUR COLLEGE OF SCIENCE & COMMERCE**

AUTONOMOUS COLLEGE, PERMANENTLY AFFILIATED TO UNIVERSITY OF MUMBAI

NAAC Accredited Grade 'A' (3<sup>rd</sup> Cycle) & ISO 9001: 2015 (Certified)

**Azadi Ka Amrit Mahotsav Best College Award by University of Mumbai for the Year 2018-2019**

**tcsc**

**CELEBRATING  
25 YEARS OF GLORY**

**Department of Computer Science  
Journal Certificate**

Class: M.Sc. Computer Science – Part I (Semester: II)

UID: 2240426

Roll No: 026

**Academic Year: 2022-2023**

This is to certify that the work entered in this journal is the work of Mr/Ms Amal Sathyam, who has worked for academic year 2022-2023 in the computer laboratory. He / She has completed prescribed practical of following course satisfactorily.

Course Title: : Cyber & Information Security (Network & Comm. Security)

**Teacher In-charge**

**Head of Department**

**Date:**

**Examiner**

## **INDEX**

<b>Sr. No.</b>	<b>Experiments</b>	<b>Page No</b>	<b>Date</b>
1	Prerequisites for Database applications.	4	
2	Develop an application to implement OLAP, roll-up, drill-down, slice, and dice operations	10	
3	Design and create cube by identifying measures and dimensions for star schema	13	
4	Create and uses excel pivot table report based on data cube operations.	25	
5	Develop an application to import/export data from external sources..	33	
6	Design and create cube by identifying measures and dimensions for snow flake schema.	44	
7	Develop an application to construct a multidimensional data.	51	
8	Develop an application to implement defining subject areas, design of fact and dimension tables, data marts.	60	



# Practical 1

## Aim: Prerequisites for Database applications.

### Queries:

#### Creating employee table:

```
create table employee(  
    empid number(3) primary key,  
    empname varchar2(25),  
    age number(3));
```

#### Creating client table:

```
create table client(  
    clientid number(3) primary key,  
    clientname varchar(25),  
    location varchar2(20));
```

#### Creating project table:

```
create table project(  
    projid number(5) primary key,  
    empid number(3),  
    projname varchar2(25),  
    clientid number(3));
```

#### Inserting records in table employee:

```
insert into employee values(1,'Vardhan',22);  
insert into employee values(2,'Khushi',30);  
insert into employee values(3,'Simran',40);  
insert into employee values(4,'Chetan',20);  
insert into employee values(5,'Swati',27);  
insert into employee values(10,'Vedant',33);  
insert into employee values(15,'Gauri',35);  
insert into employee values(16,'ABC',36);  
insert into employee values(17,'PQR',37);
```

#### Inserting records in table client:

```
insert into client values(33,'will','mumbai');  
insert into client values(34,'jones','bangalore');  
insert into client values(35,'deep','kolkata');  
insert into client values(36,'smith','delhi');  
insert into client values(37,'marin','mumbai');  
insert into client values(38,'jason','mumbai');
```

#### Inserting records in table project:

```
insert into project values(111,1,'abc',33);  
insert into project values(222,2,'def',34);  
insert into project values(333,3,'xyz',35);  
insert into project values(444,4,'pqr',36);  
insert into project values(555,5,'mnq',37);  
insert into project values(666,6,'efg',38);  
insert into project values(777,7,'lmn',39);  
insert into project values(888,8,'ppd',40);
```

```
SQL> desc employee;
Name                           Null?    Type
-----                         -----
EMPID                          NOT NULL NUMBER(3)
EMPNAME                         VARCHAR2(25)
AGE                            NUMBER(3)
```

```
SQL> desc client;
Name                           Null?    Type
-----                         -----
CLIENTID                        NOT NULL NUMBER(3)
CLIENTNAME                       VARCHAR2(25)
LOCATION                          VARCHAR2(20)
```

```
SQL> desc project;
Name                           Null?    Type
-----                         -----
PROJID                          NOT NULL NUMBER(5)
EMPID                           NUMBER(3)
PROJNAME                         VARCHAR2(25)
CLIENTID                         NUMBER(3)
```

```
SQL> select * from employee;
      EMPID  EMPNAME          AGE
-----  -----
      1 Vardhan           22
      2 Khushi            30
      3 Simran            40
      4 Chetan             20
      5 Swati             27
```

```
SQL> select * from client;
      CLIENTID  CLIENTNAME          LOCATION
-----  -----
      33 will        mumbai
      34 jones       bangalore
      35 deep        kolkata
      36 smith       delhi
      37 marin       mumbai
      38 jason       mumbai
```

```
SQL> insert into project values(111,1,'abc',33);
1 row created.
```

```

SQL> select * from project;
      PROJID    EMPID PROJNAME          CLIENTID
-----  -----  -----
      111        1  abc                  33
      222        2  def                  34
      333        3  xyz                  35
      444        4  pqr                  36
      555        5  mnq                  37
      666        6  efg                  38
      777        7  lmn                  39
      888        8  ppd                  40
8 rows selected.

```

```

SQL> select employee.empid,employee.empname,project.clientid
  2  from employee
  3  INNER JOIN
  4  project
  5  ON employee.empid=project.empid;

      EMPID  EMPNAME          CLIENTID
-----  -----
      1  Vardhan                33
      2  Khushi                 34
      3  Simran                 35
      4  Chetan                 36
      5  Swati                  37

```

```

SQL> select employee.empid,employee.empname,project.clientid
  2  from employee
  3  FULL JOIN
  4  project
  5  ON employee.empid=project.empid;

      EMPID  EMPNAME          CLIENTID
-----  -----
      1  Vardhan                33
      2  Khushi                 34
      3  Simran                 35
      4  Chetan                 36
      5  Swati                  37
                           38
                           39
                           40
8 rows selected.

```

```

SQL> select employee.empid, employee.empname,project.clientid
  2  from employee
  3  LEFT JOIN
  4  project
  5  ON employee.empid=project.empid;

      EMPID  EMPNAME          CLIENTID
-----  -----  -----
        1  Vardhan            33
        2  Khushi              34
        3  Simran              35
        4  Chetan              36
        5  Swati               37

```

```

SQL> select employee.empid, employee.empname,project.clientid
  2  from employee
  3  RIGHT JOIN
  4  project
  5  ON employee.empid=project.empid;

      EMPID  EMPNAME          CLIENTID
-----  -----  -----
        1  Vardhan            33
        2  Khushi              34
        3  Simran              35
        4  Chetan              36
        5  Swati               37
                           40
                           38
                           39

8 rows selected.

```

## Transactions

```

SQL> set transaction read write;
Transaction set.

SQL> UPDATE EMPLOYEE
  2  SET EMPID = 201
  3  WHERE EMPID = 1;

1 row updated.

SQL> commit;

Commit complete.

```

```

SQL> insert into employee values(16,'ABC',36);
1 row created.

SQL> insert into employee values(17,'PQR',37);
1 row created.

```

## Savepoint

```
SQL> savepoint s1;
Savepoint created.

SQL> insert into employee values(6,'ABC',20);
1 row created.
```

## Rollback

```
SQL> rollback to s1;
Rollback complete.

SQL> select * from employee;

  EMPID EMPNAME          AGE
----- -----
  201  Vardhan           22
    2  Khushi             30
    3  Simran            40
    4  Chetan            20
    5  Swati             27
```

## Wild cards

```
SQL> select * from employee
  2 where empname like '%an';

  EMPID EMPNAME          AGE
----- -----
  201  Vardhan           22
    3  Simran            40
    4  Chetan            20
```

```
SQL> select * from employee
  2 where empname like 'Khu%';

  EMPID EMPNAME          AGE
----- -----
    2  Khushi             30
```

```
SQL> select * from employee
  2 where empname like'%ar%';

  EMPID EMPNAME          AGE
----- -----
  201  Vardhan           22
```

## Distinct

```
SQL> insert into employee values(6,'ABC',20);
1 row created.

SQL> insert into employee values(7,'ABC',22);
1 row created.

SQL> select distinct age from employee;

      AGE
-----
      22
      30
      20
      40
      27
```

**Conclusion:** Pre requisites done.

## Practical 2

**Aim: Develop an application to implement OLAP, roll-up, drill-down, slice, and dice operations.**

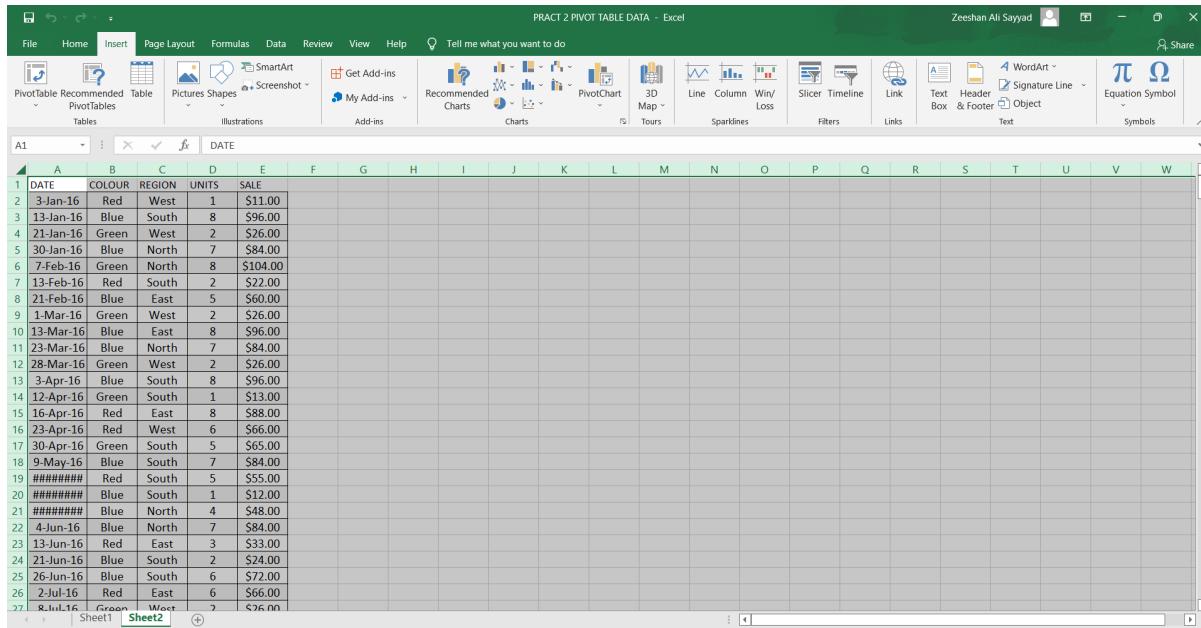
### Theory:

A pivot table is a statistics tool that summarizes and reorganizes selected columns and rows of data in a spreadsheet or database table to obtain a desired report. The tool does not actually change the spreadsheet or database itself, it simply “pivots” or turns the data to view it from different perspectives.

Pivot tables are especially useful with large amounts of data that would be time-consuming to calculate by hand. A few data processing functions a pivot table can perform include identifying sums, averages, ranges or outliers. The table then arranges this information in a simple, meaningful layout that draws attention to key values.

Steps:

1) Select source data in excel sheet.



The screenshot shows a Microsoft Excel window titled "PRACT 2 PIVOT TABLE DATA - Excel". The ribbon menu is visible at the top, with the "Insert" tab selected. Below the ribbon, there is a toolbar with various icons for inserting shapes, charts, and filters. The main area of the screen displays a data table in a grid format. The columns are labeled A through W, and the rows are numbered 1 through 27. The data includes columns for DATE, COLOUR, REGION, UNITS, and SALE. The "SALE" column contains numerical values such as \$11.00, \$96.00, \$26.00, etc. The "DATE" column shows dates from January 3, 2016, to July 8, 2016. The "COLOUR" and "REGION" columns categorize the data. The bottom of the screen shows the Excel ribbon tabs and the status bar.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	DATE	COLOUR	REGION	UNITS	SALE																	
2	3-Jan-16	Red	West	1	\$11.00																	
3	13-Jan-16	Blue	South	8	\$96.00																	
4	21-Jan-16	Green	West	2	\$26.00																	
5	30-Jan-16	Blue	North	7	\$84.00																	
6	7-Feb-16	Green	North	8	\$104.00																	
7	13-Feb-16	Red	South	2	\$22.00																	
8	21-Feb-16	Blue	East	5	\$60.00																	
9	1-Mar-16	Green	West	2	\$26.00																	
10	13-Mar-16	Blue	East	8	\$96.00																	
11	23-Mar-16	Blue	North	7	\$84.00																	
12	28-Mar-16	Green	West	2	\$26.00																	
13	3-Apr-16	Blue	South	8	\$96.00																	
14	12-Apr-16	Green	South	1	\$13.00																	
15	16-Apr-16	Red	East	8	\$88.00																	
16	23-Apr-16	Red	West	6	\$66.00																	
17	30-Apr-16	Green	South	5	\$65.00																	
18	9-May-16	Blue	South	7	\$84.00																	
19	#####	Red	South	5	\$55.00																	
20	#####	Blue	South	1	\$12.00																	
21	#####	Blue	North	4	\$48.00																	
22	4-Jun-16	Blue	North	7	\$84.00																	
23	13-Jun-16	Red	East	3	\$33.00																	
24	21-Jun-16	Blue	South	2	\$24.00																	
25	26-Jun-16	Blue	South	6	\$72.00																	
26	2-Jul-16	Red	East	6	\$66.00																	
27	8-Jul-16	Green	West	2	\$26.00																	

2) Click on Insert tab > Pivot Table > Select source and location > Click Ok

PRACT 2 PIVOT TABLE DATA - Excel

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

PivotTable Recommended PivotTables Tables Illustrations Add-ins Recommended Charts Charts Tours Sparklines Filters Links

M1 DATE COLOUR REGION UNITS SALE

3-Jan-16 Red West 1 \$11.00

13-Jan-16 Blue South 8 \$96.00

21-Jan-16 Green West 2 \$26.00

30-Jan-16 Blue North 7 \$84.00

7-Feb-16 Green North 8 \$104.00

13-Feb-16 Red South 2 \$22.00

21-Feb-16 Blue East 5 \$60.00

1-Mar-16 Green West 2 \$26.00

13-Mar-16 Blue East 8 \$96.00

23-Mar-16 Blue North 7 \$84.00

28-Mar-16 Green West 2 \$26.00

3-Apr-16 Blue South 8 \$96.00

12-Apr-16 Green South 1 \$13.00

16-Apr-16 Red East 8 \$88.00

23-Apr-16 Red West 6 \$66.00

30-Apr-16 Green South 5 \$65.00

9-May-16 Blue South 7 \$84.00

##### Red South 5 \$55.00

##### Blue South 1 \$12.00

##### Blue North 4 \$48.00

4-Jun-16 Blue North 7 \$84.00

13-Jun-16 Red East 3 \$33.00

21-Jun-16 Blue South 2 \$24.00

26-Jun-16 Blue South 6 \$72.00

2-Jul-16 Red East 6 \$66.00

8-Jul-16 Green West 2 \$26.00

3) Now Drop PivotTable Fields into Filters, Columns, Rows, Values Respectively.

PRACT 2 PIVOT TABLE DATA - Excel

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

PivotTable Name: Active Field: Sum of UNITS

PivotTable1 Active Field: Sum of UNITS

Group Selection

Insert Slicer Filter Connections

Refresh Change Data Clear Select Move PivotTable

Fields, Items, & Sets OLAP Relationships Calculations

PivotChart Recommended PivotTables Tools

Field List Buttons Headers Show

L5 DATE COLOUR REGION UNITS SALE

3-Jan-16 Red West 1 \$11.00

13-Jan-16 Blue South 8 \$96.00

21-Jan-16 Green West 2 \$26.00

30-Jan-16 Blue North 7 \$84.00

7-Feb-16 Green North 8 \$104.00

13-Feb-16 Red South 2 \$22.00

21-Feb-16 Blue East 5 \$60.00

1-Mar-16 Green West 2 \$26.00

13-Mar-16 Blue East 8 \$96.00

23-Mar-16 Blue North 7 \$84.00

28-Mar-16 Green West 2 \$26.00

3-Apr-16 Blue South 8 \$96.00

12-Apr-16 Green South 1 \$13.00

16-Apr-16 Red East 8 \$88.00

23-Apr-16 Red West 6 \$66.00

30-Apr-16 Green South 5 \$65.00

9-May-16 Blue South 7 \$84.00

##### Red South 5 \$55.00

##### Blue South 1 \$12.00

##### Blue North 4 \$48.00

4-Jun-16 Blue North 7 \$84.00

13-Jun-16 Red East 3 \$33.00

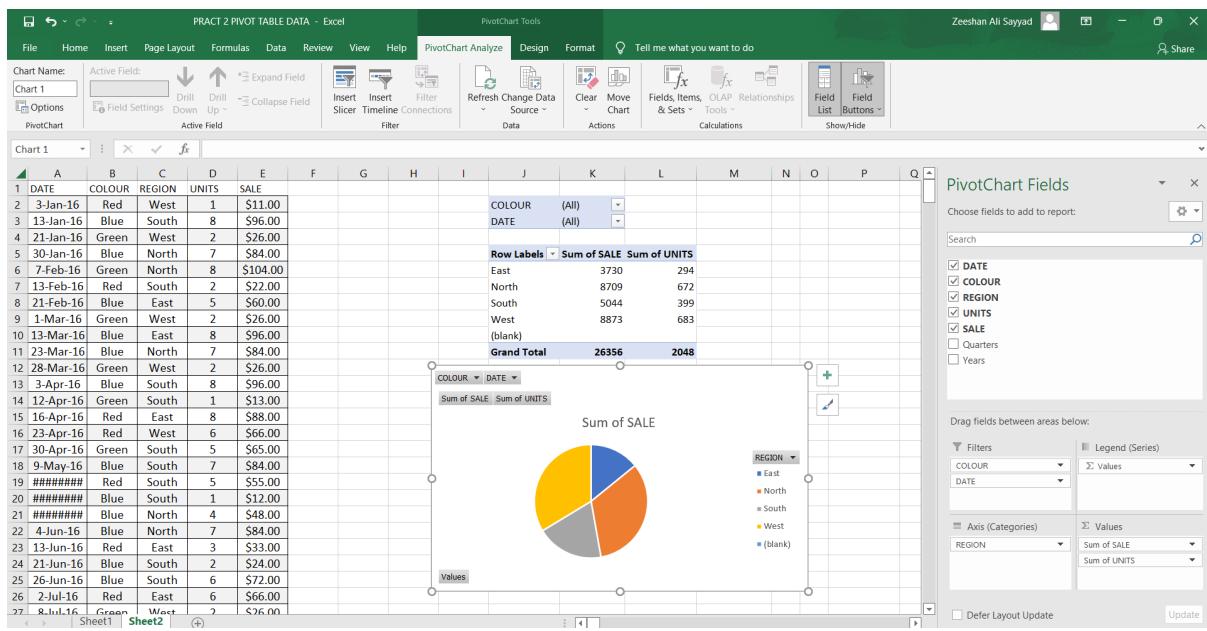
21-Jun-16 Blue South 2 \$24.00

26-Jun-16 Blue South 6 \$72.00

2-Jul-16 Red East 6 \$66.00

8-Jul-16 Green West 2 \$26.00

#### 4) Insert Pie Chart or Bar graph.



**Conclusion:** Pivot table implemented successfull

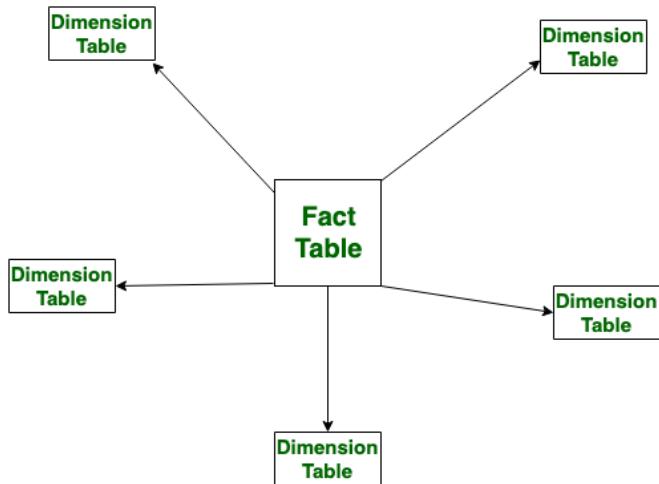
## Practical 3

**Aim:** Design and create cube by identifying measures and dimensions for star schema.

**Theory:**

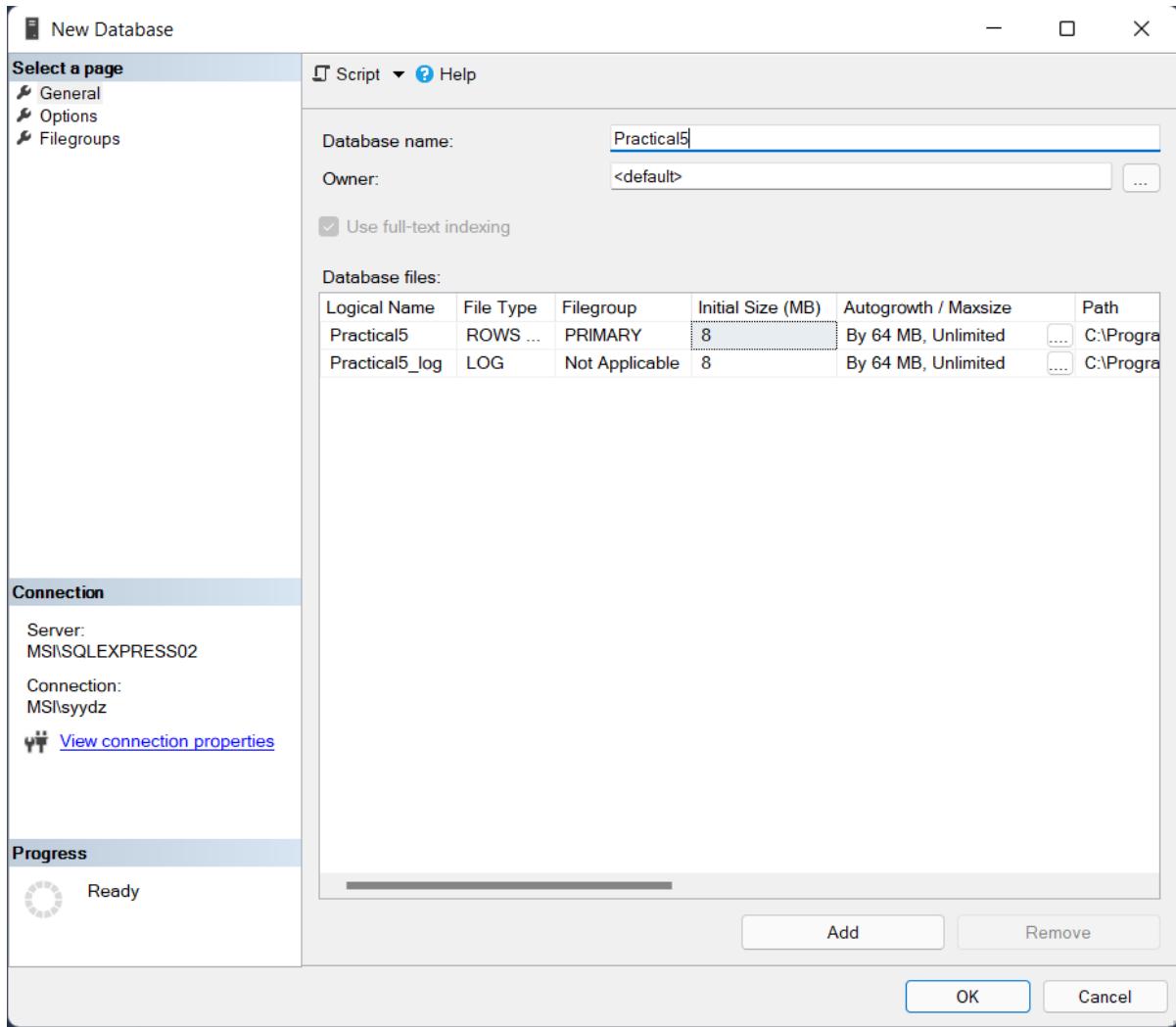
**Star Schema:**

Star schema is the type of multidimensional model which is used for data warehouse. In star schema, the fact tables and the dimension tables are contained. In this schema fewer foreign-key join is used. This schema forms a star with fact table and dimension tables.

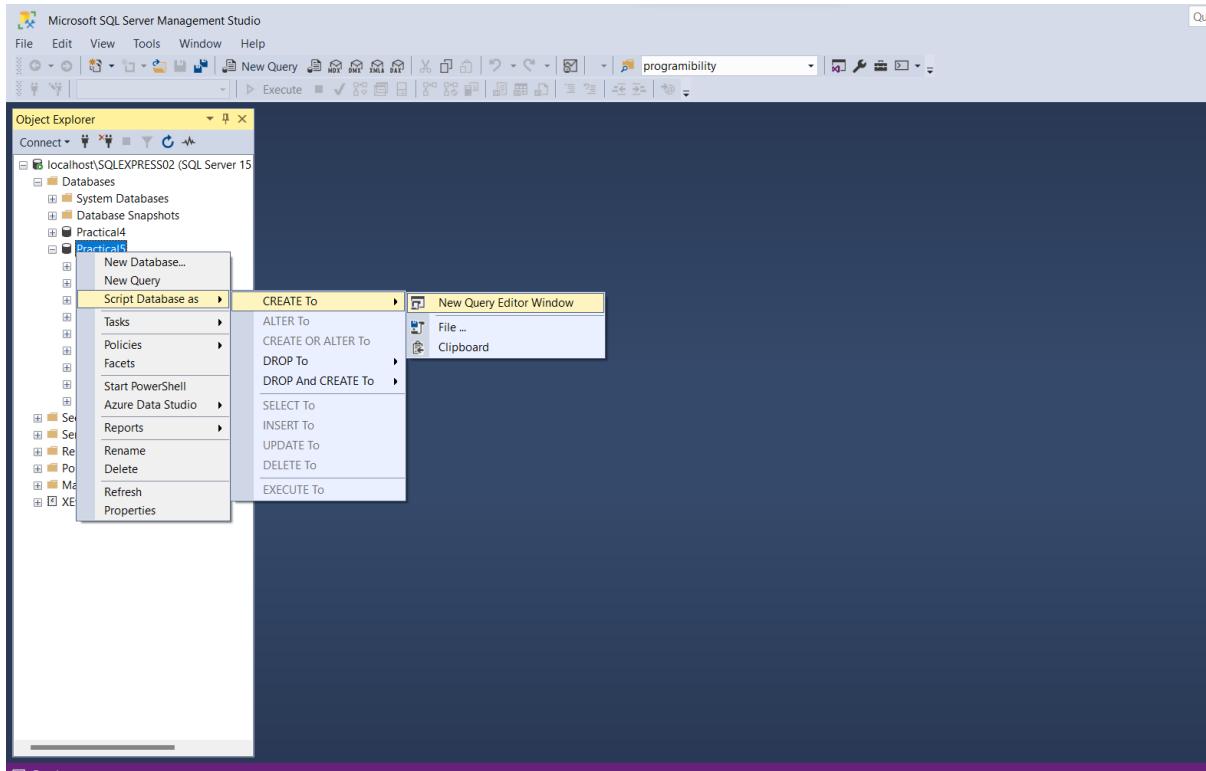


#Steps

- 1) Create New Database



2) Right Click on Database > Script Database as > Create to > Select: New Query Editor Window.



3) Remove pre-written queries

4) Create tables:

```
CREATE TABLE DIM_EMPLOYEE
(EMP_ID INT,
EMP_NAME VARCHAR(25))
CREATE TABLE DIM_BRANCH
(BRANCH_ID INT,
BRANCH_NAME VARCHAR(25))
CREATE TABLE DIM_PRODUCT
(PROD_ID INT,
PROD_NAME VARCHAR(25))
CREATE TABLE DIM_CUSTOMER
(CUST_ID INT,
CUST_NAME VARCHAR(25))
CREATE TABLE FACT_SHOP
(EMP_ID INT,
BRANCH_ID INT,
PROD_ID INT,
CUST_ID INT)
```

4) Execute queries

```
CREATE TABLE DIM_EMPLOYEE
(EMP_ID INT,
EMP_NAME VARCHAR(25))

CREATE TABLE DIM_BRANCH
(BRANCH_ID INT,
BRANCH_NAME VARCHAR(25))

CREATE TABLE DIM_PRODUCT
(PROD_ID INT,
PROD_NAME VARCHAR(25))

CREATE TABLE DIM_CUSTOMER
(CUST_ID INT,
CUST_NAME VARCHAR(25))

CREATE TABLE FACT_SHOP
(EMP_ID INT,
BRANCH_ID INT,
PROD_ID INT,
CUST_ID INT)
```

Messages

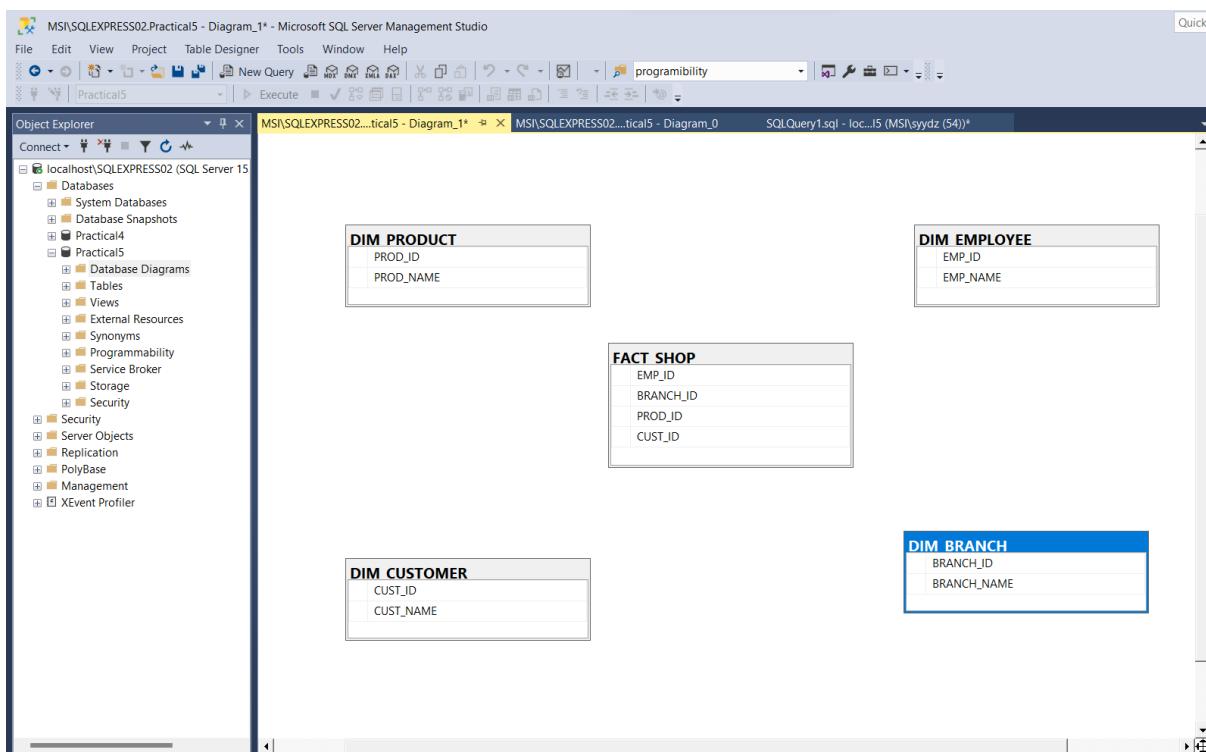
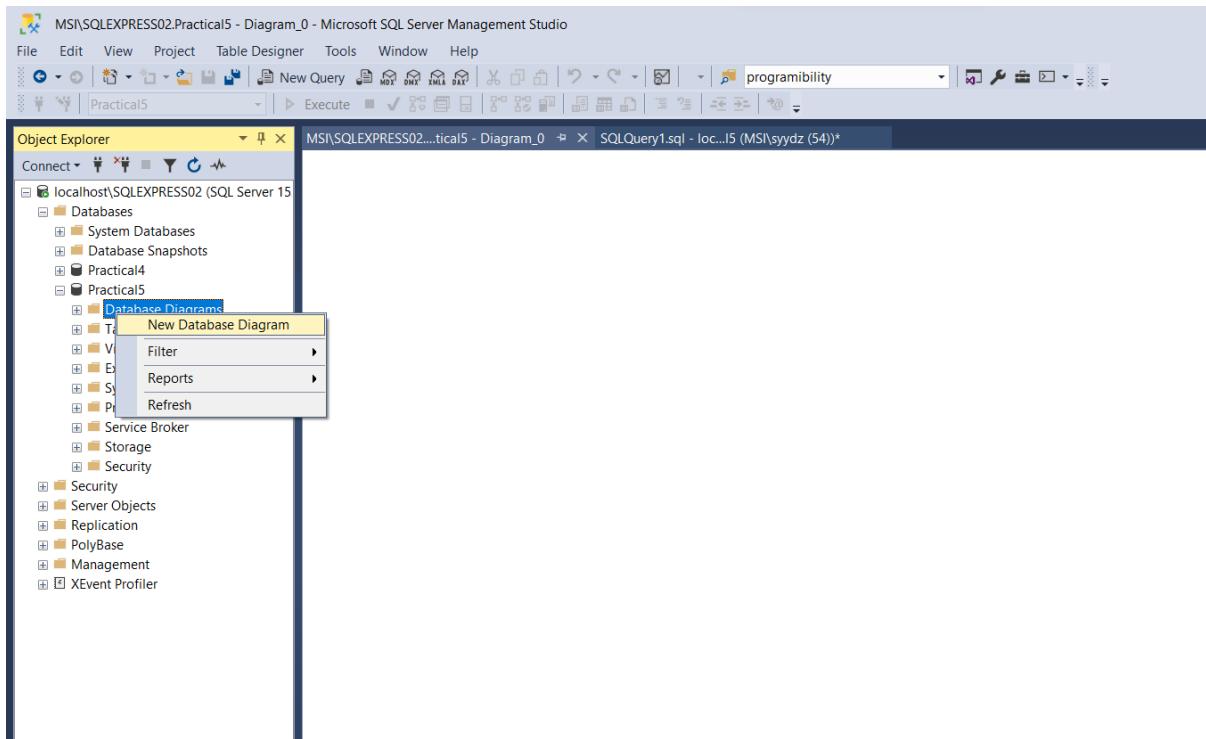
Commands completed successfully.

Completion time: 2022-04-08T22:55:46.9780615+08:30

Query executed successfully.

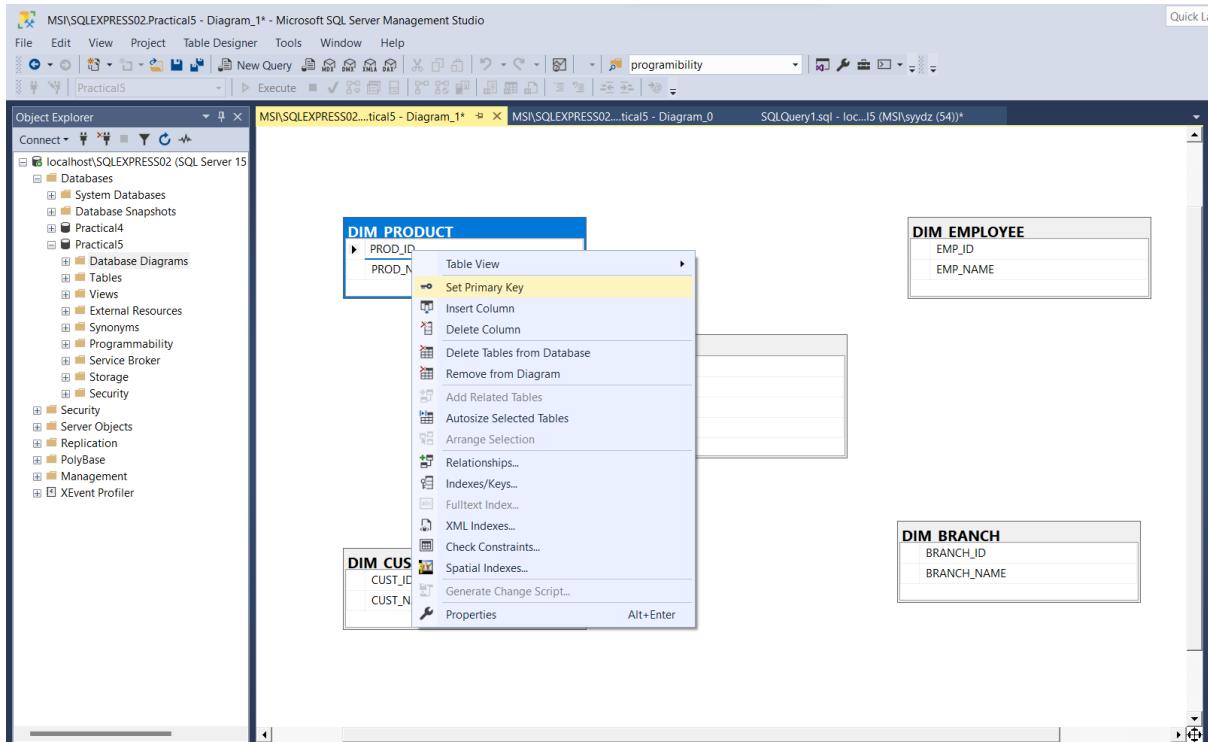
## 6) Create Database diagrams:

Right Click On Database Diagrams > New Database Diagram



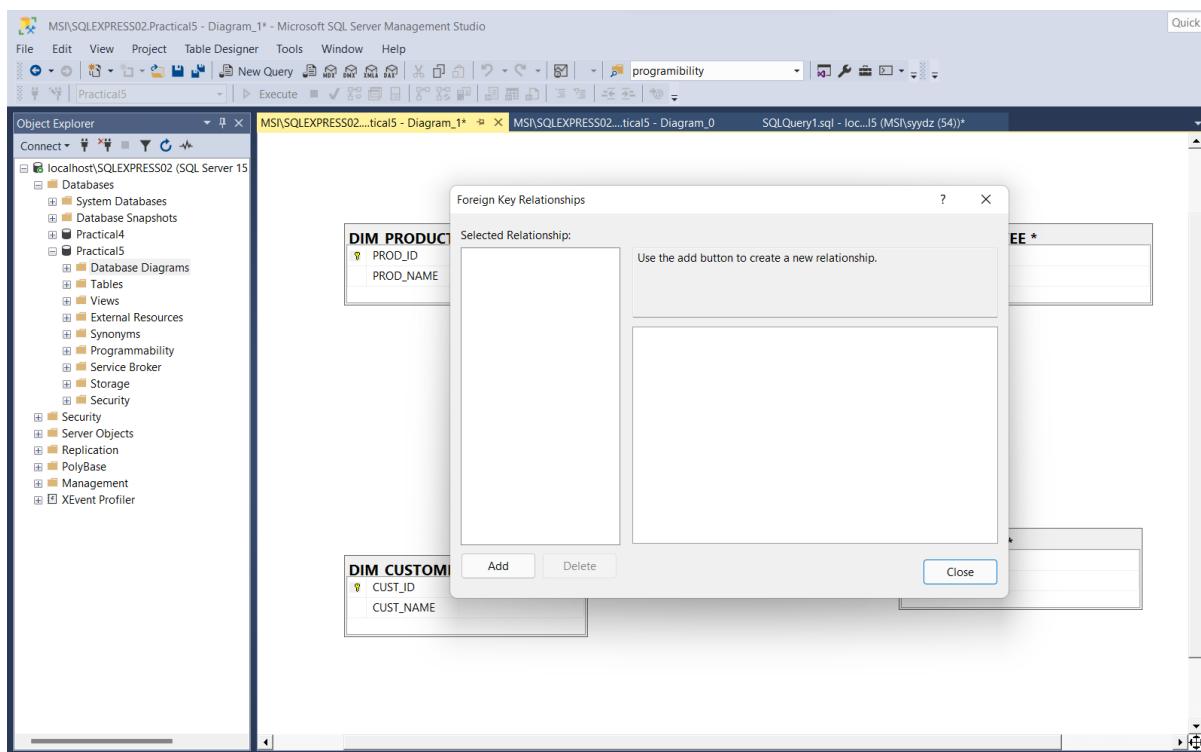
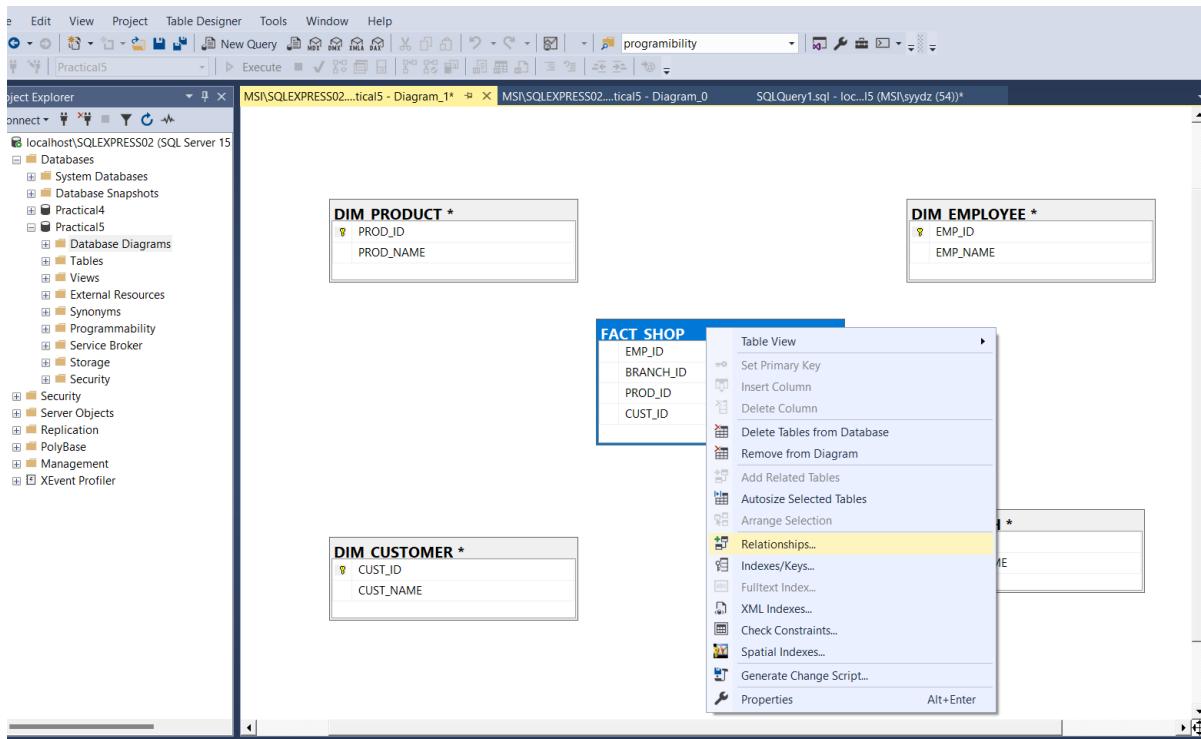
7) Set primary keys :

PROD\_ID, EMP\_ID, CUST\_ID, BRANCH\_ID

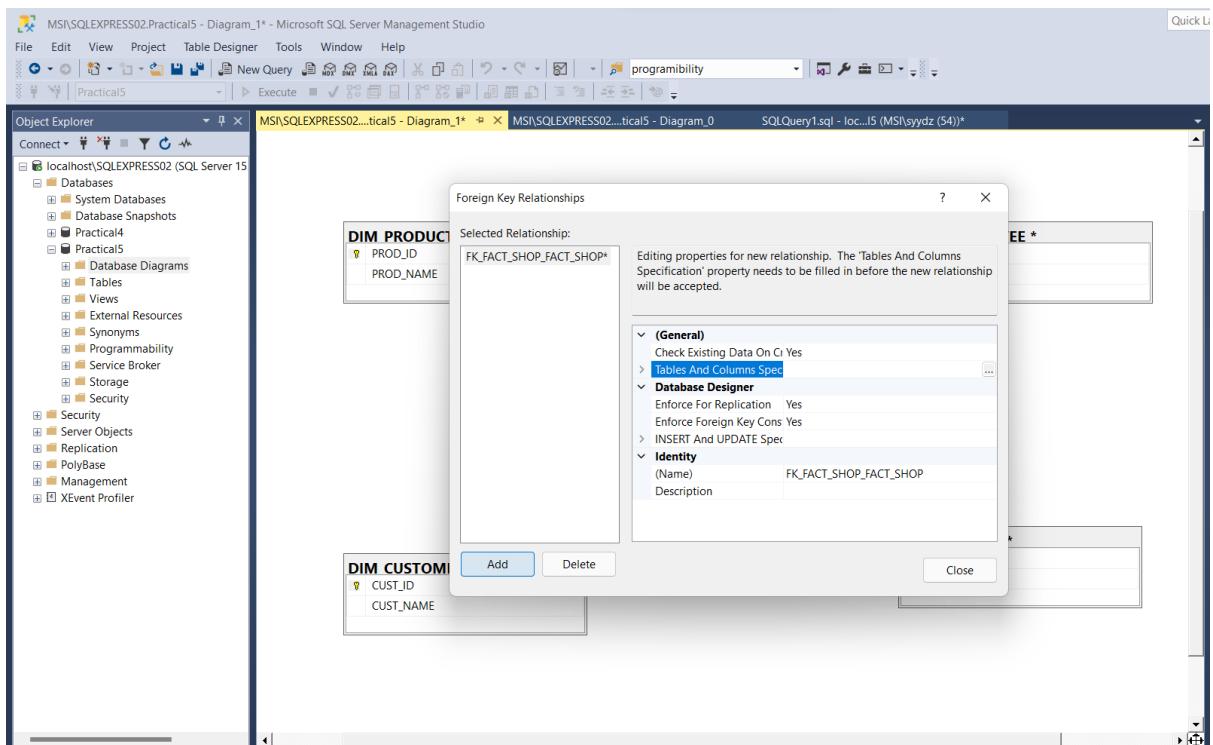


8) Create Relationships:

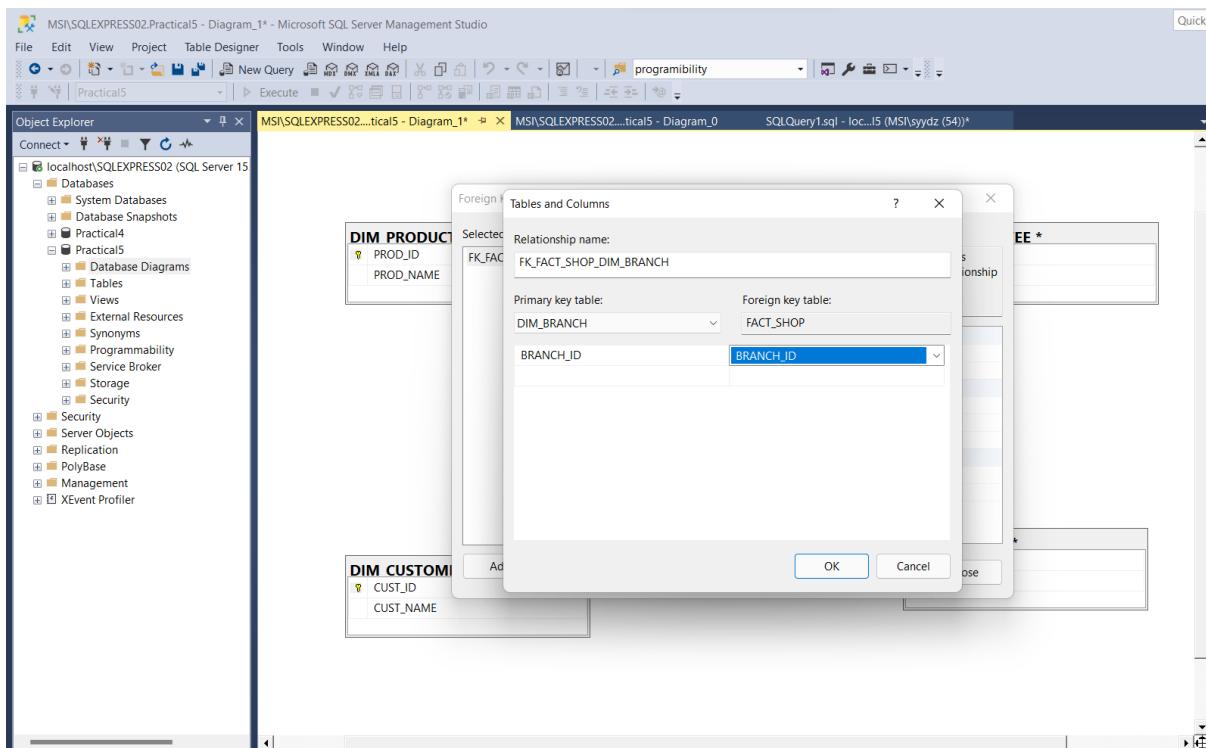
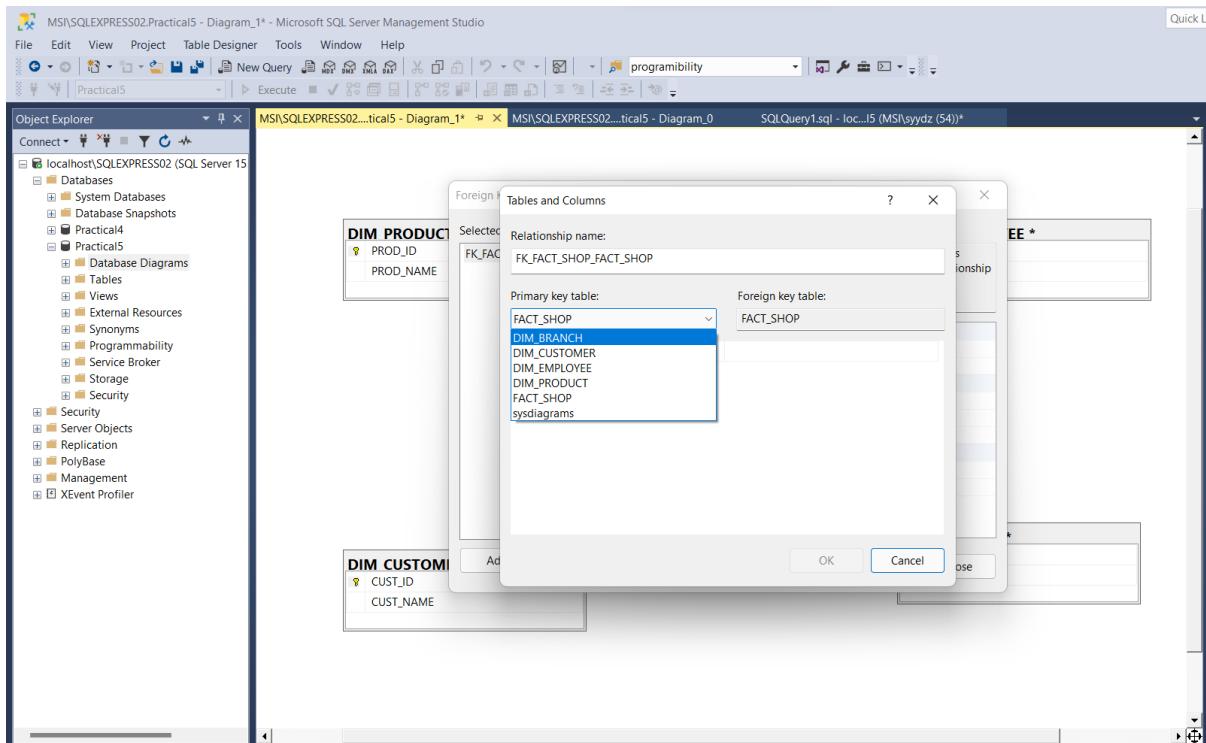
Right Click on Fact Shop table > Select Relationships > Add relations.

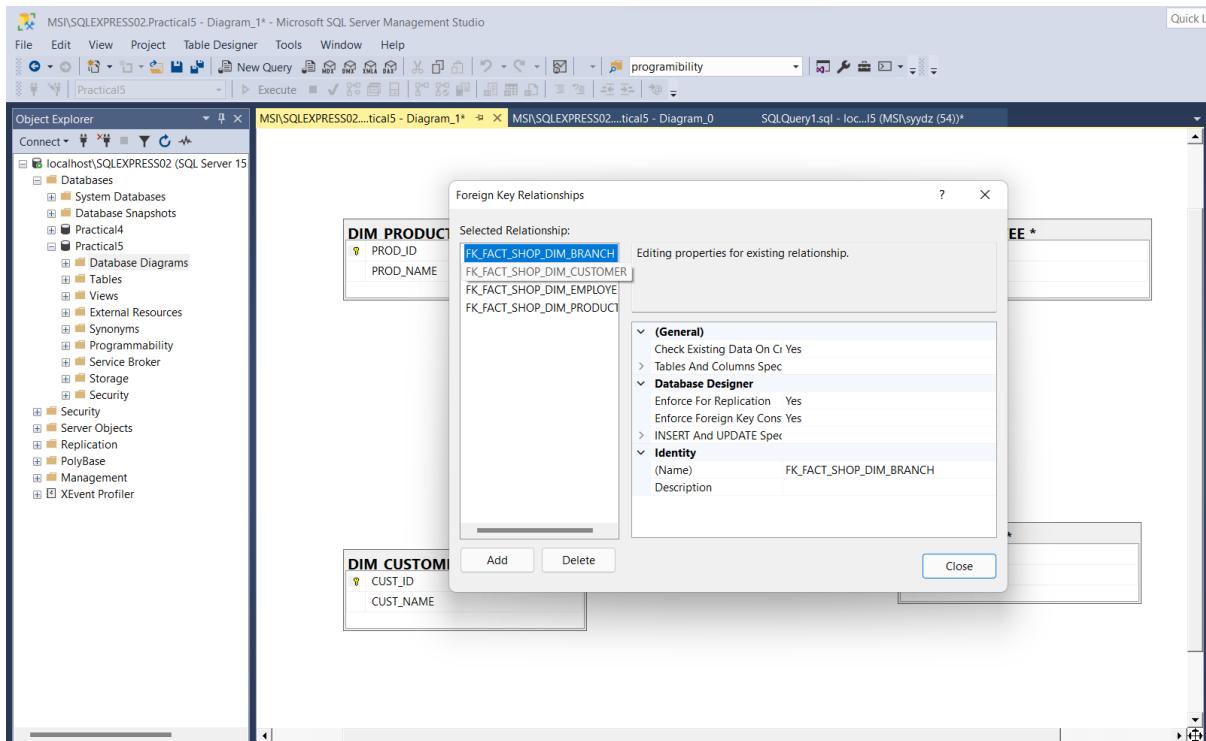


8) Click on Tables And Columns Sec( Expand ...)

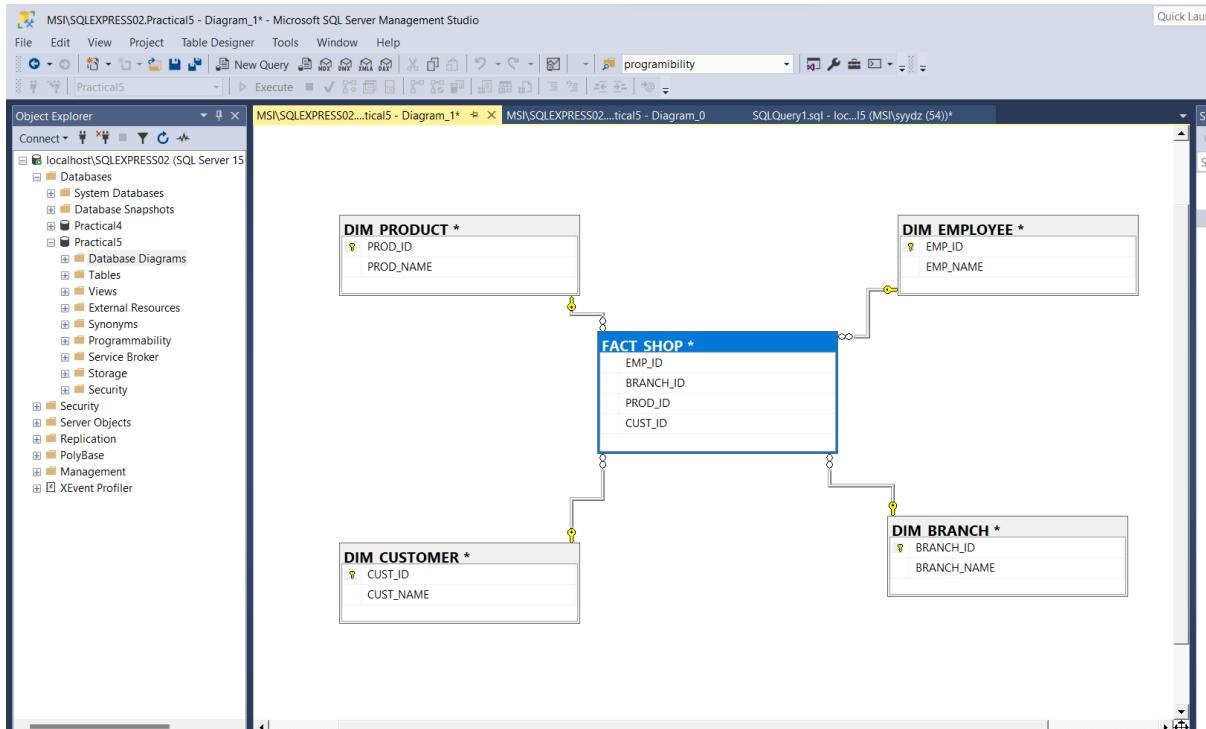


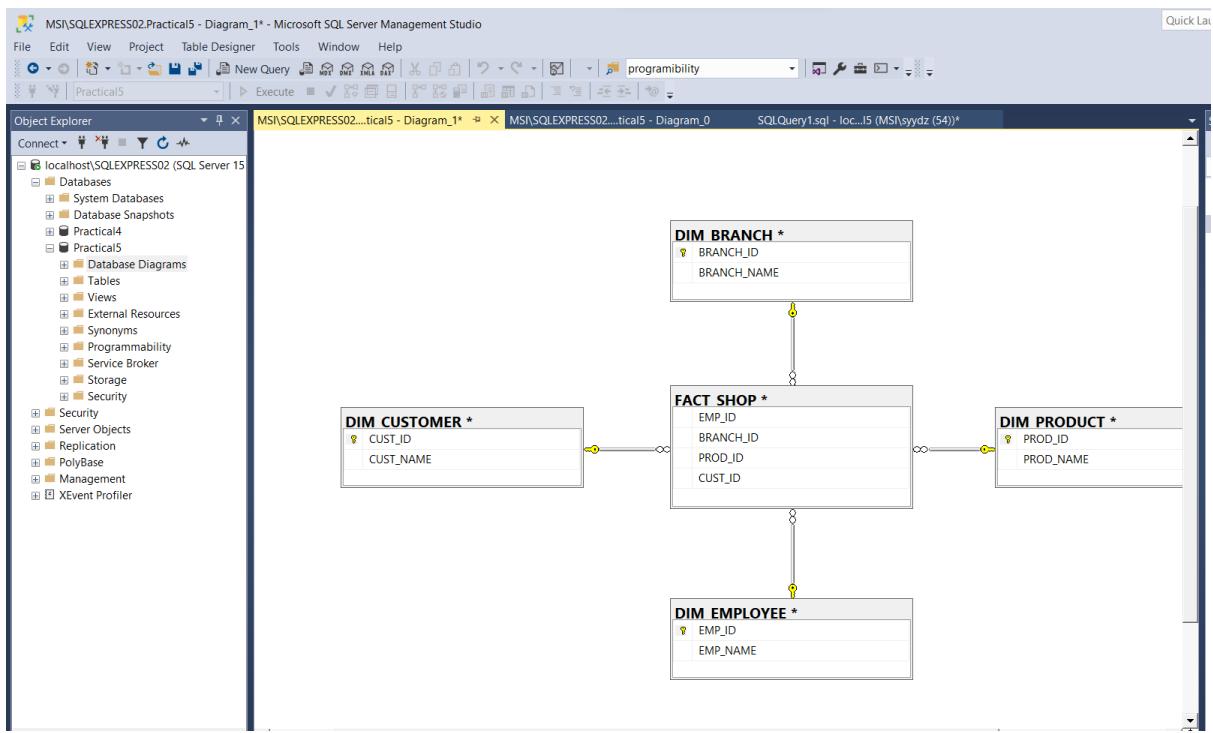
9) Select Primary key tables for every relation





## 10) Close Relationships window





Conclusion: Star schema executed successfull

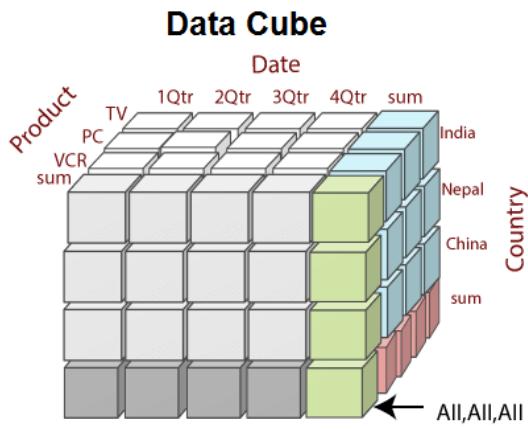
## Practical 4

**Aim:** Create and uses excel pivot table report based on data cube operations.

### Theory:

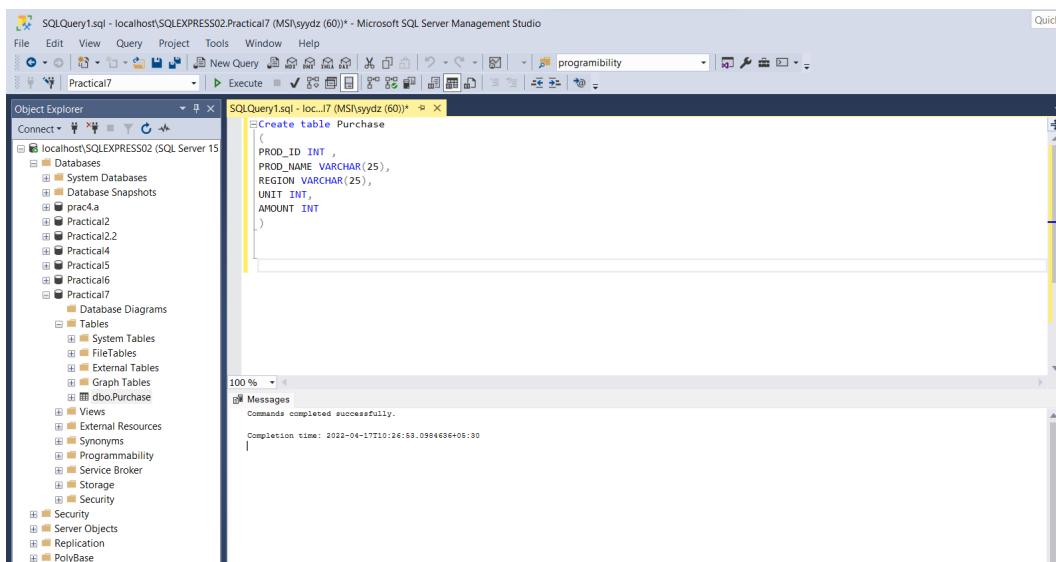
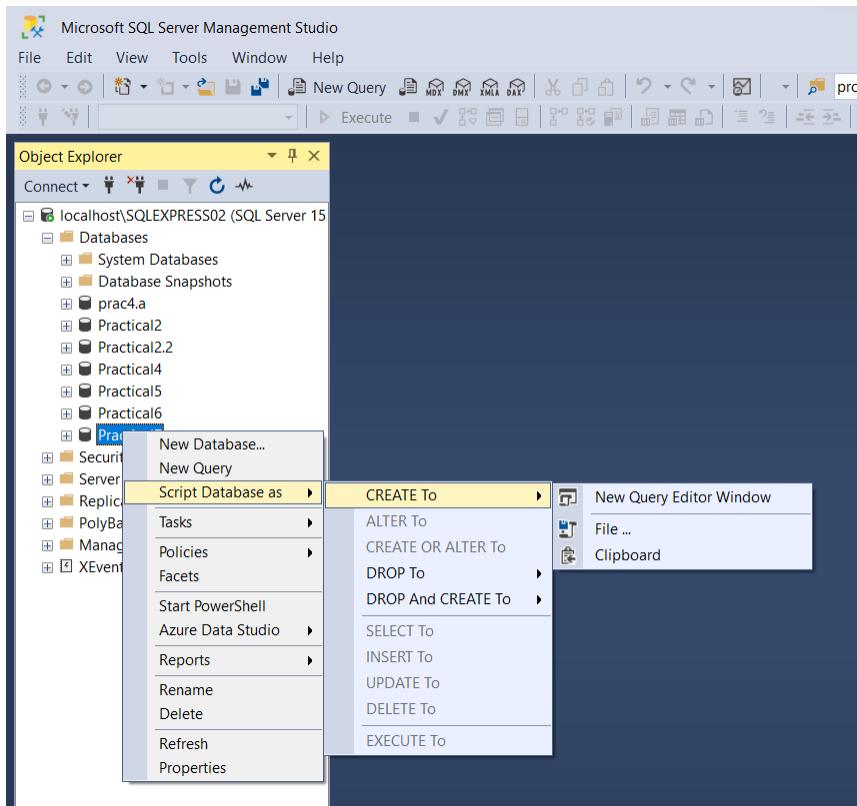
#### What is Data Cube?

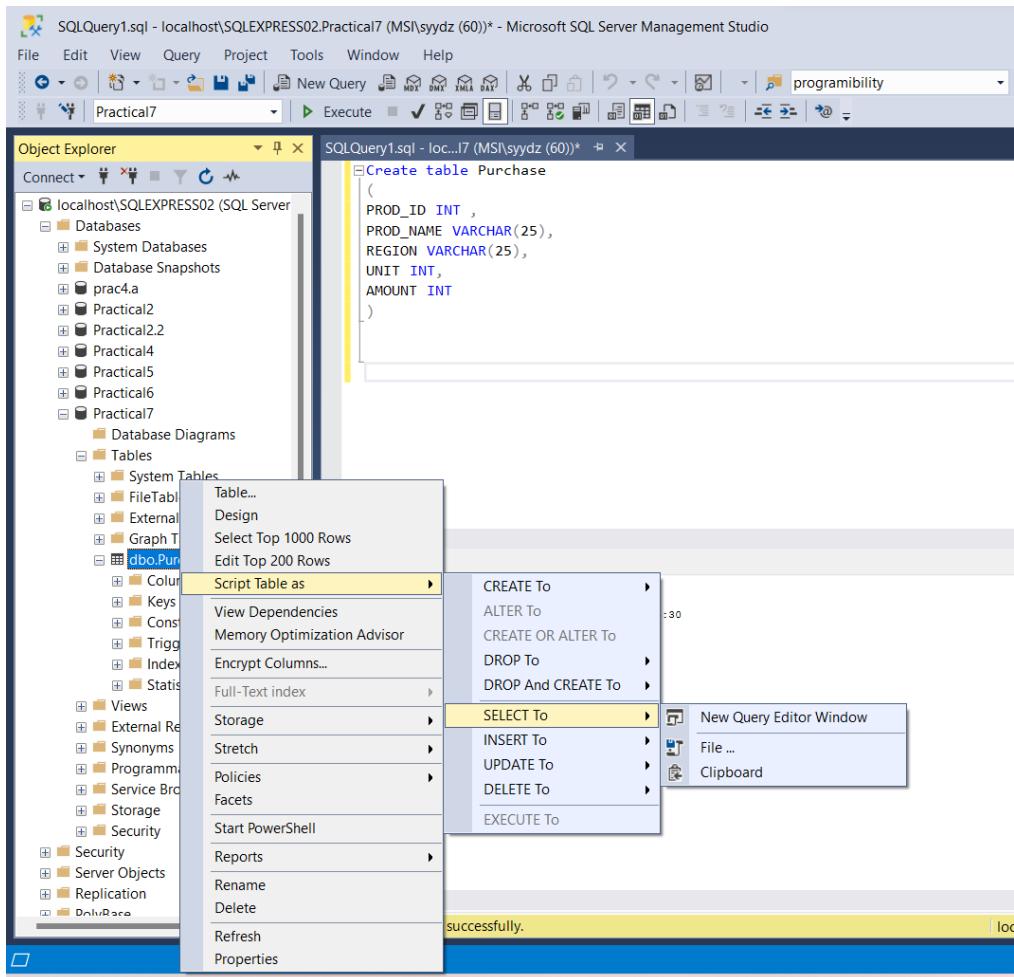
When data is grouped or combined in multidimensional matrices called Data Cubes. The data cube method has a few alternative names or a few variants, such as "Multidimensional databases," "materialized views," and "OLAP (On-Line Analytical Processing)."



Steps:

- 1) Create new database





## 2) Insert data in table

SQLQuery3.sql - localhost\SQLEXPRESS02.Practical7 (MSI\syydz (58))\* - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Object Explorer

localhost\SQLEXPRESS02 (SQL Server)

Databases

Tables

dbo.Purchase

Views

Security

SQLQuery3.sql - loc...l7 (MSI\syydz (58))\* X SQLQuery1.sql - loc...l7 (MSI\syydz (58))\*

```
USE [Practical7]
Insert into[dbo].[Purchase]
Values (
1,
'Honor',
'Mira-road',
20,
20000
)
Insert into[dbo].[Purchase]
Values (
2,
'Samsung',
'Mumbai',
30,
10000
)
Insert into[dbo].[Purchase]
Values (
3,
'Nokia',
'Ireland',
40,
60000
)
Insert into[dbo].[Purchase]
Values (
4,
'LG',
'Budapest',
50,
70000
)
Insert into[dbo].[Purchase]
Values (
5,
'Apple',
'Scotland',
20,
90000
)
```

SQLQuery3.sql - localhost\SQLEXPRESS02.Practical7 (MSI\syydz (58))\* - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Object Explorer

localhost\SQLEXPRESS02 (SQL Server)

Databases

Tables

dbo.Purchase

Views

Security

SQLQuery3.sql - loc...l7 (MSI\syydz (58))\* X SQLQuery1.sql - loc...l7 (MSI\syydz (60))\*

```
USE [Practical7]
Insert into[dbo].[Purchase]
Values (
1,
'Honor',
'Mira-road',
20,
20000
)
Insert into[dbo].[Purchase]
Values (
2,
'Samsung',
'Mumbai',
30,
10000
)
```

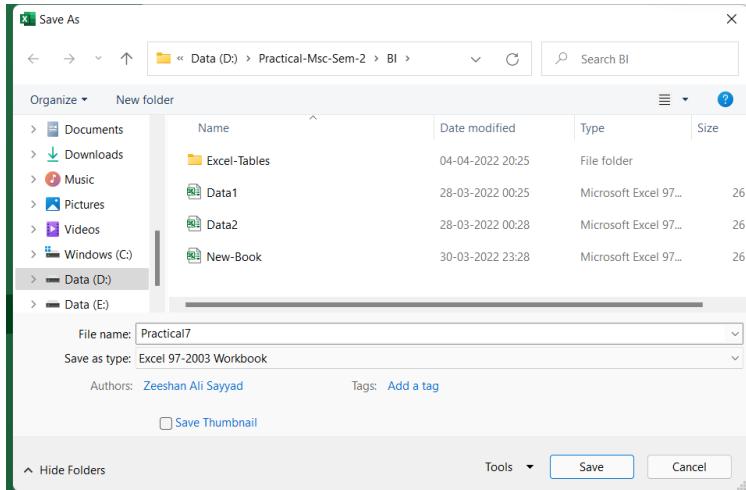
Messages

(1 row affected)

Completion time: 2022-04-17T10:39:42.4071716+05:30

Query executed successfully.

### 3) Create a new Excel File



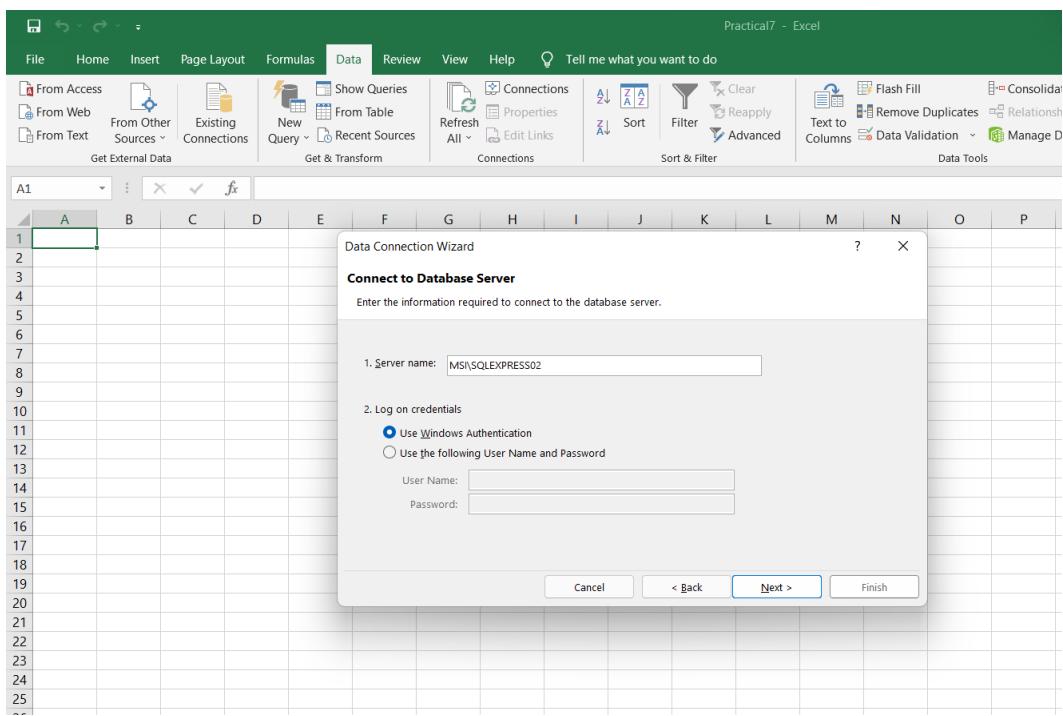
### 4) Go to Data > From Other Sources> Click on From SQL Server

The screenshot shows the Microsoft Excel ribbon with the 'Data' tab selected. In the 'Get Data' section of the ribbon, the 'From Other Sources' button is highlighted, and a dropdown menu is open. The menu lists several options:

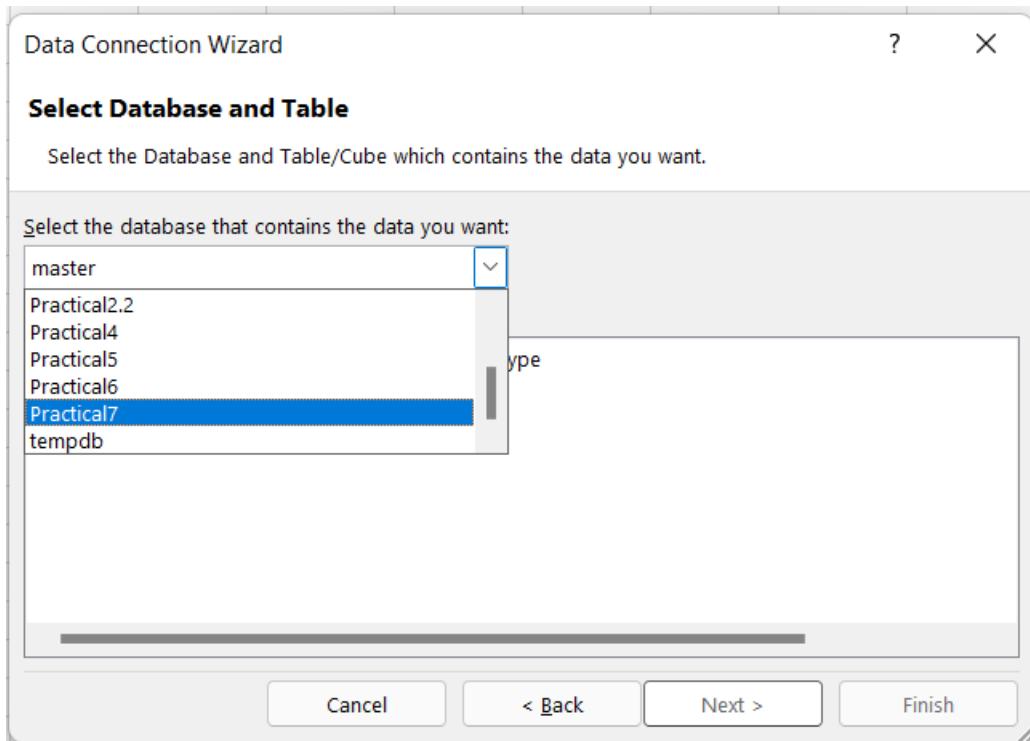
- From Access
- From Web
- From Text
- From SQL Server
- Existing Connections
- New Query
- Recent Sources

The 'From SQL Server' option is currently selected. Its description is displayed below the menu:  
Create a connection to a SQL Server table. Import data into Excel as a Table or PivotTable report.

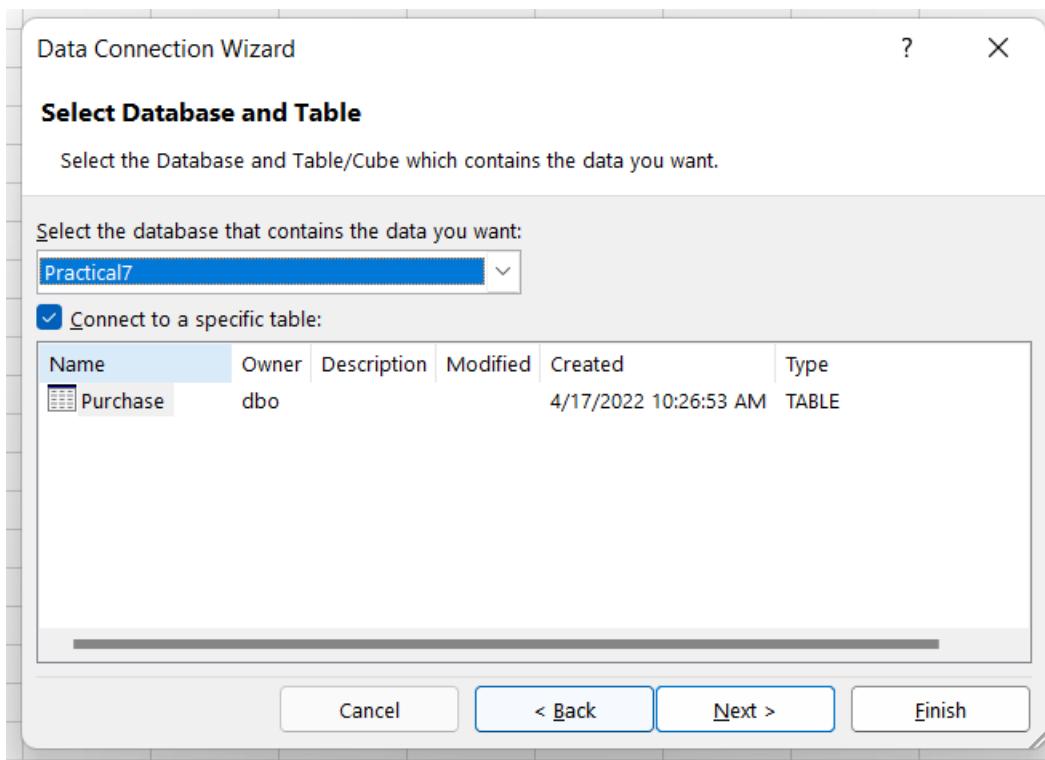
5) Connect to database using server name.



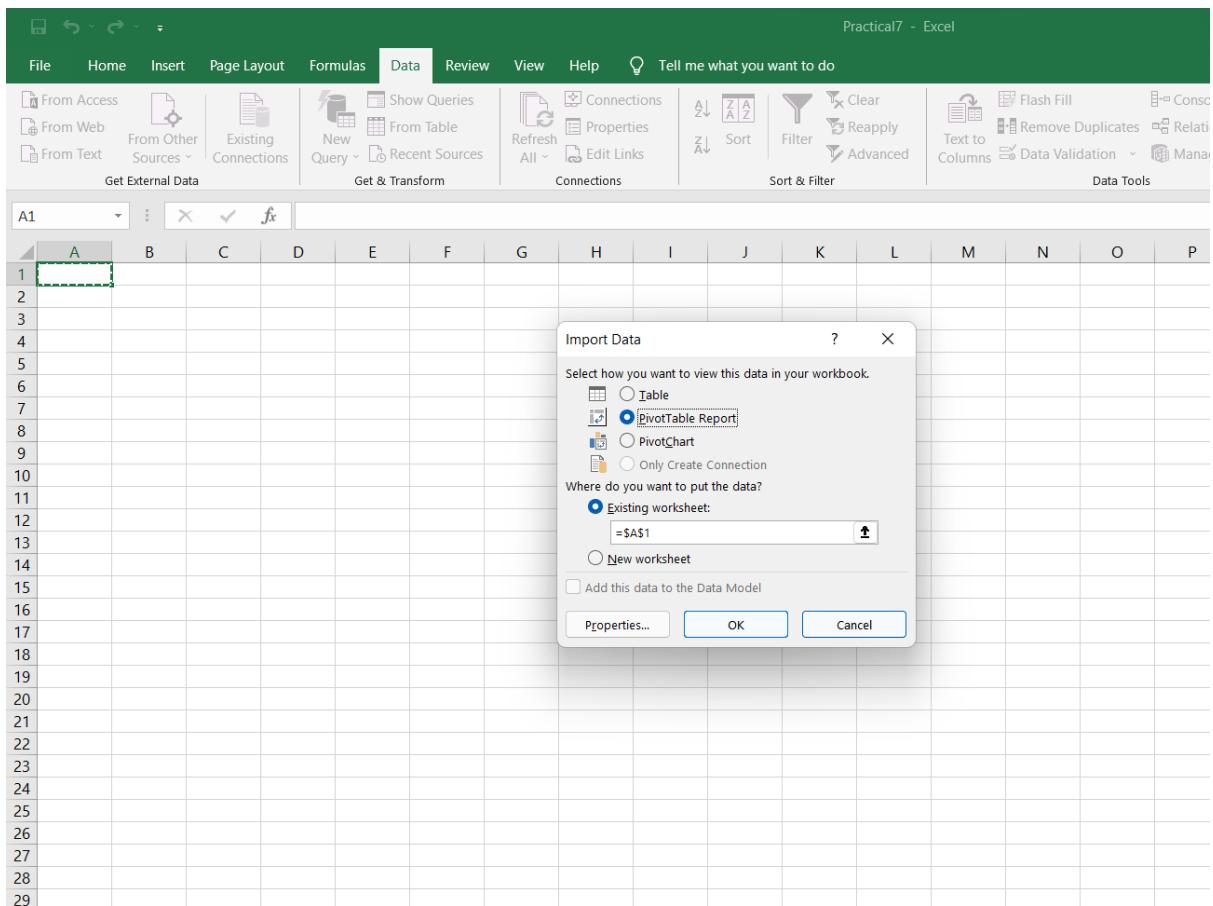
6) Select database and table we want to connect to.



7) Click on Next.



## 8) Select PivotTable report



## 9) Arrange fields PivotTable Fields respectively.

		Column Labels	Sum of AMOUNT	Total Sum of AMOUNT	Total Sum of UNIT
Row Labels		1 2 3 4 5	1 2 3 4 5		
Apple		90000	20	90000	20
Scotland		90000	20	90000	20
Honor	20000		20	20000	20
Mira-road	20000		20	20000	20
LG	70000		50	70000	50
Budapest	70000		50	70000	50
Nokia	60000		40	60000	40
Ireland	60000		40	60000	40
Samsung	10000		30	10000	30
Mumbai	10000		30	10000	30
<b>Grand Total</b>	<b>20000</b>	<b>10000</b>	<b>60000</b>	<b>70000</b>	<b>90000</b>
			20 30 40 50 20	250000	160

Conclusion: Successfully implemented pivot table report based on data cube operations

## Practical 5

**Aim: Do data preprocessing on data obtained from databases imported from external sources.**

**Theory:**

### What is data preprocessing?

Data preprocessing, a component of data preparation, describes any type of processing performed on raw data to prepare it for another data processing procedure. It has traditionally been an important preliminary step for the data mining process. More recently, data preprocessing techniques have been adapted for training machine learning models and AI models and for running inferences against them.

Data preprocessing transforms the data into a format that is more easily and effectively processed in data mining, machine learning and other data science tasks. The techniques are generally used at the earliest stages of the machine learning and AI development pipeline to ensure accurate results.

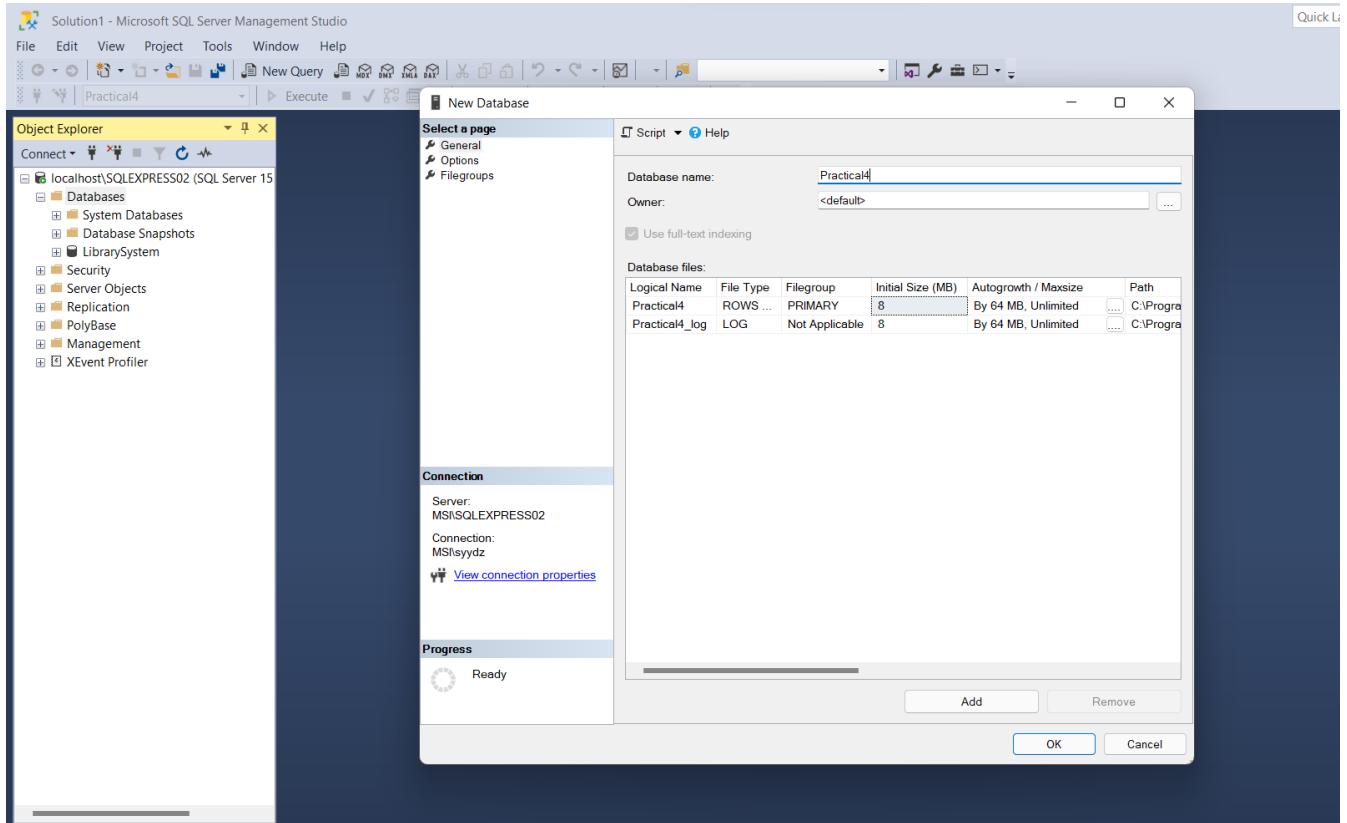
There are several different tools and methods used for preprocessing data, including the following:

- sampling, which selects a representative subset from a large population of data;
- transformation, which manipulates raw data to produce a single input;
- denoising, which removes noise from data;
- imputation, which synthesizes statistically relevant data for missing values;
- normalization, which organizes data for more efficient access; and
- feature extraction, which pulls out a relevant feature subset that is significant in a particular context.
- pulls out a relevant feature subset that is significant in a particular context.

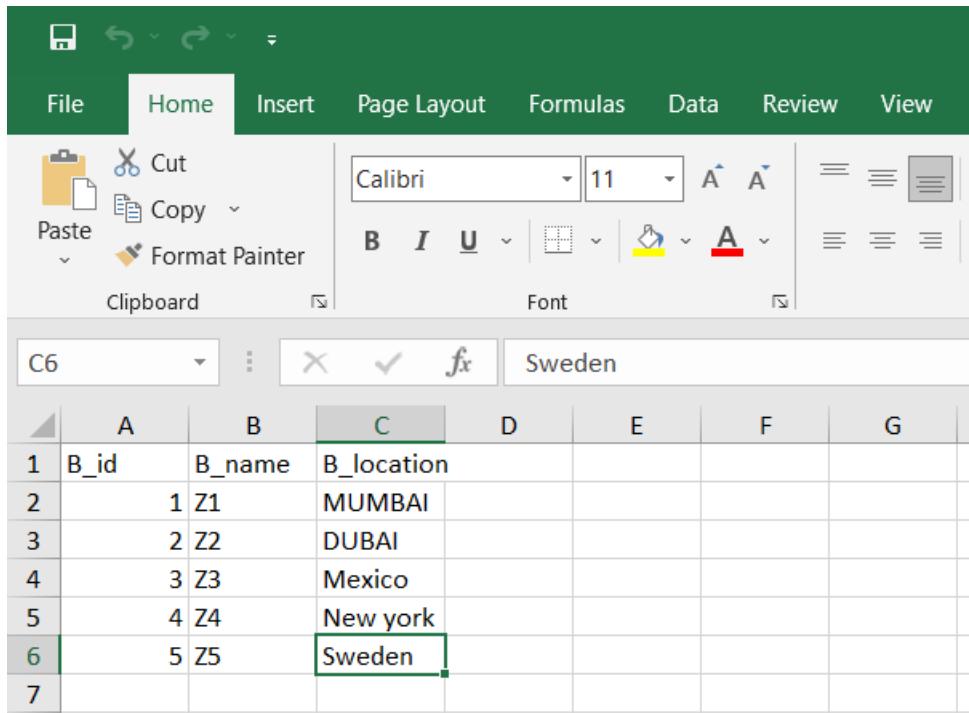
#Steps:

1) Start Microsoft SQL server

2) Right Click On Databases > New Database > Give Database Name > Click Ok



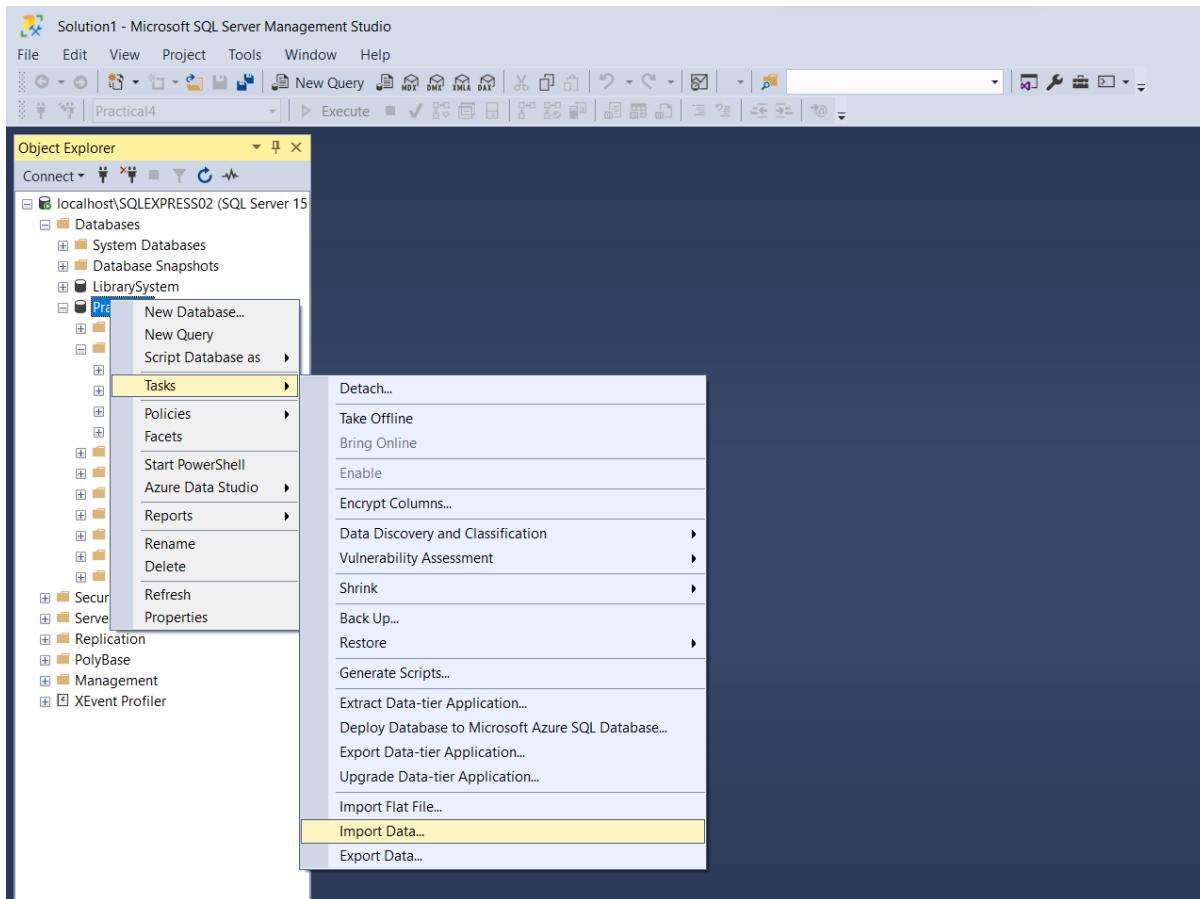
3) Create a new Microsoft Excel 97-2003 Worksheet (.xls) with columns Branch id, Branch name, Branch location and enter some data into rows.

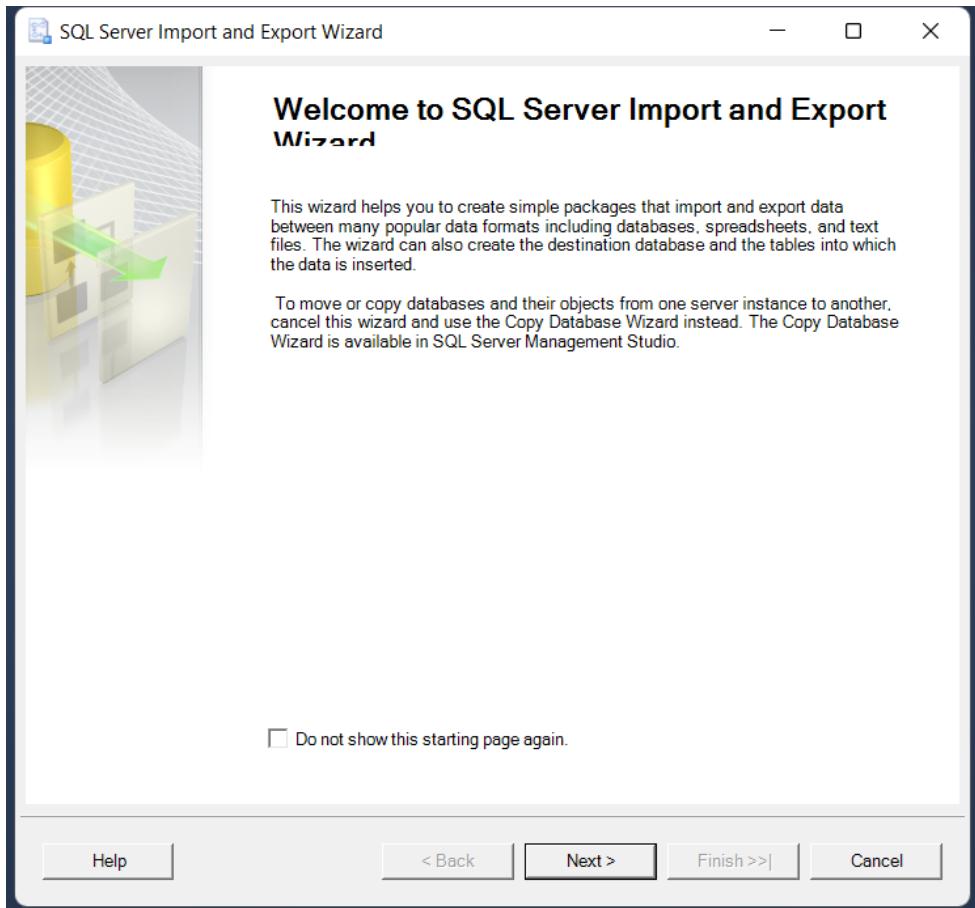


A screenshot of Microsoft Excel showing a table with data in rows 1 through 6. Row 6 is selected, highlighted with a green border. The table has columns A, B, and C. The data is as follows:

	A	B	C	D	E	F	G
1	B_id	B_name	B_location				
2	1	Z1	MUMBAI				
3	2	Z2	DUBAI				
4	3	Z3	Mexico				
5	4	Z4	New york				
6	5	Z5	Sweden				
7							

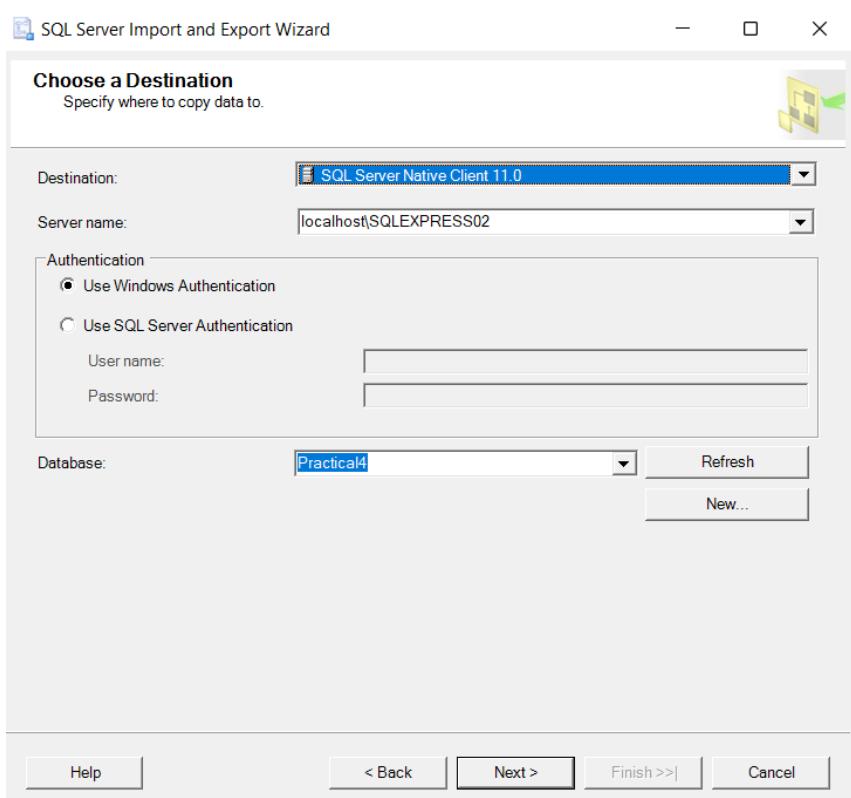
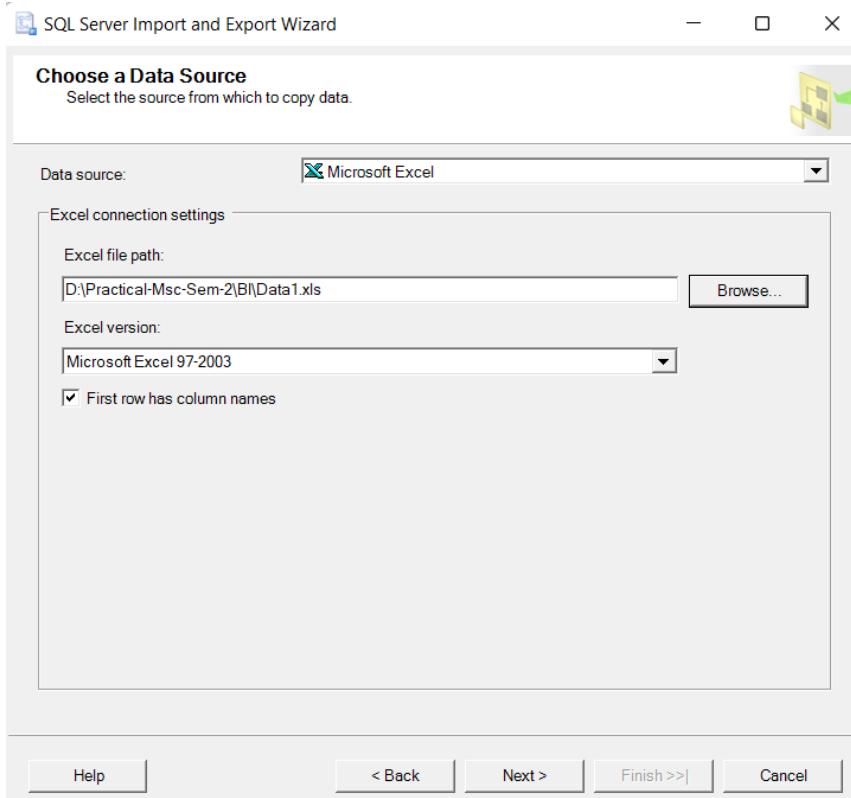
4) Right Click On Practical4 > Tasks > Select: Import Data

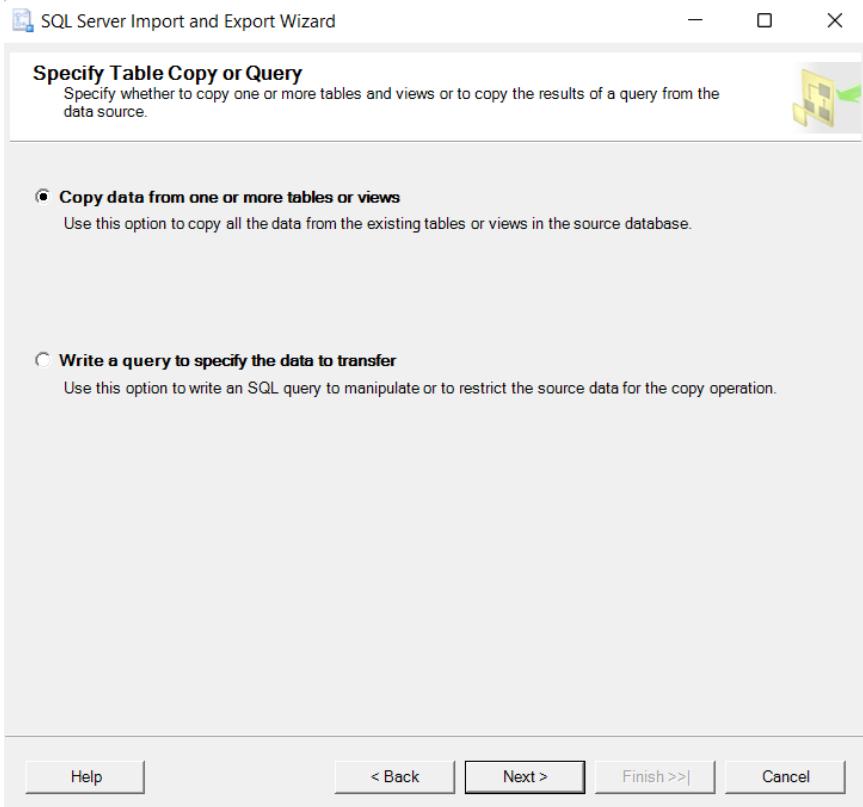


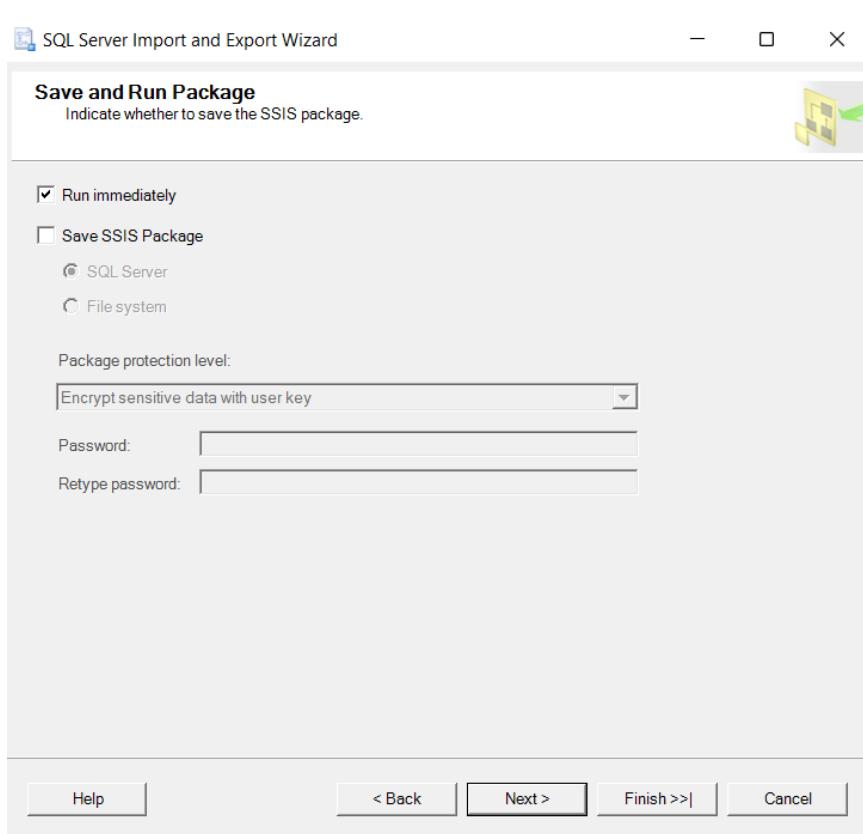
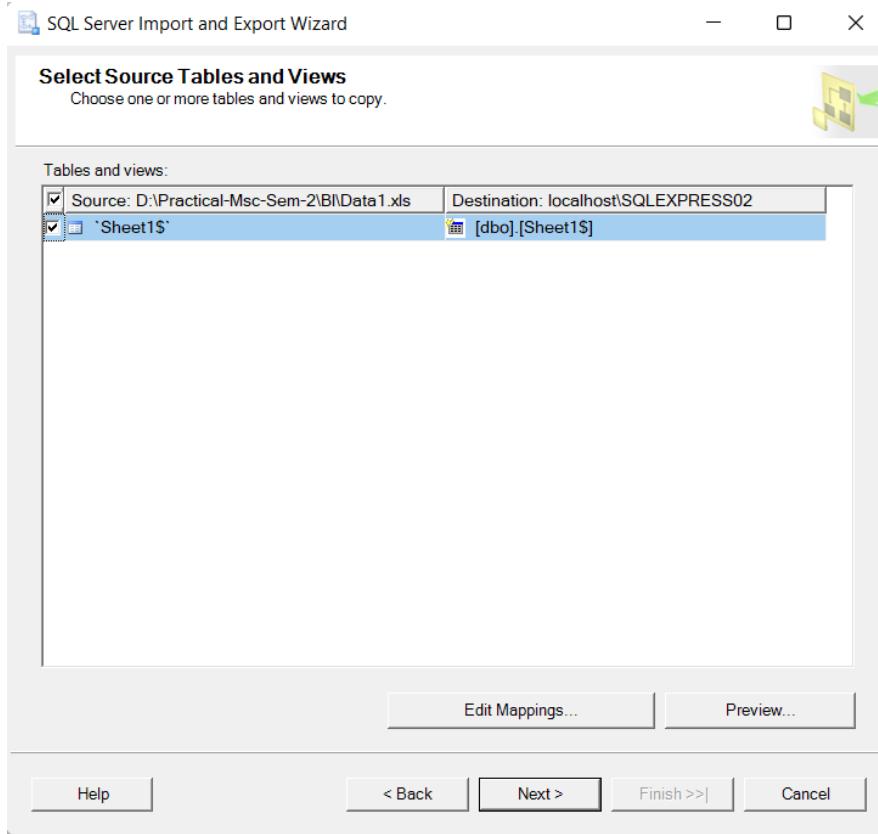


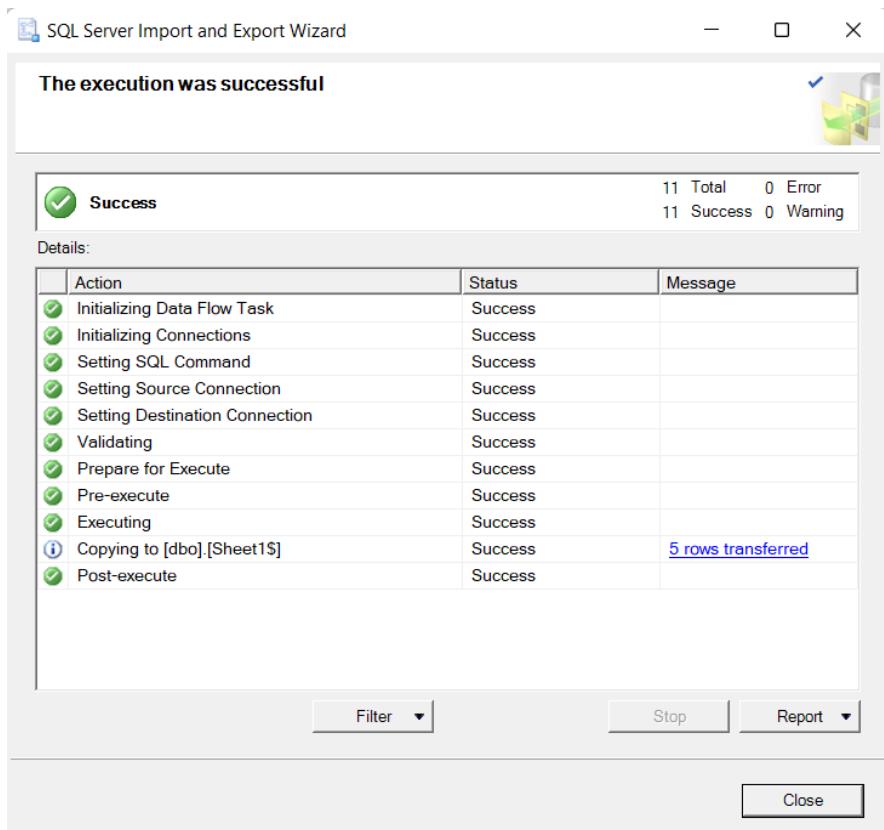
**5) Click on Next > Select Data Source as: Microsoft Excel > Give Excel File Path > Select Destination as: SQL Server Native Client 11.0 > Click Next > Select Sheet 1 > Click Next > Click Finish > Close the window after successful execution.**

**Note: Close the excel file before using it as Source File**

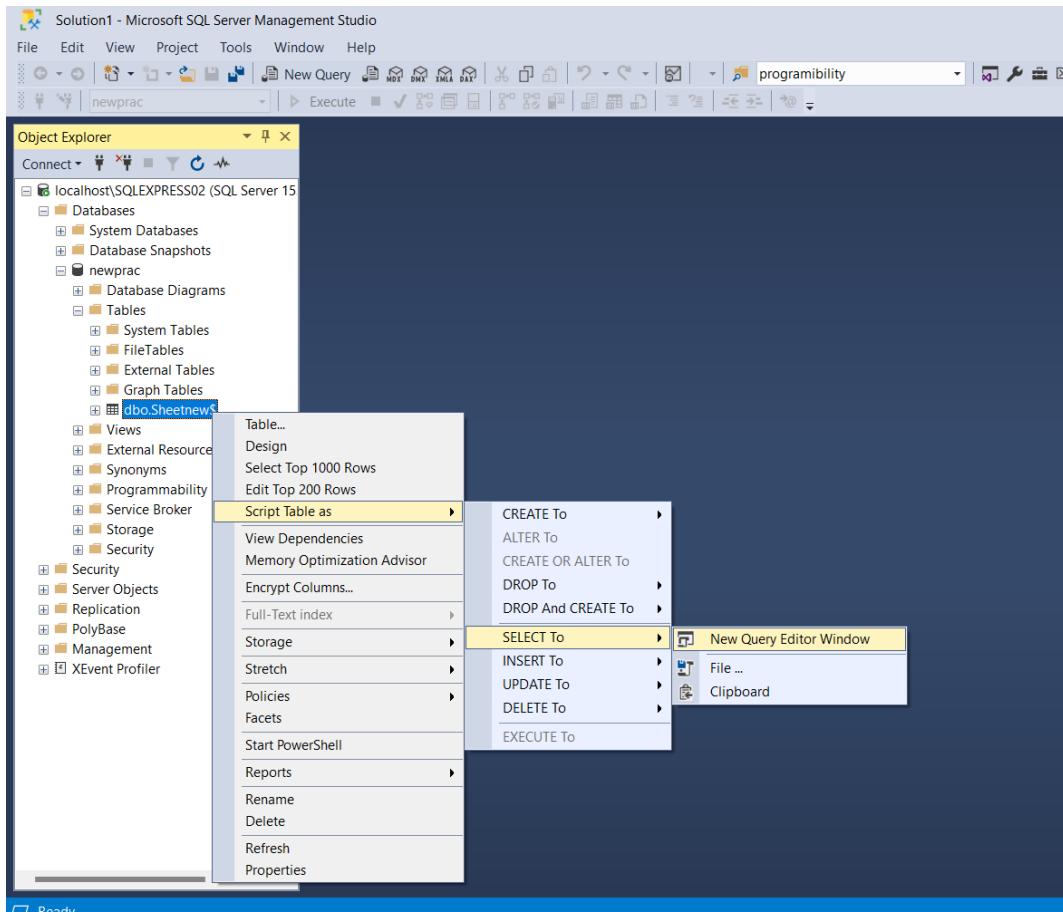








6) Now expand the Tables Section > Right Click on dbo.Sheet\$1 > Script table as > Select To > Click on New Query Editor Window.



SQLQuery2.sql - localhost\SQLEXPRESS02.Practical4 (MSI\syydz (53)) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

localhost\SQLEXPRESS02 (SQL Server 15)

Databases System Databases Database Snapshots LibrarySystem Practical4

Tables Database Diagrams

System Tables FileTables External Tables Graph Tables

Views External Resources Synonyms Programmability Service Broker Storage Security

Server Objects Replication PolyBase Management XEvent Profiler

Object Explorer

Connect ▾

localhost\SQLEXPRESS02 (SQL Server 15)

Databases System Databases Database Snapshots LibrarySystem Practical4

Tables Database Diagrams

System Tables FileTables External Tables Graph Tables

Views External Resources Synonyms Programmability Service Broker Storage Security

Server Objects Replication PolyBase Management XEvent Profiler

SQLQuery2.sql - loc...J4 (MSI\syydz (53))

```
USE [Practical4]
GO

/***** Object: Table [dbo].[Sheet1$]    Script Date: 28-03-2022 00:04:58 *****/
SET ANSI_NULLS ON
GO

SET QUOTED_IDENTIFIER ON
GO

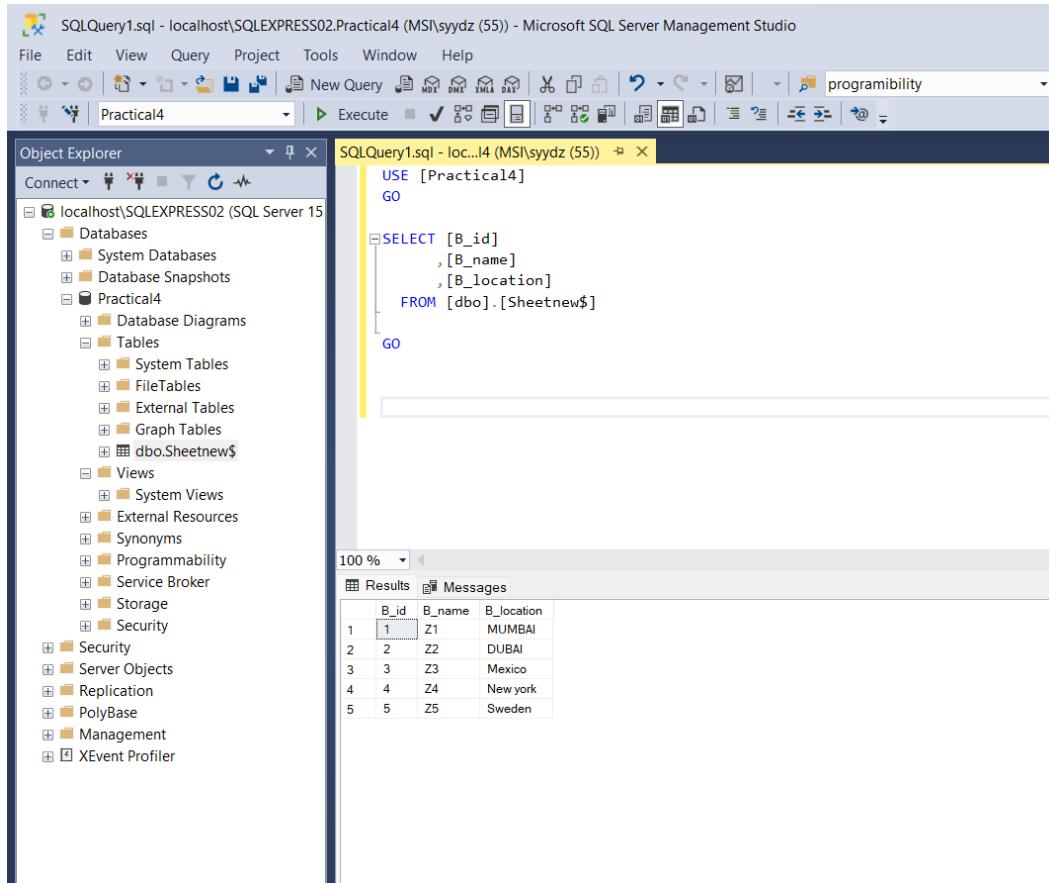
CREATE TABLE [dbo].[Sheet1$](
    [B_id] [float] NULL,
    [B_name] [nvarchar](255) NULL,
    [B_location] [nvarchar](255) NULL
) ON [PRIMARY]
GO
```

100 %

Connected. (1/1)

localhost\SQLEXPRESS02 (15...) MSI\syydz (53) Practical4 000000 0 rows

## 7) Click on Execute



The screenshot shows the Microsoft SQL Server Management Studio interface. The Object Explorer on the left shows the database structure for 'localhost\SQLEXPRESS02'. The 'Practical4' database is selected, and under 'Tables', a table named 'dbo.Sheetnew\$' is listed. The 'Script' button next to it is highlighted. The central pane displays a query window titled 'SQLQuery1.sql - loc...l4 (MSI\syydz (55))'. The query is:

```
USE [Practical4]
GO

SELECT [B_id]
      ,[B_name]
      ,[B_location]
 FROM [dbo].[Sheetnew$]
GO
```

The 'Results' tab at the bottom shows the output of the query:

	B_id	B_name	B_location
1	1	Z1	MUMBAI
2	2	Z2	DUBAI
3	3	Z3	Mexico
4	4	Z4	Newyork
5	5	Z5	Sweden

**Conclusion:** Data pre-processing on imported data from external sources implemented Successfully

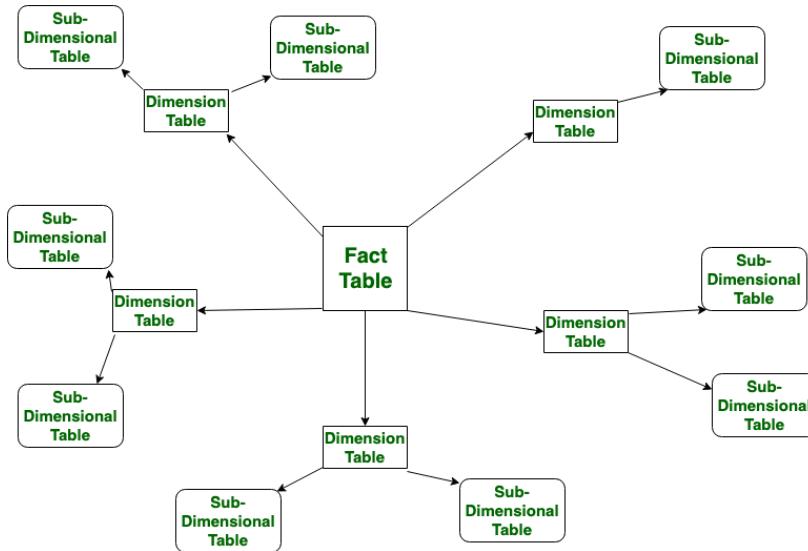
## Practical 6

**Aim:** Design and create cube by identifying measures and dimensions for snow flake schema.

**Theory:**

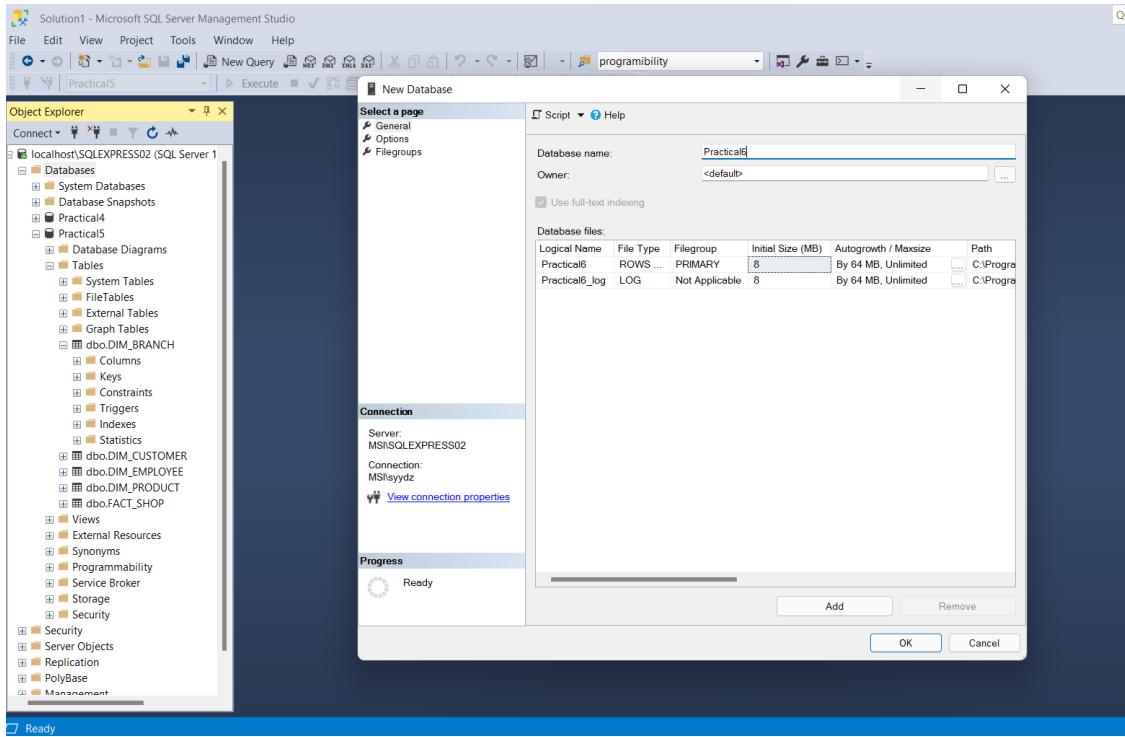
### Snowflake Schema:

Snowflake Schema is also the type of multidimensional model which is used for data warehouse. In snowflake schema, The fact tables, dimension tables as well as sub dimension tables are contained. This schema forms a snowflake with fact tables, dimension tables as well as sub-dimension tables.

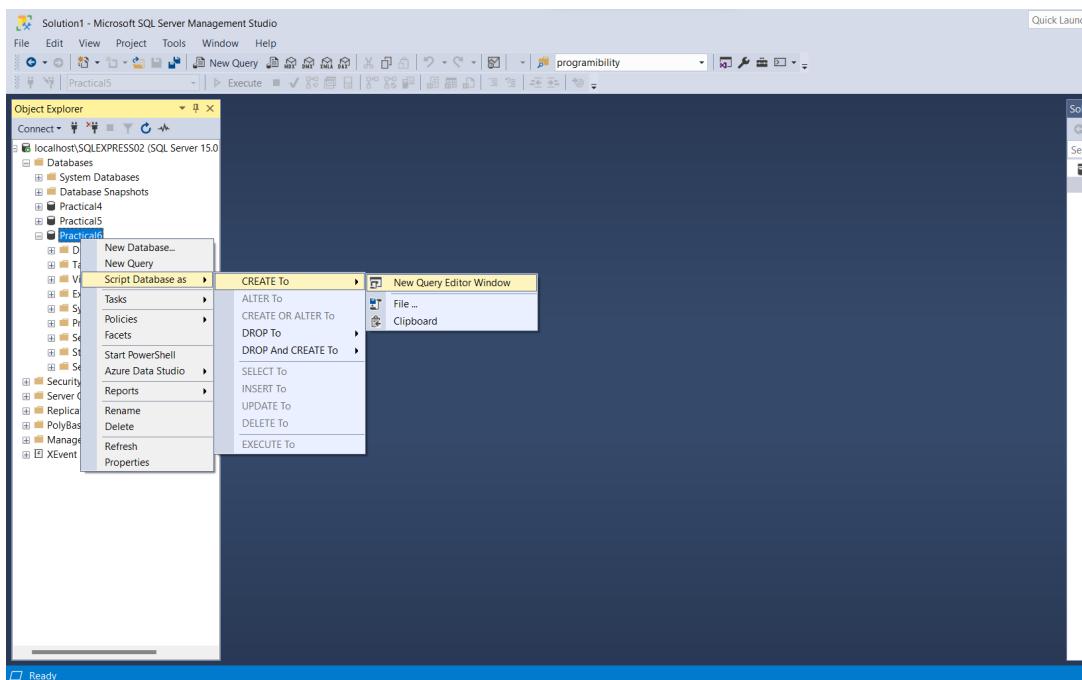


Steps:

- 1) Create database.



## 2) Script Database as > Create To > New Query Window

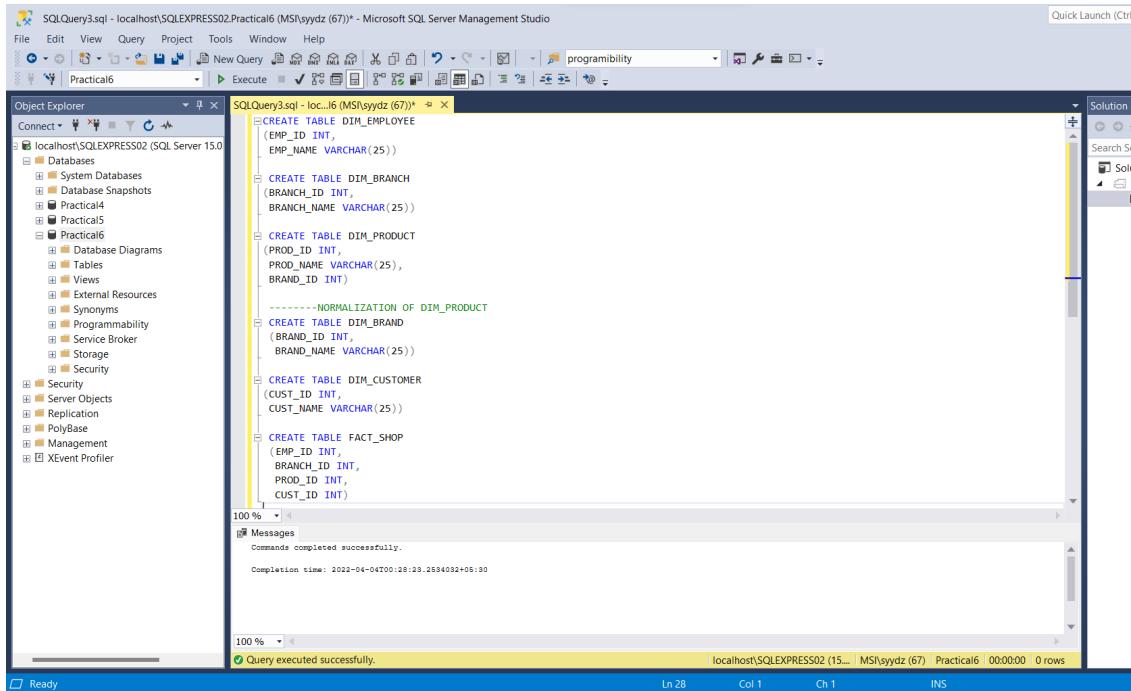


## 3) Create tables:

```
CREATE TABLE DIM_EMPLOYEE
```

```
(EMP_ID INT,  
EMP_NAME VARCHAR(25))
```

```
CREATE TABLE DIM_BRANCH  
(BRANCH_ID INT,  
BRANCH_NAME VARCHAR(25))  
CREATE TABLE DIM_PRODUCT  
(PROD_ID INT,  
PROD_NAME VARCHAR(25),  
BRAND_ID INT)  
CREATE TABLE DIM_BRAND  
(BRAND_ID INT,  
BRAND_NAME VARCHAR(25))  
CREATE TABLE DIM_CUSTOMER  
(CUST_ID INT,  
CUST_NAME VARCHAR(25))  
CREATE TABLE FACT_SHOP  
(EMP_ID INT,  
BRANCH_ID INT,  
PROD_ID INT,  
CUST_ID INT)
```



#### 4) Create New database diagram

The screenshot shows the Microsoft SQL Server Management Studio interface. The title bar reads "SQLQuery3.sql - localhost\SQLEXPRESS02.Practical6 (MSI\syydz (67))\* - Microsoft SQL Server Management Studio". The menu bar includes File, Edit, View, Query, Project, Tools, Window, and Help. The toolbar has various icons for connecting, executing queries, and managing databases. The Object Explorer on the left shows the database structure under "localhost\SQLEXPRESS02 (SQL Server 15.0)". A context menu is open over the "Practical6" database node, with "New Database Diagram" highlighted. The main pane displays a query window titled "SQLQuery3.sql - loc...l6 (MSI\syydz (67))\*". The query contains several CREATE TABLE statements:

```
CREATE TABLE DIM_EMPLOYEE
(EMP_ID INT,
EMP_NAME VARCHAR(25))

CREATE TABLE DIM_BRANCH
(BRANCH_ID INT,
BRANCH_NAME VARCHAR(25))

CREATE TABLE DIM_PRODUCT
PROD_ID INT,
PROD_NAME VARCHAR(25),
BRAND_ID INT)

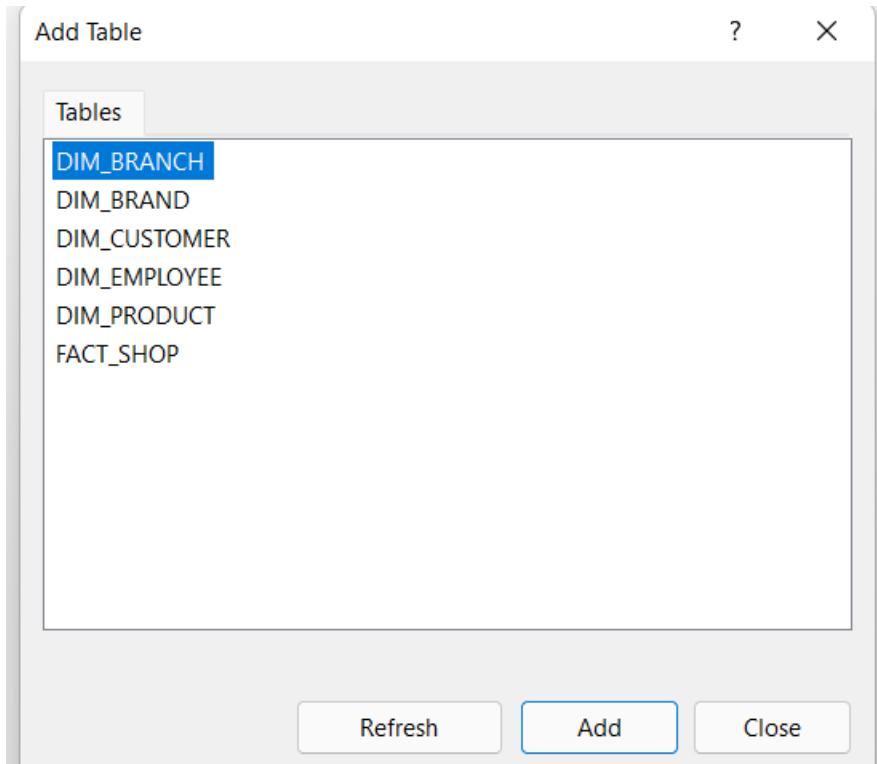
-----NORMALIZATION OF DIM_PRODUCT
CREATE TABLE DIM_BRAND
(BRAND_ID INT,
BRAND_NAME VARCHAR(25))

CREATE TABLE DIM_CUSTOMER
(CUST_ID INT,
CUST_NAME VARCHAR(25))

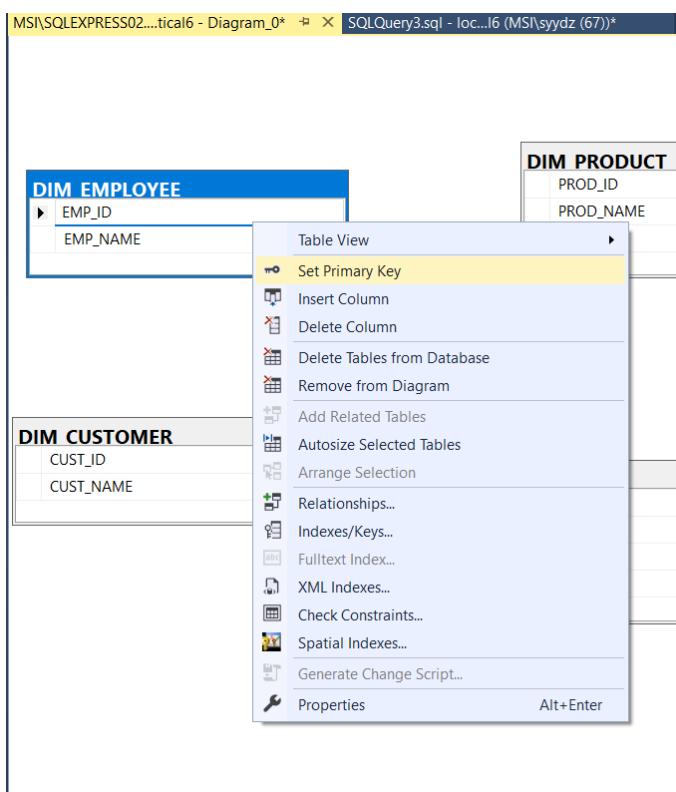
CREATE TABLE FACT_SHOP
(EMP_ID INT,
BRANCH_ID INT,
PROD_ID INT,
CUST_ID INT)
```

The status bar at the bottom shows "Commands completed successfully." and "Completion time: 2022-04-04T00:28:23.2534032+05:30".

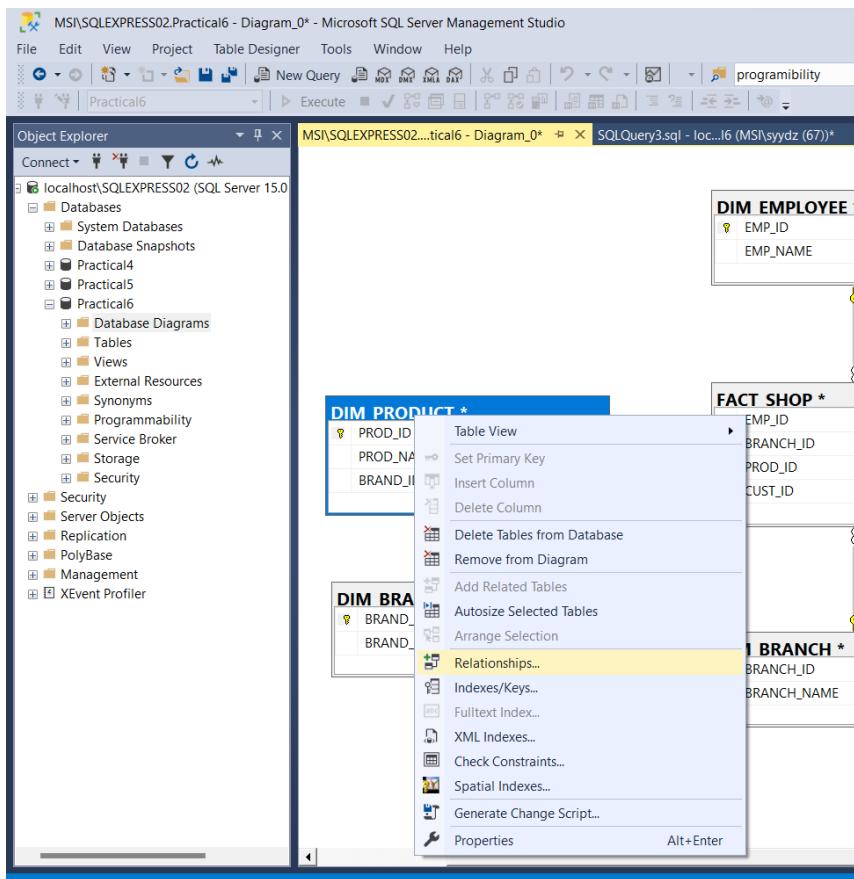
5) Add tables



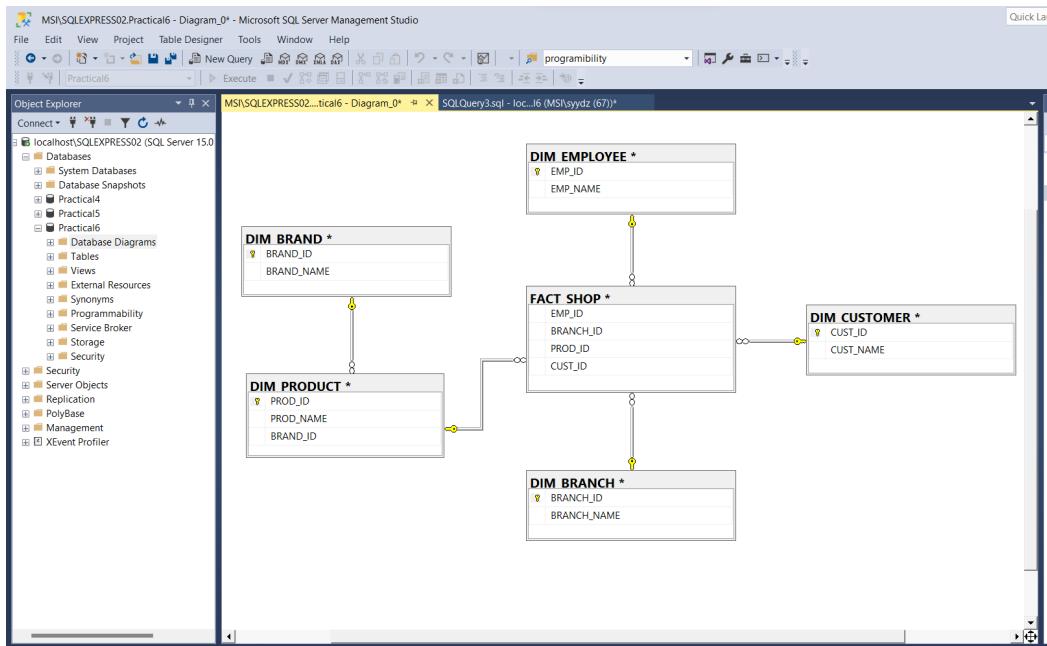
6) Set Primary key for all tables:



7) Add relationship of tables with fact shop.



8) Add relationship of product and brand .



Conclusion: Snowflake schema implemented successfully.

## Practical 7

**Aim: Develop an application to construct a multidimensional data.**

**Theory:**

**What is Multi-Dimensional Data Model?**

A multidimensional model views data in the form of a data-cube. A data cube enables data to be modeled and viewed in multiple dimensions. It is defined by dimensions and facts.

The dimensions are the perspectives or entities concerning which an organization keeps records. For example, a shop may create a sales data warehouse to keep records of the store's sales for the dimension time, item, and location. These dimensions allow the save to keep track of things, for example, monthly sales of items and the locations at which the items were sold. Each dimension has a table related to it, called a dimensional table, which describes the dimension further. For example, a dimensional table for an item may contain the attributes item\_name, brand, and type.

A multidimensional data model is organized around a central theme, for example, sales. This theme is represented by a fact table. Facts are numerical measures. The fact table contains the names of the facts or measures of the related dimensional tables.

#Steps:

1) Create new database.

```
CREATE TABLE DIM_customer
(
    custid VARCHAR(6),
    fname VARCHAR(30),
    mname VARCHAR(30),
    lname VARCHAR(30),
    city VARCHAR(15),
    mobileno VARCHAR(10),
    occupation VARCHAR(10),
    dob DATE
)

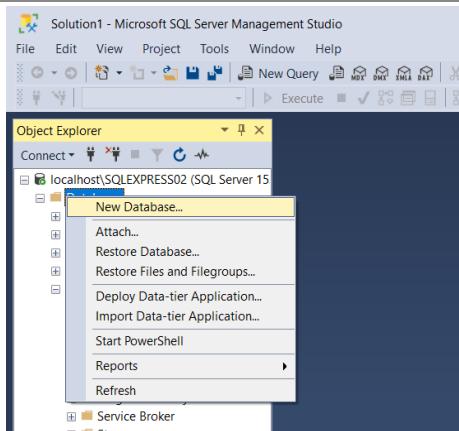
CREATE TABLE DIM_branch
(
    bid VARCHAR(6),
    bname VARCHAR(30),
    bcity VARCHAR(30),
)

CREATE TABLE DIM_account
(
    acnumber VARCHAR(6),
    custid VARCHAR(6),
    bid VARCHAR(6),
    opening_balance VARCHAR(7),
    aod DATE,
    atype VARCHAR(10),
```

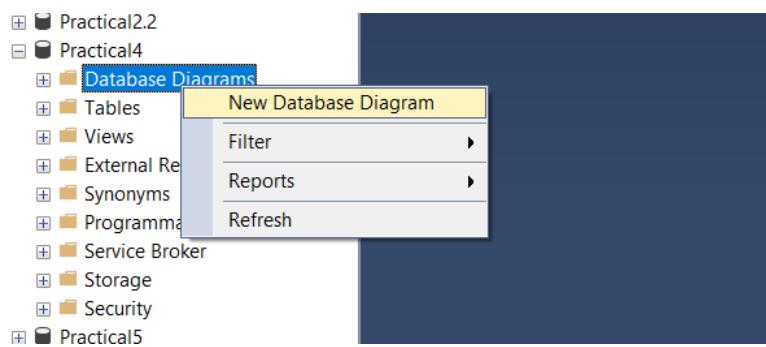
```

        astatus VARCHAR(10)
    )
CREATE TABLE DIM_trandetails
(
tnumber VARCHAR(6),
acnumber VARCHAR(6),
dot DATE,
medium_of_transaction VARCHAR(20),
transaction_type VARCHAR(20),
transaction_amount VARCHAR(7)
)
CREATE TABLE DIM_loan
(
loan_id VARCHAR(10),
custid VARCHAR(6),
bid VARCHAR(6),
loan_amount VARCHAR(7)
)
CREATE TABLE FACT_BANK
(
custid VARCHAR(6),
bid VARCHAR(6),
acnumber VARCHAR(6),
tnumber VARCHAR(6),
loan_id VARCHAR(10)
)

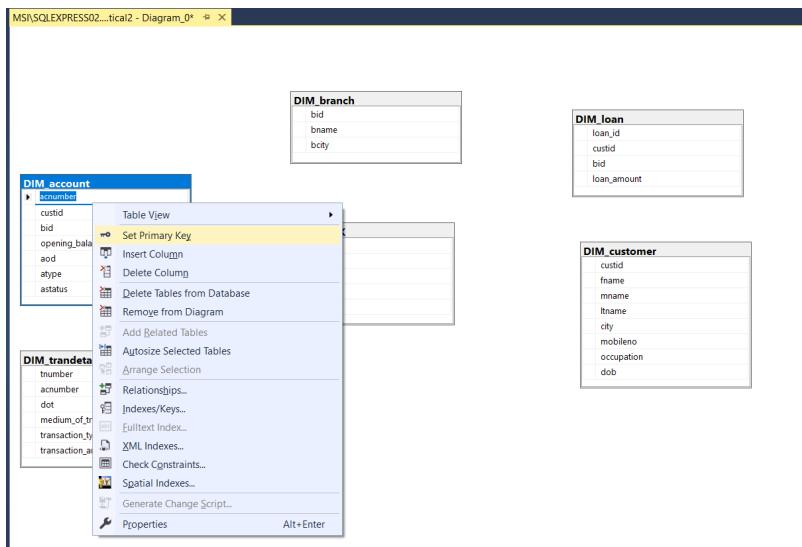
```



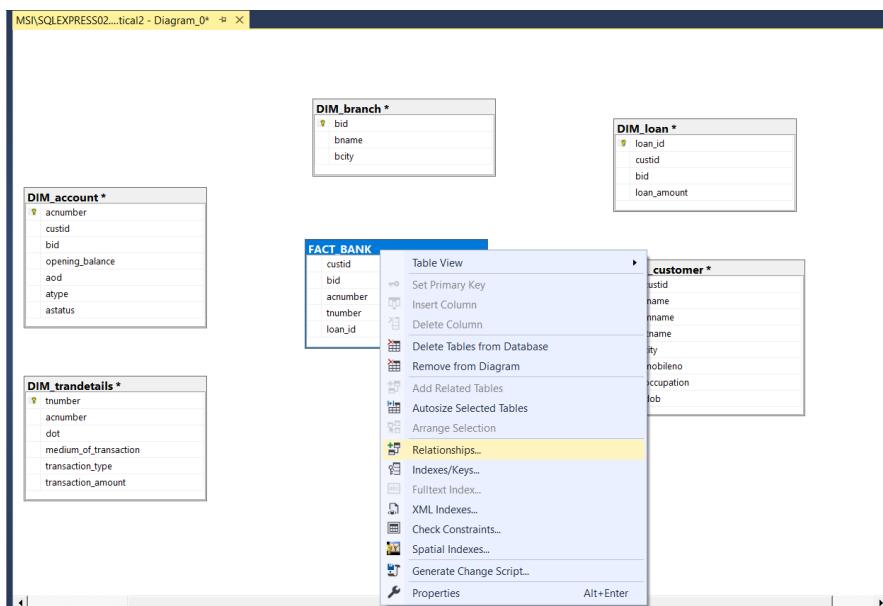
## 2) Create database diagram.

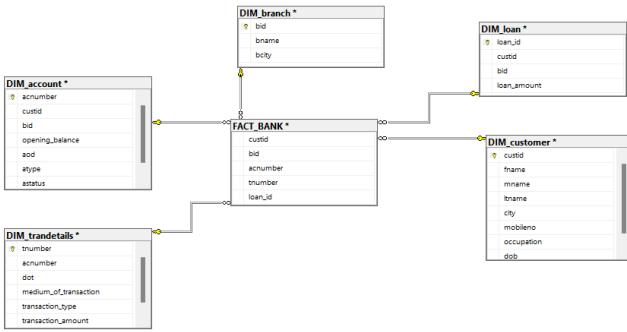


### 3) Set primary key for each table:

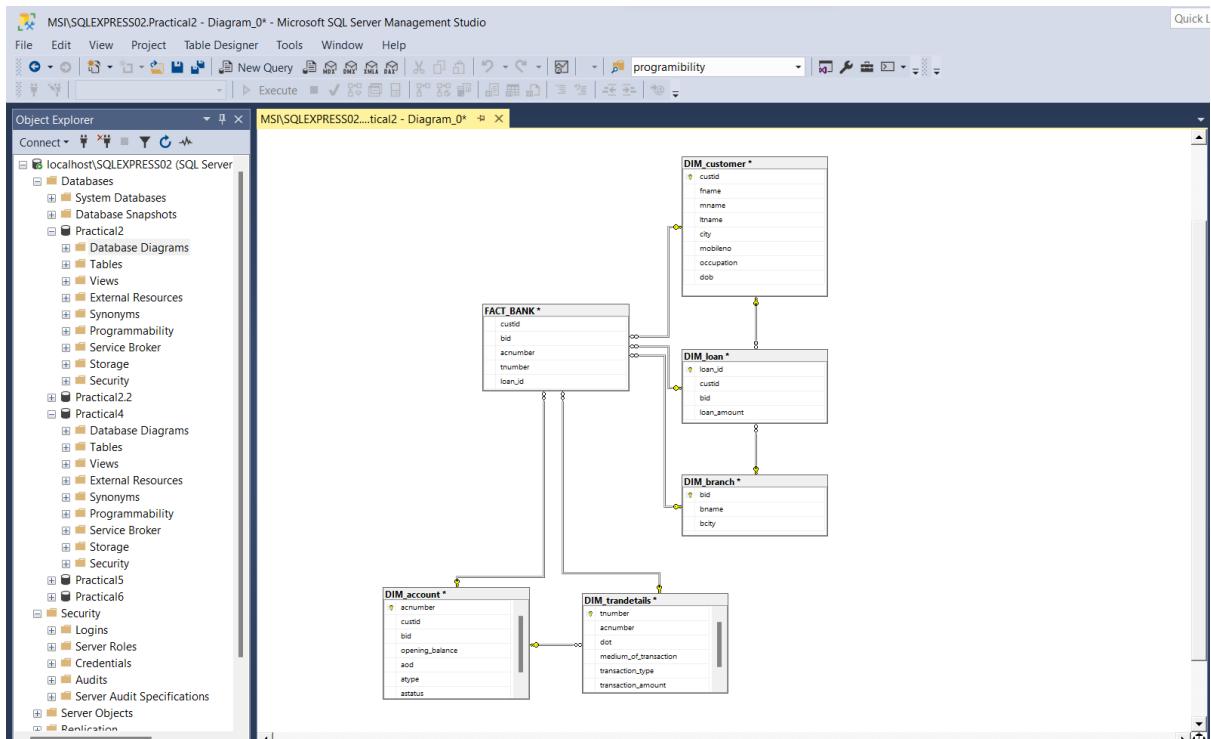


### 4) Set Relationships for fact table.





## 5) Set relationships of dimensional table with each other.



**Conclusion:** Multidimensional data constructed successfully.

## **Practical 8**

**Aim: Develop an application to implement defining subject areas, design of fact and dimension tables, data marts**

### **Theory:**

#### **Fact Table**

A Fact Table is a central table in a star schema of a data warehouse. It is an important concept required for Data Warehousing and BI . A fact table stores quantitative information for analysis and is often denormalized. A fact table works with dimension tables and it holds the data to be analyzed and a dimension table stores data about the ways in which the data can be analyzed.

Thus, a fact table consists of two types of columns. The foreign keys column allows to join with dimension tables and the measure columns contain the data that is being analyzed.

A dimension table is a table in a star schema of a data warehouse. A dimension table stores attributes, or dimensions, that describe the objects in a fact table.

#### **Dimension Table**

In data warehousing, a dimension is a collection of reference information about a measurable event. These events are known as facts and are stored in a fact table. Dimensions categorize and describe data warehouse facts and measures in ways that support meaningful answers to business questions. They form the very core of dimensional modeling.

#### **Data Mart**

A data mart is a simple form of data warehouse focused on a single subject or line of business. With a data mart, teams can access data and gain insights faster, because they don't have to spend time searching within a more complex data warehouse or manually aggregating data from different sources.

#Steps:

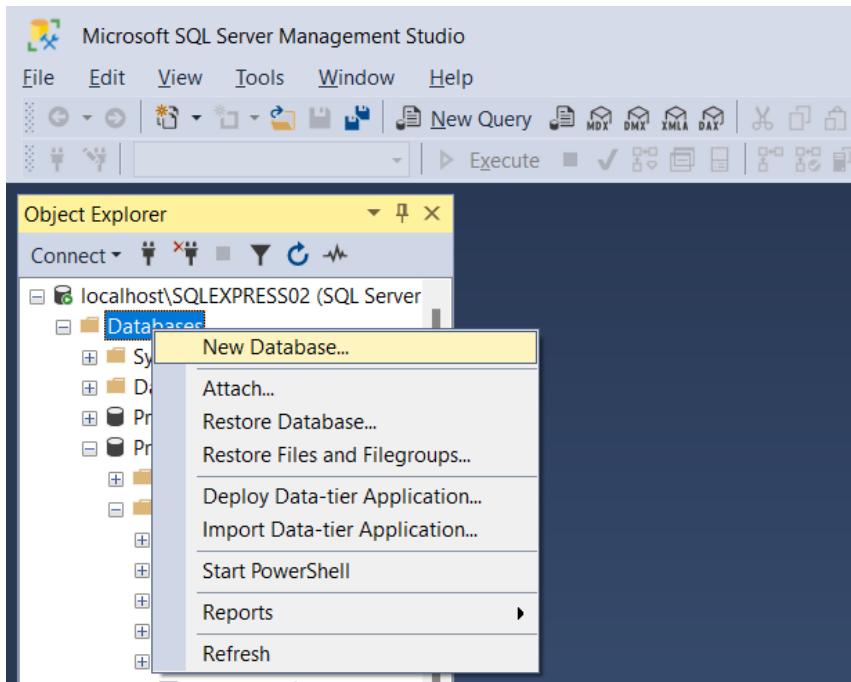
- 1) Create fact and dimensions tables in excel(.xls) (Change sheet name for every sheet)

	A	B	C	D	E	F
1	T_id	C_id	Stud_id	R_id		
2	1	1	1	1	1	
3	2	2	2	2	2	
4	3	3	3	3	3	
5	4	4	4	4	4	
6						
7						

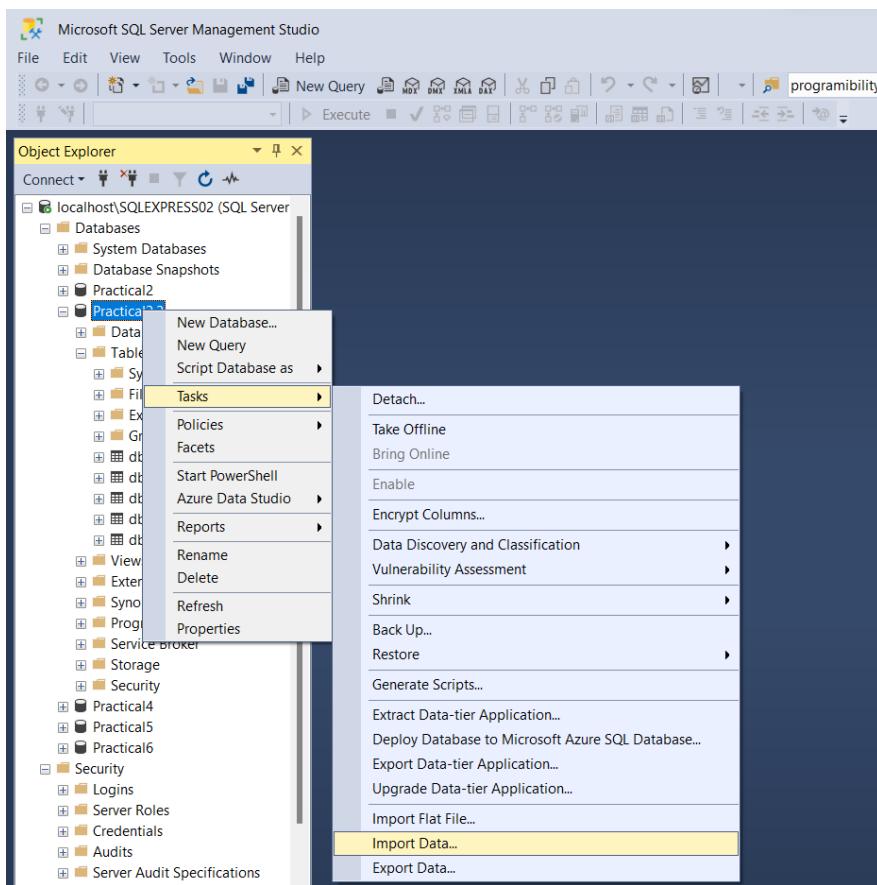
> This PC > Data (D:) > Practical-Msc-Sem-2 > BI > Excel-Tables

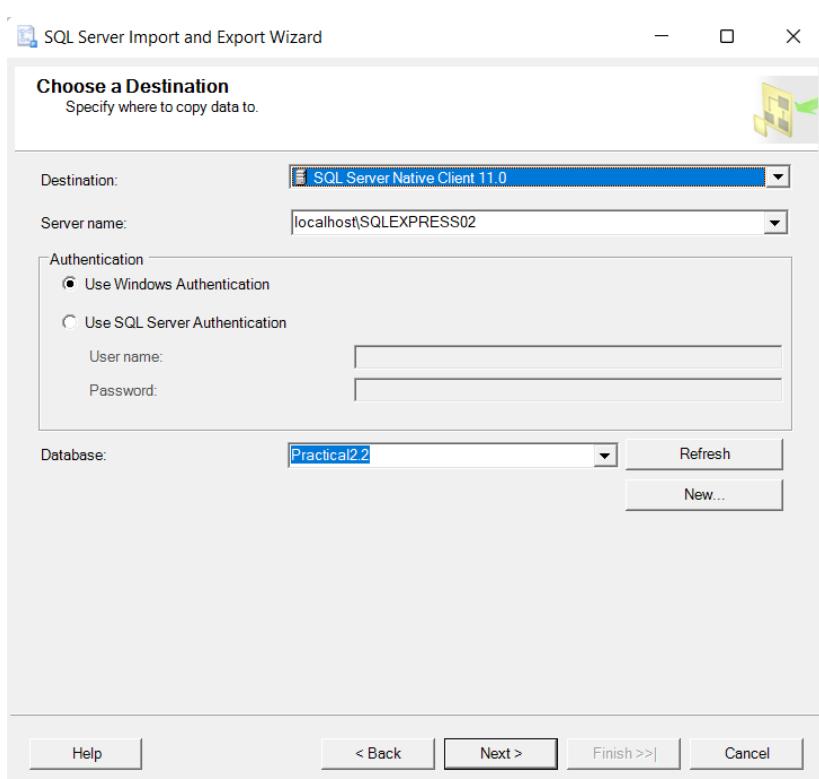
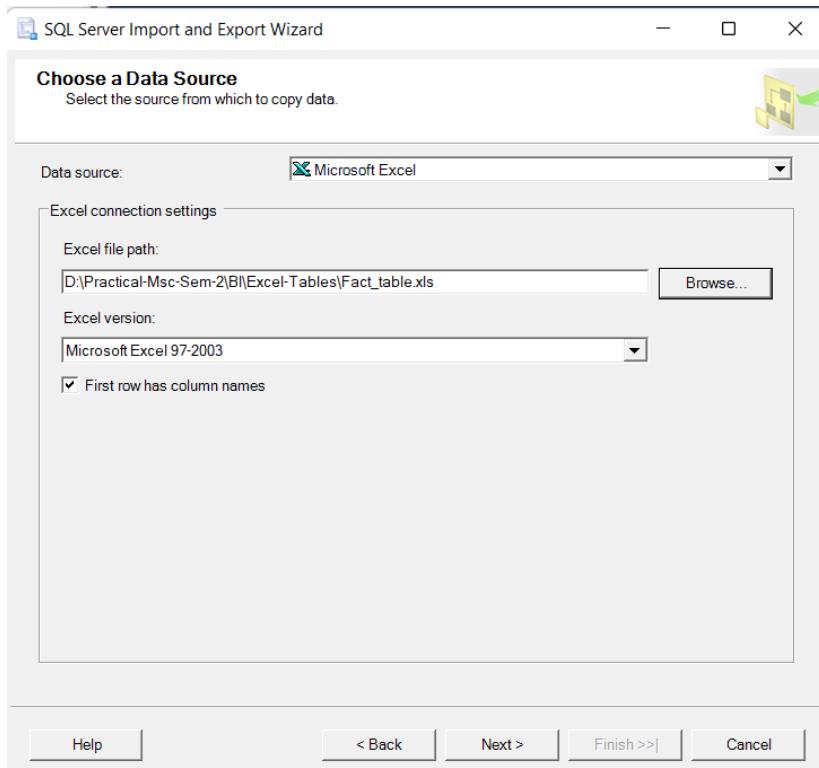
Name	Date modified	Type	Size
dimension_1	04-04-2022 08:09	Microsoft Excel 97...	25 KB
dimension_2	04-04-2022 08:20	Microsoft Excel 97...	25 KB
dimension_3	04-04-2022 08:20	Microsoft Excel 97...	25 KB
dimension_4	04-04-2022 08:20	Microsoft Excel 97...	25 KB
Fact_table	04-04-2022 19:58	Microsoft Excel 97...	25 KB

- 2) Create new database.

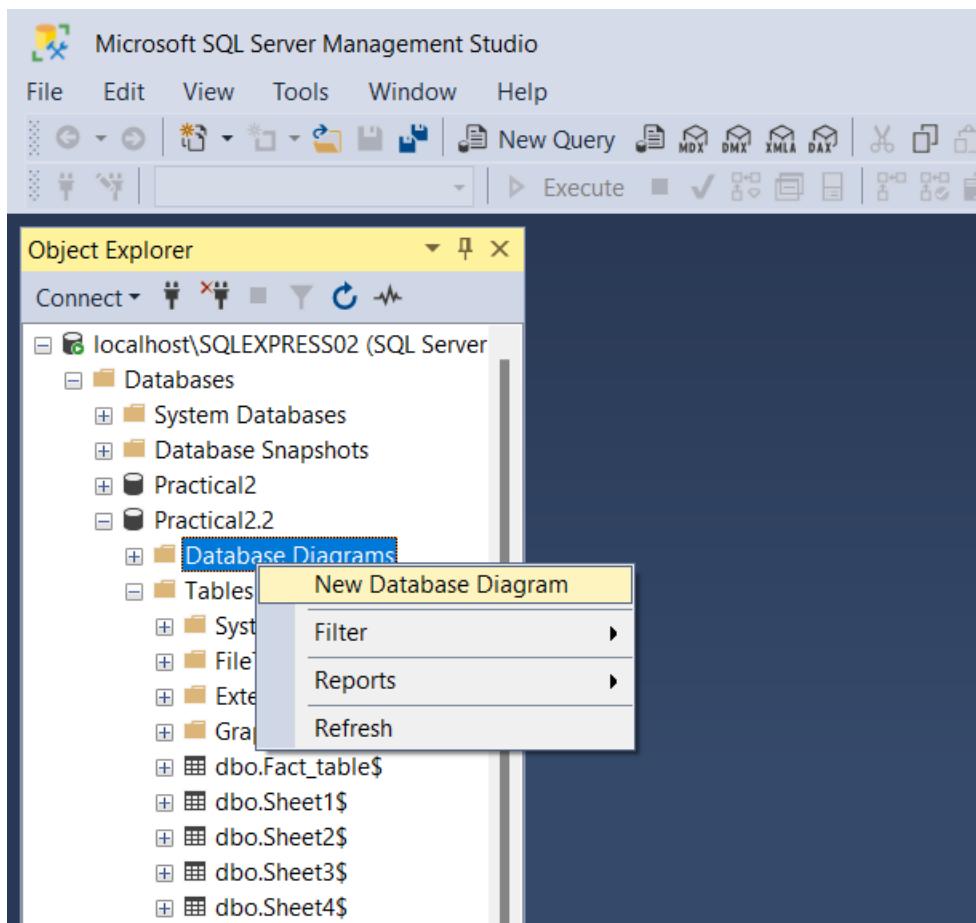


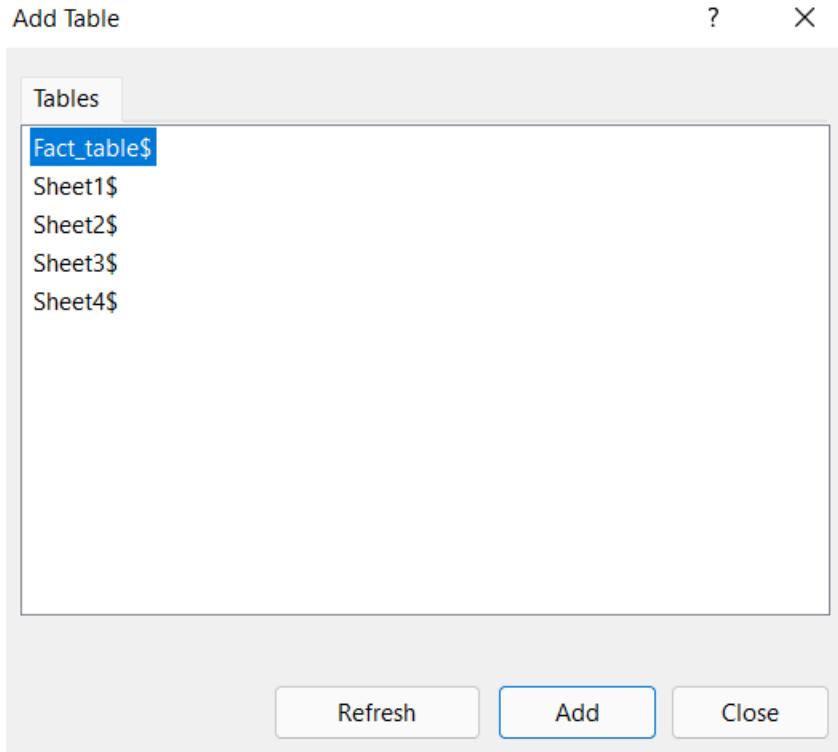
3) Import excel files > Choose data source > Choose destination source >Finish



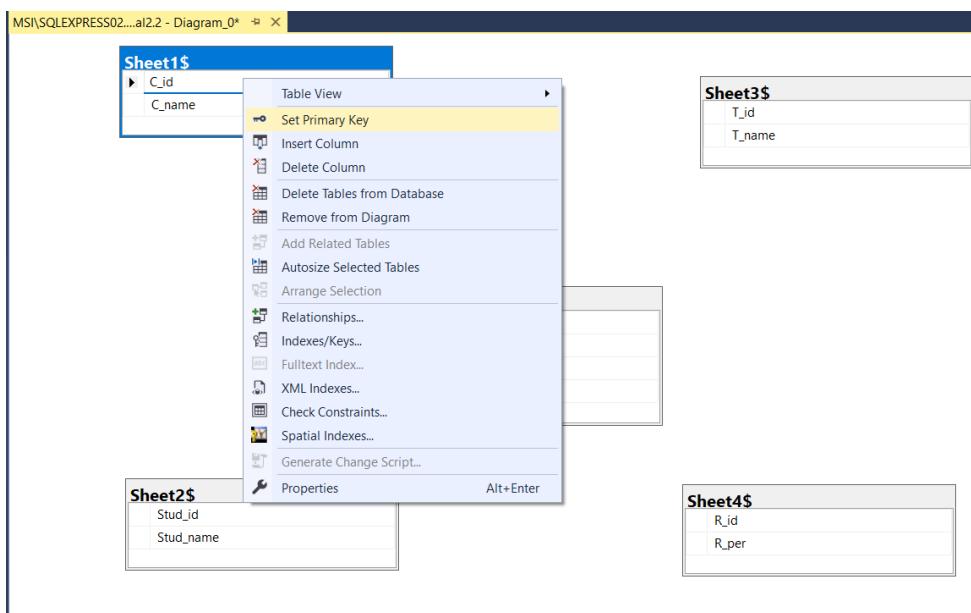


#### 4) Create database diagrams

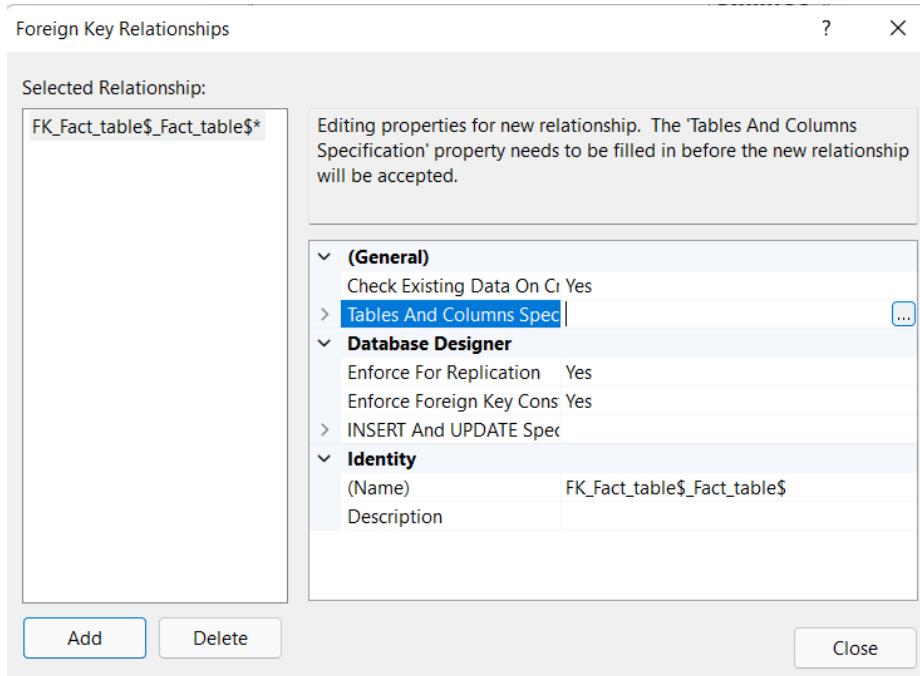
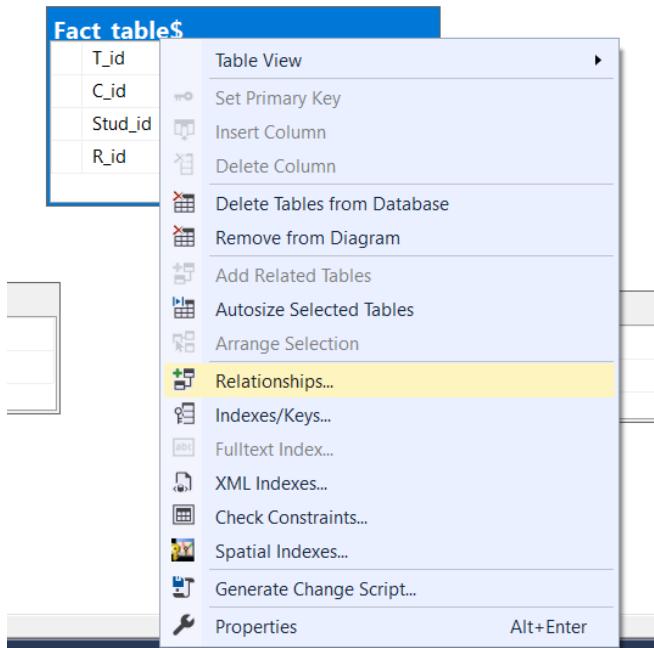


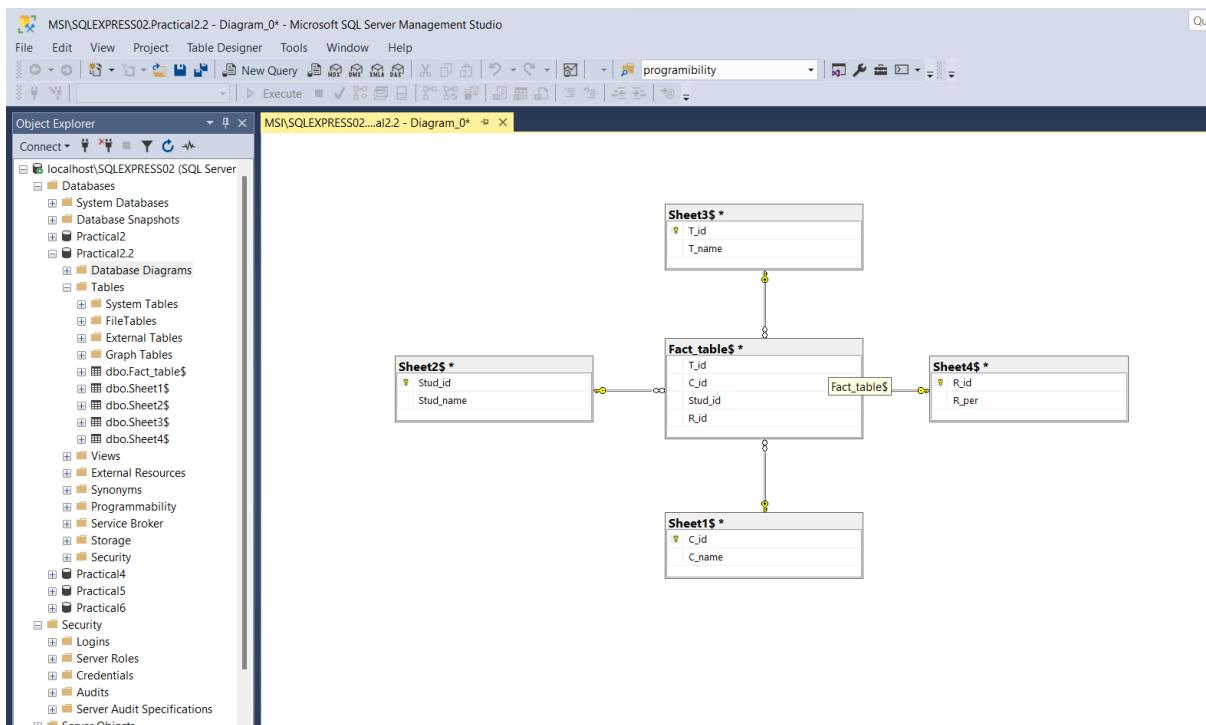
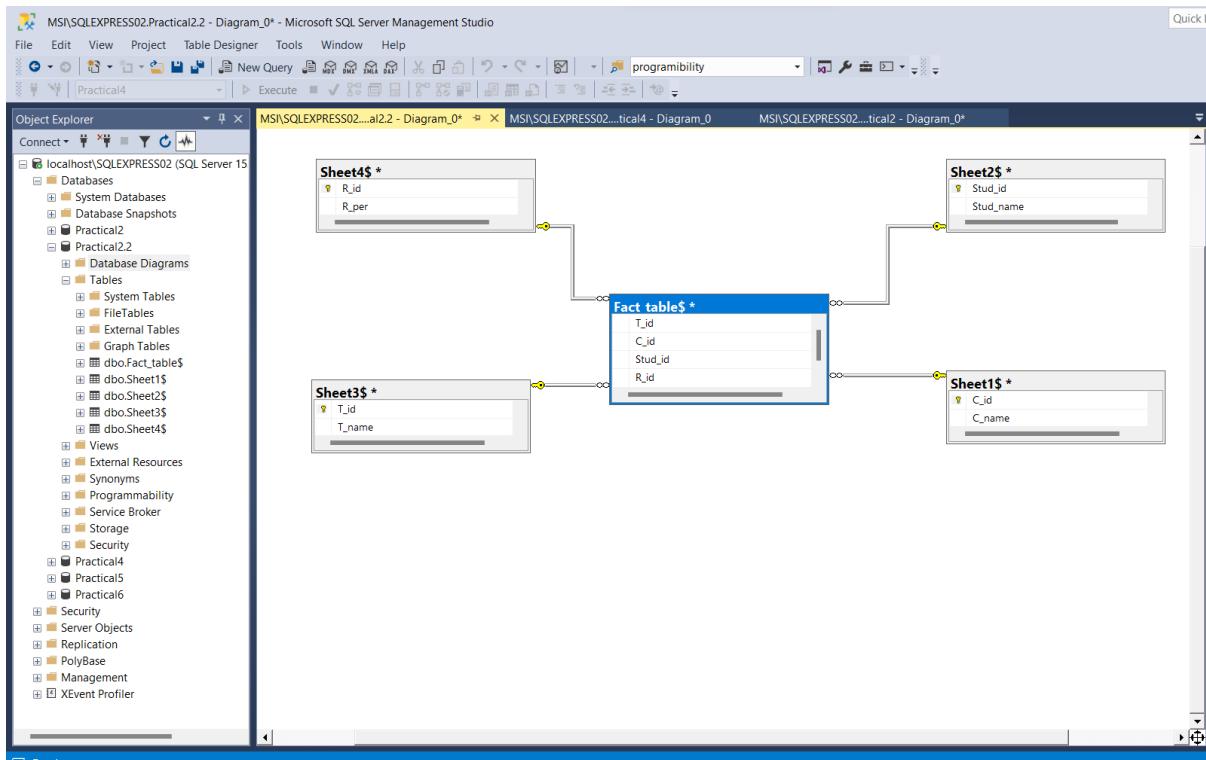


### 5) Set primary keys



### 6) Create relationships for fact table





Conclusion: Successfully Implemented fact, dimension table and data marts in MS SQL.

