



Fall 2023

CIS 611 - Enterprise Database and Warehouse

Project Report

Topic: Business Intelligence and Visualization Project Using Multidimensional OLAP Cube and MS Data Tool

Guided By:

Prof. Sunnie S. Chung

Submitted By:

Kajal Shrivastava

CSU ID: 2861721

Mohammed Ovesh, Shaikh

CSU ID: 2860166

Index

No.	Content	Page no.
1.	Platform/System Set up Procedures/Instructions	3
2.	Dataset Details	3
3.	Designing and development of CUBE <ul style="list-style-type: none">• Give the Project name and create the project.• Creating New Data Source• Creating Data Source View• Creating the Cube• Building the Cube• Cube Deployment	4 9 9 10 11 13 13
4.	MDX Query for analysis	14
5.	Analysis and Visualization over the CUBE <ul style="list-style-type: none">• Queries	17 17
6.	Data Mining Performed	22
7.	References	24

1 . Platform/System Set up Procedures/Instructions: -

For this project, we have used Microsoft Business Intelligence Data Mining Platform with:

1. Visual Studio 2019 with Analysis Services extension

Downloading link:

<https://marketplace.visualstudio.com/items?itemName=ProBITools.MicrosoftAnalysisServicesModelingProjects>

2. Microsoft SQL Server for Analysis Service (Multi-Dimensional OLAP Server) for Business Intelligence (SSDT BI) and Data Integration Service: SQL Server Data Tools (SSDT) for Visual Studio 2019

Downloading link:

<https://www.microsoft.com/en-us/sql-server/sql-server-downloads>

3. SQL Server Management Studio (SSMS)

Downloading link:

<https://learn.microsoft.com/en-us/sql/ssms/download-sql-server-management-studio-ssms?view=sql-server-ver16>

2 . Dataset Details: -

For this project, we have taken dataset **ContosoRetailDW.bak** from ContosoBIdemoBAK.exe

Downloading link:

<https://www.microsoft.com/en-us/download/details.aspx?id=18279>

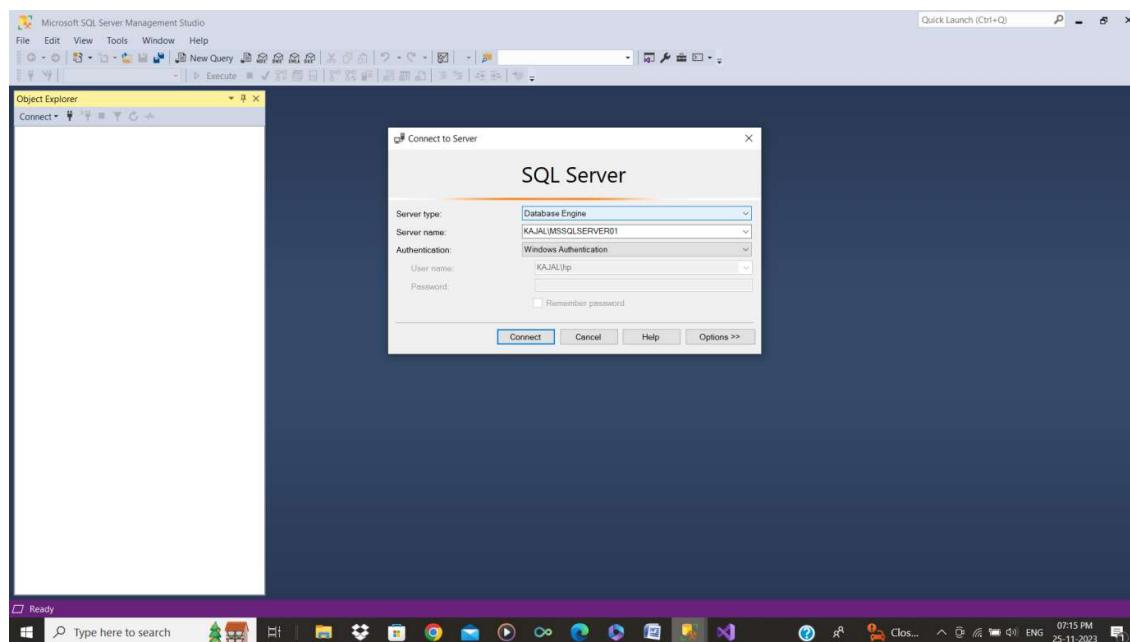
- Contoso Database is a sample dataset that Microsoft provides as a data analysis study. It simulates a company's sales data.
- The Contoso BI Demo dataset is used to demonstrate DW/BI functionalities across the entire Microsoft Office product family. This dataset includes C-level, sales/marketing, IT, and common finance scenarios for the retail industry and support map integration.
- The dataset consists of 17 Dimension Tables and 8 Fact Tables.
- It consists of tables such as DimAccount, DimCustomer, DimGeography, DimProduct, DimDate, DimPromotion, FactSales, FactInventory, FactOnlineSales, etc.
- Sample 10 rows of DimCustomer is shown below:

	CustomerKey	GeographyKey	CustomerLabel	Title	FirstName	MiddleName	LastName	NameStyle	BirthDate	MaritalStatus	Suffix	Gender	EmailAddress	YearlyIncome	TotalChildren	NumberChildrenAtHome	Education
1	1	680	11000	NULL	Jon	V	Yang	0	1966-04-08	M	NULL	M	jon24@adventure-works.com	90000.00	2	0	Bachelor
2	2	692	11001	NULL	Eugene	L	Huang	0	1965-05-14	S	NULL	M	eugene10@adventure-works.com	60000.00	3	3	Bachelor
3	3	493	11002	NULL	Ruben	NULL	Torres	0	1965-08-12	M	NULL	M	ruben35@adventure-works.com	60000.00	3	3	Bachelor
4	4	519	11003	NULL	Christy	NULL	Zhu	0	1968-02-15	S	NULL	F	christy12@adventure-works.com	70000.00	0	0	Bachelor
5	5	706	11004	NULL	Elizabeth	NULL	Johnson	0	1968-08-08	S	NULL	F	elizabeth5@adventure-works.com	80000.00	5	5	Bachelor
6	6	478	11005	NULL	Julio	NULL	Ruiz	0	1965-08-05	S	NULL	M	julio18@adventure-works.com	70000.00	0	0	Bachelor
7	7	509	11006	NULL	Janet	G	Alvarez	0	1965-12-06	S	NULL	F	jane98@adventure-works.com	70000.00	0	0	Bachelor
8	8	568	11007	NULL	Marco	NULL	Mehta	0	1964-05-09	M	NULL	M	marco14@adventure-works.com	60000.00	3	3	Bachelor
9	9	425	11008	NULL	Rob	NULL	Verhoff	0	1964-07-07	S	NULL	F	rob1@adventure-works.com	60000.00	4	4	Bachelor
10	10	492	11009	NULL	Shannon	C	Carlson	0	1964-04-01	S	NULL	M	shannon38@adventure-works.com	70000.00	0	0	Bachelor

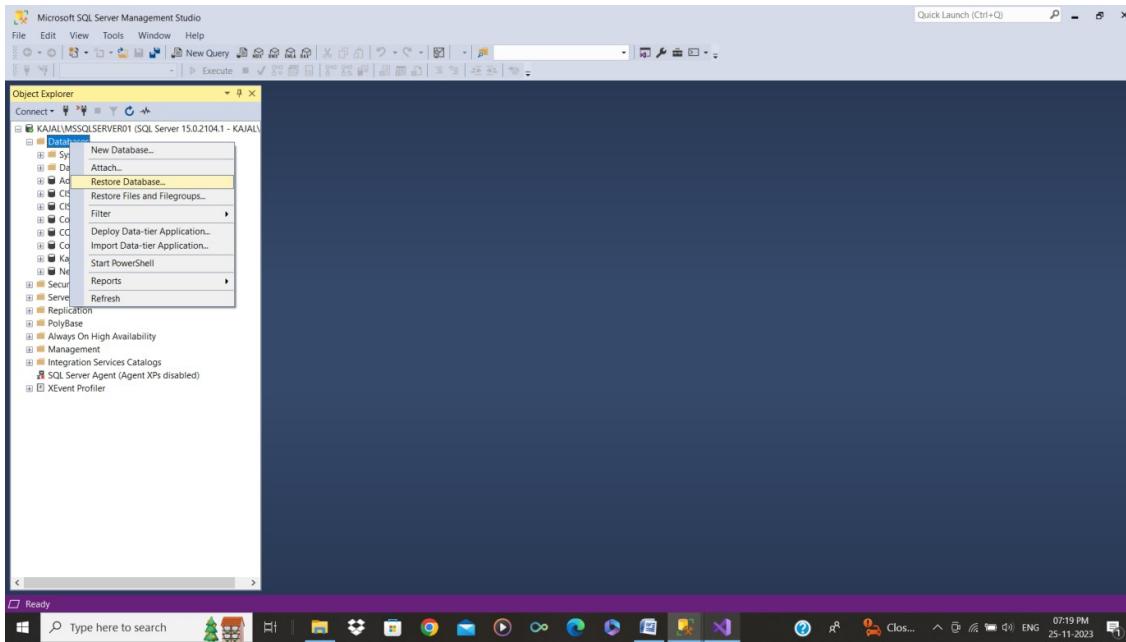
3. Designing and development of CUBE: -

Cube Design:

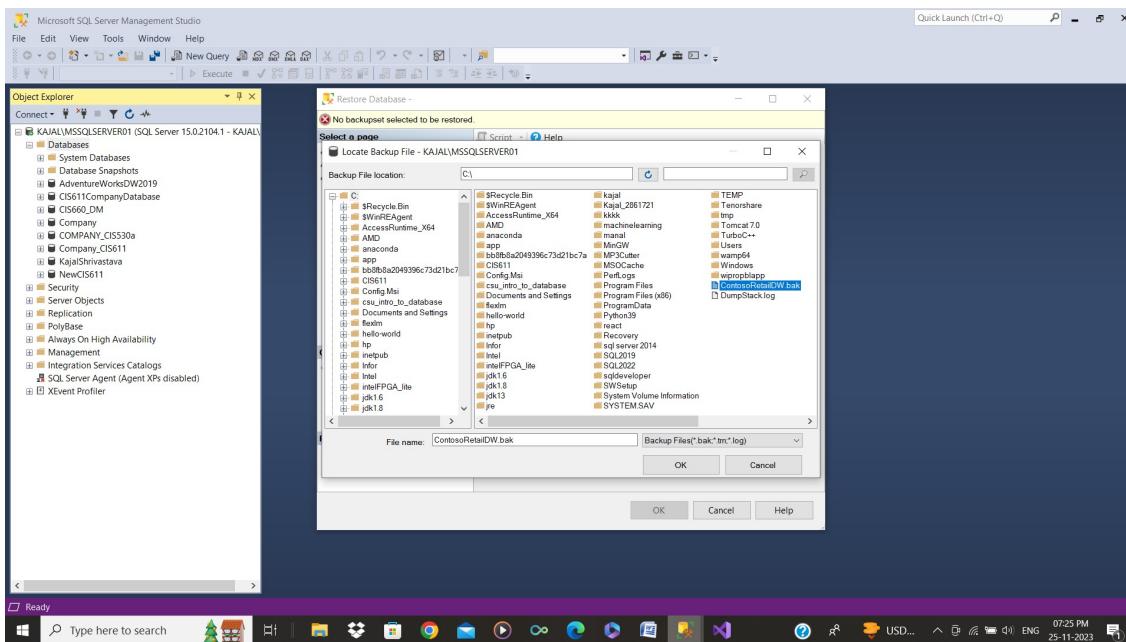
- Open SSMS and select the Database Engine from where you want to import the database for analysis.

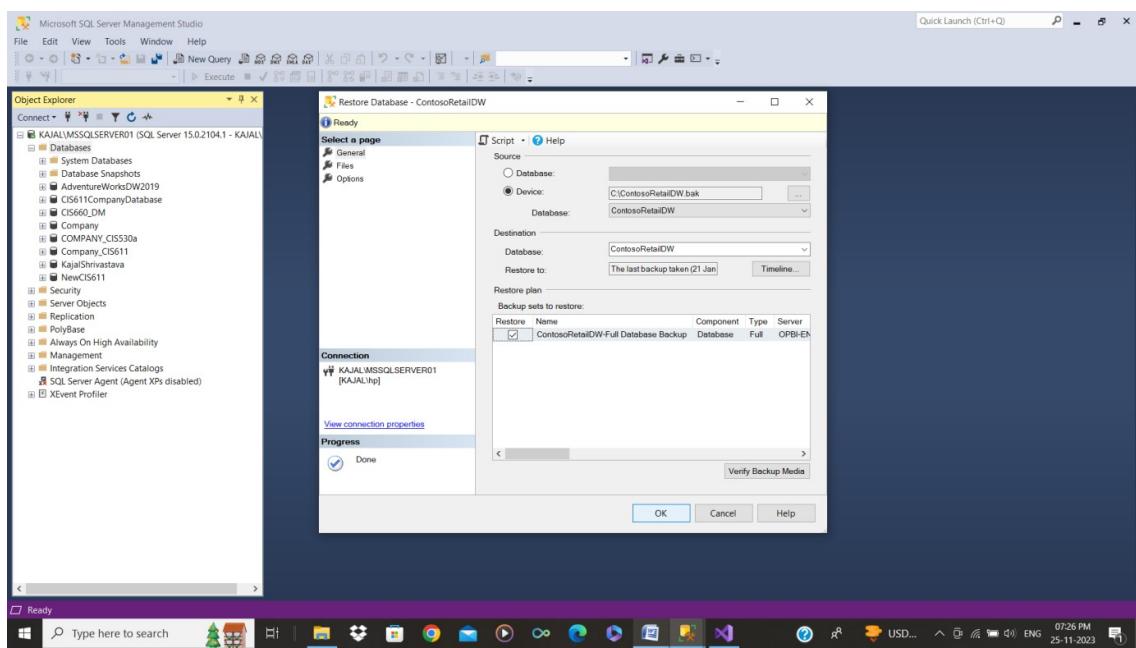
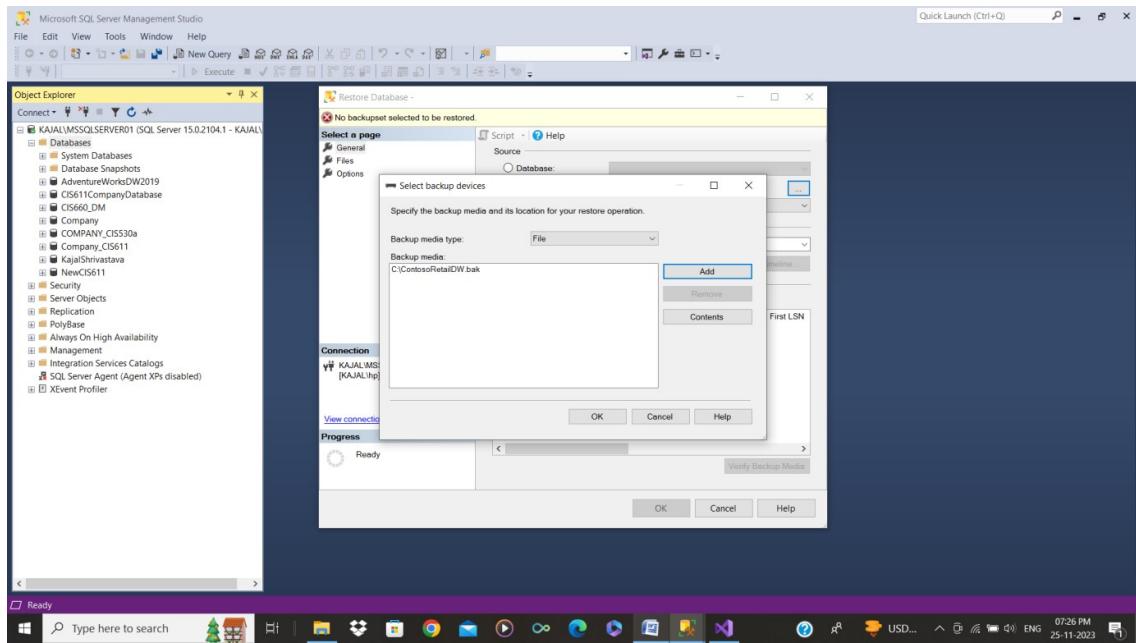


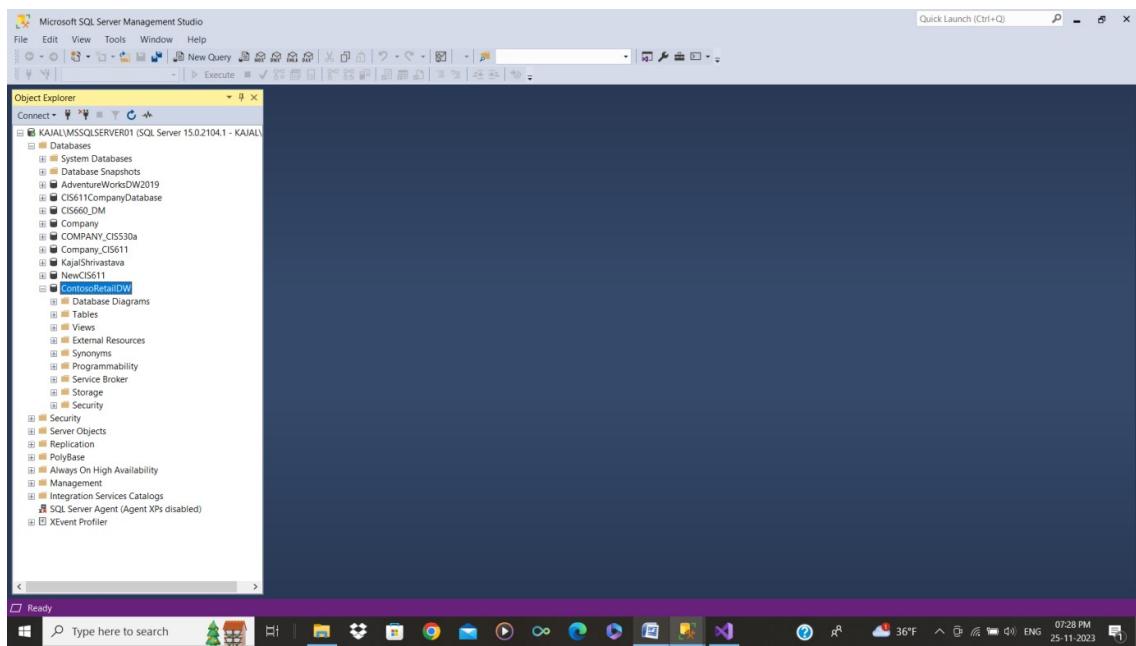
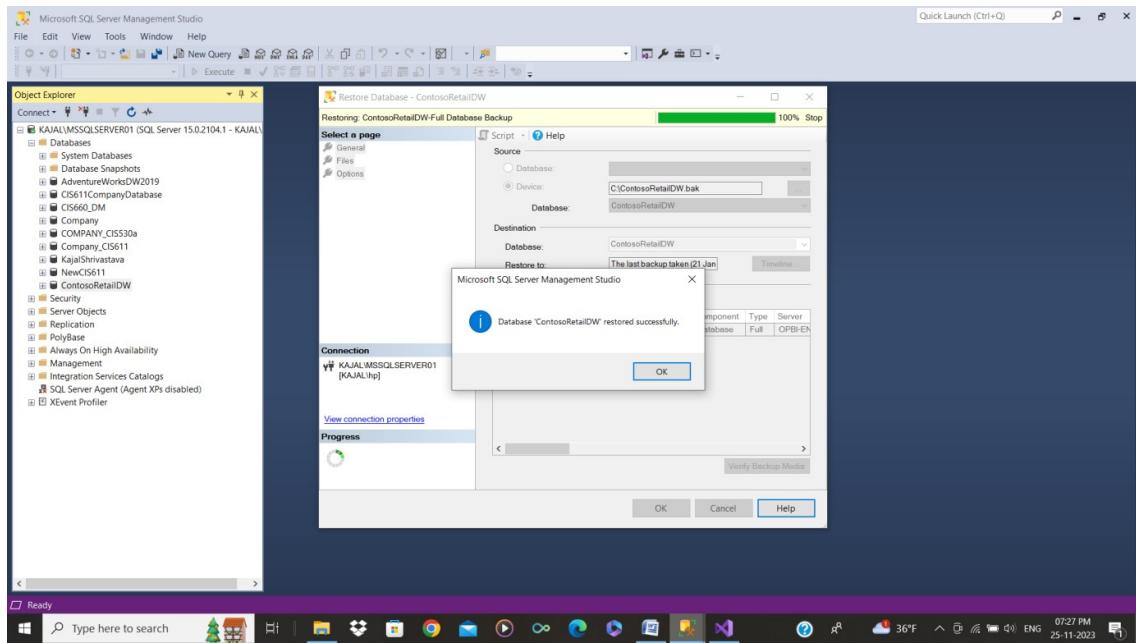
- Restore the Database file **ContosoRetailDW.bak**



- Select the Database file **ContosoRetailDW.bak**

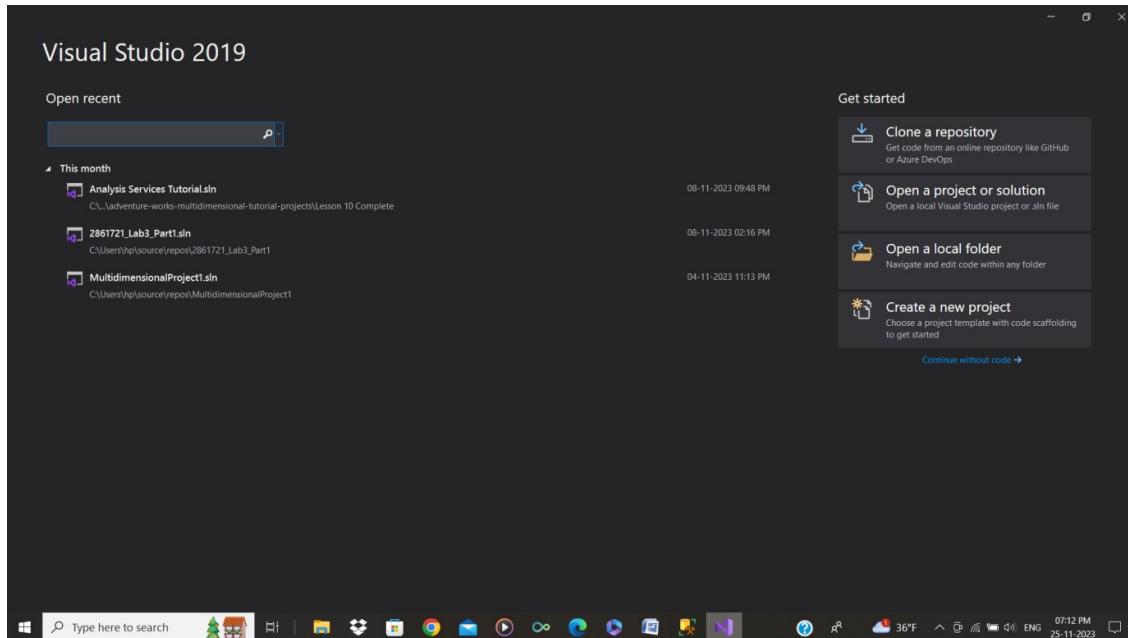




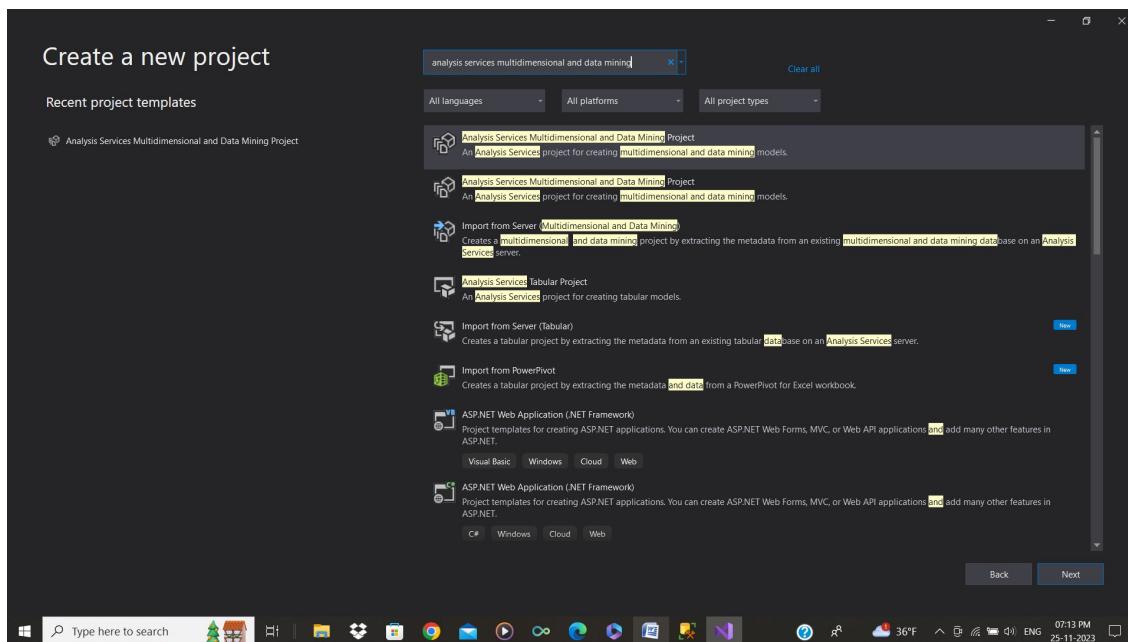


- Starting with the designing of the cube:

1. Open Visual Studio
2. Create a New Project

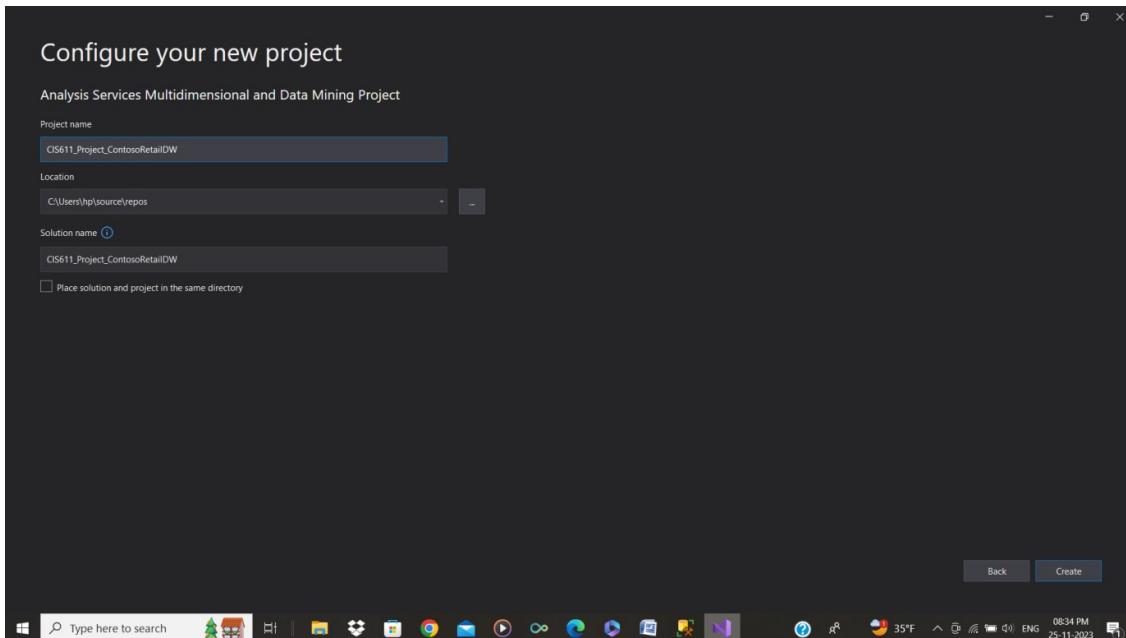


3. Search for Analysis Service Multidimensional and Data Mining Project. Select and Click Next.



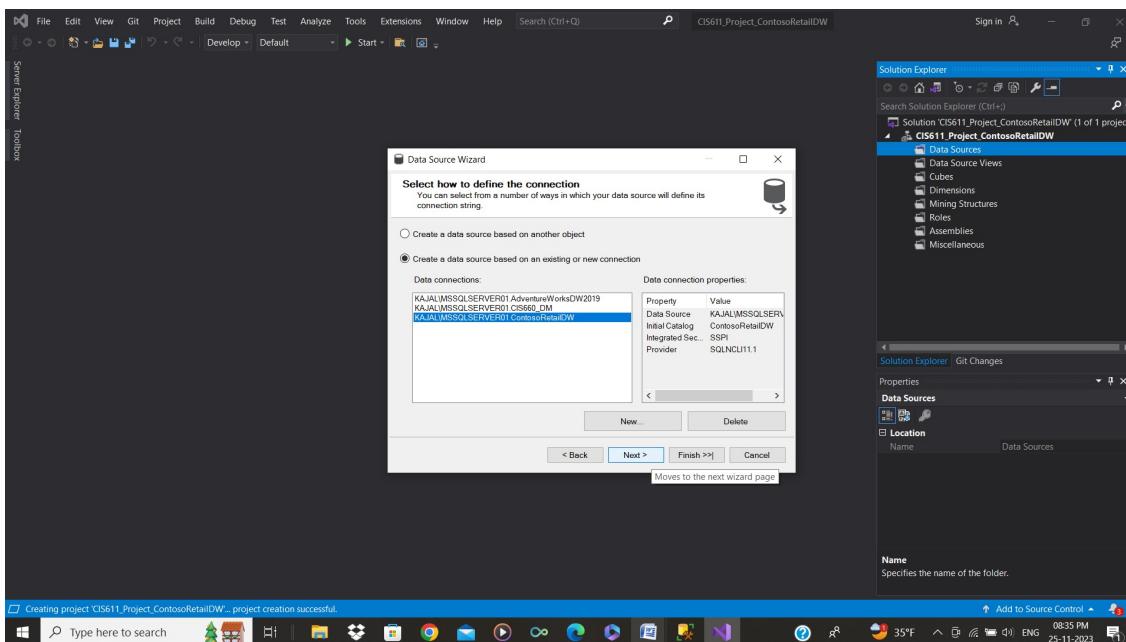
4. Give the Project name and create the project.

- Created a new project with name of ‘**CIS_611_ContosoRetailDW**’ in Analysis Services Multidimensional and Data Mining Project.

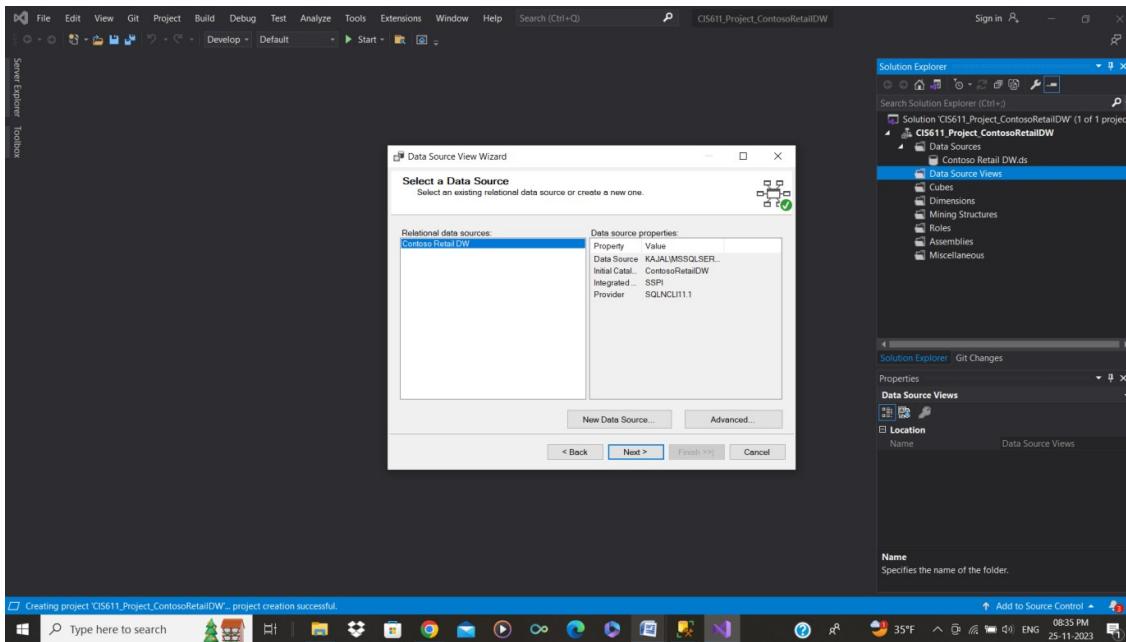


5. Creating New Data Source

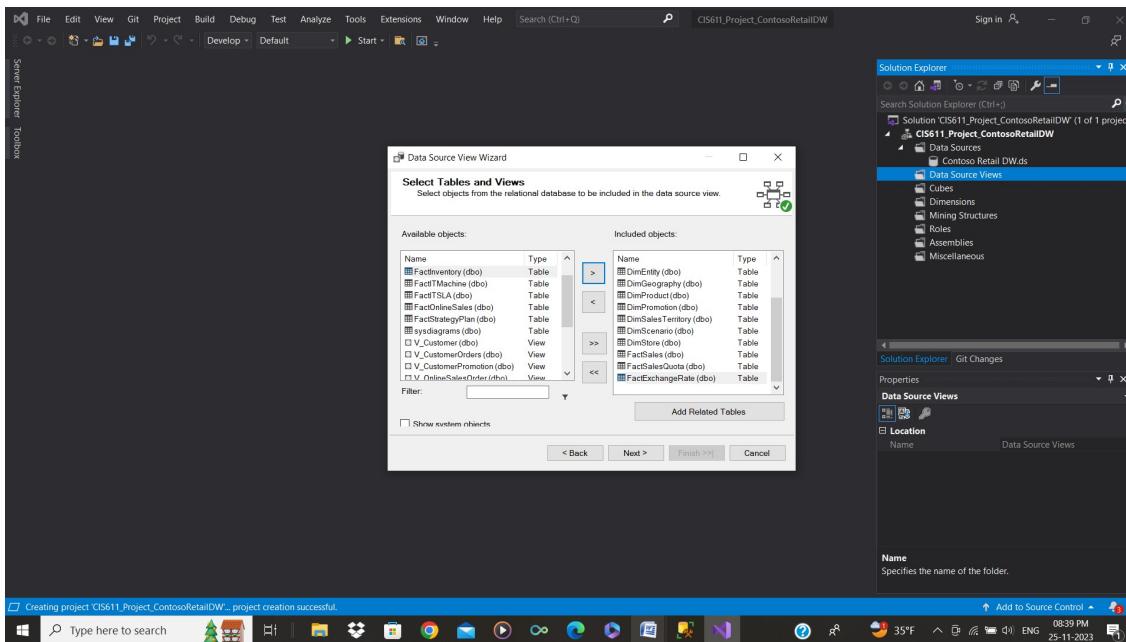
- This window will ask you to add new connection from existing connection.
- Especially select to Use the service account and then Click NEXT.



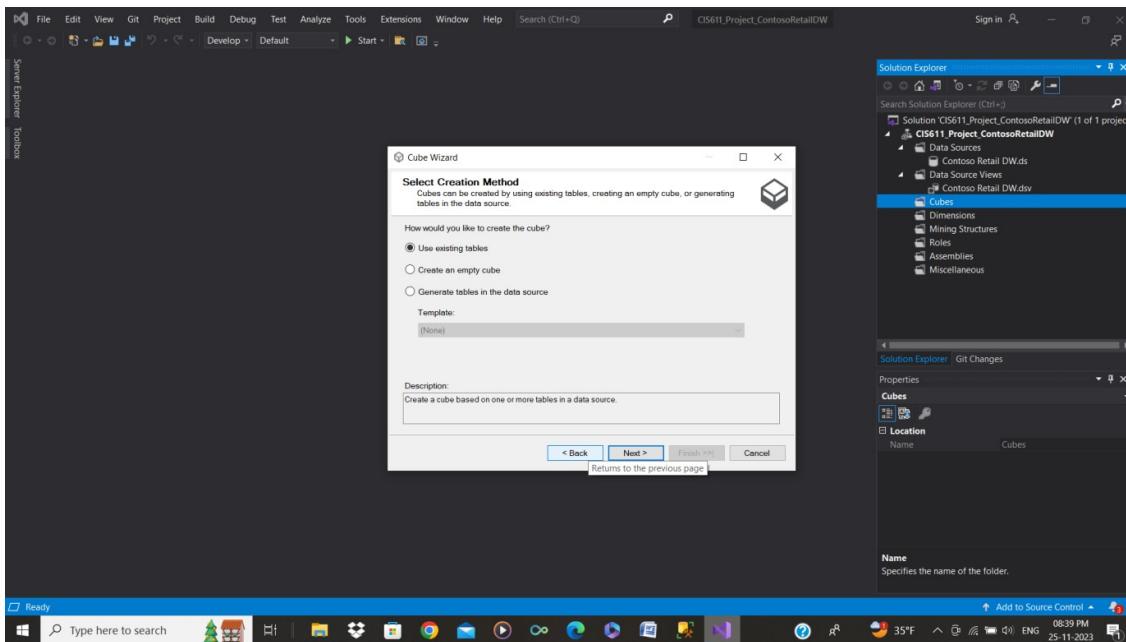
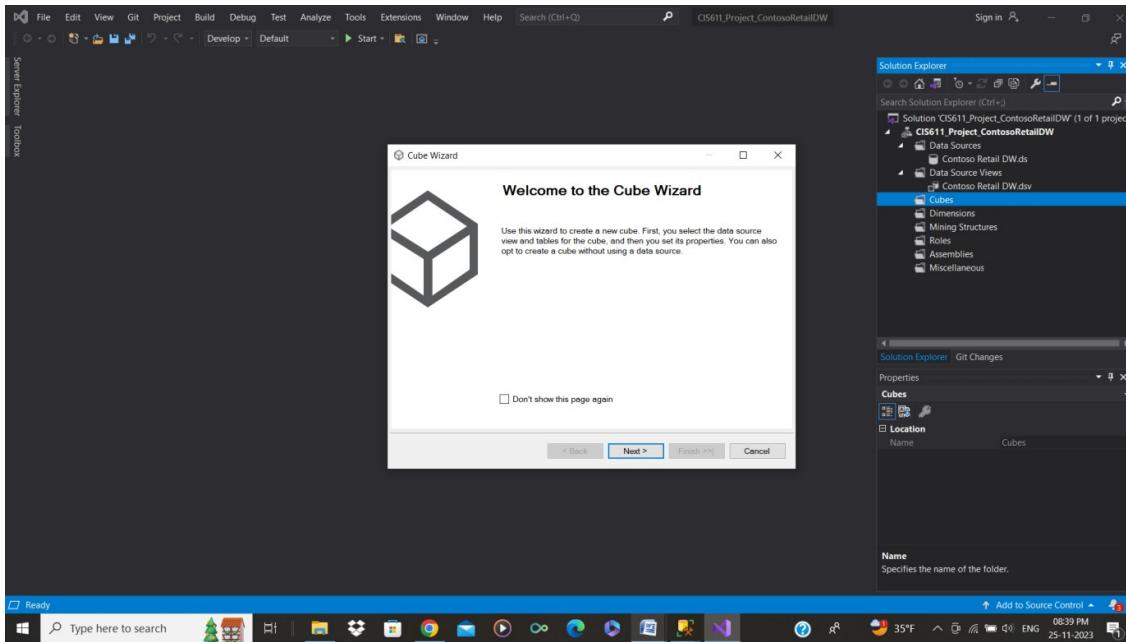
6. Creating Data Source View



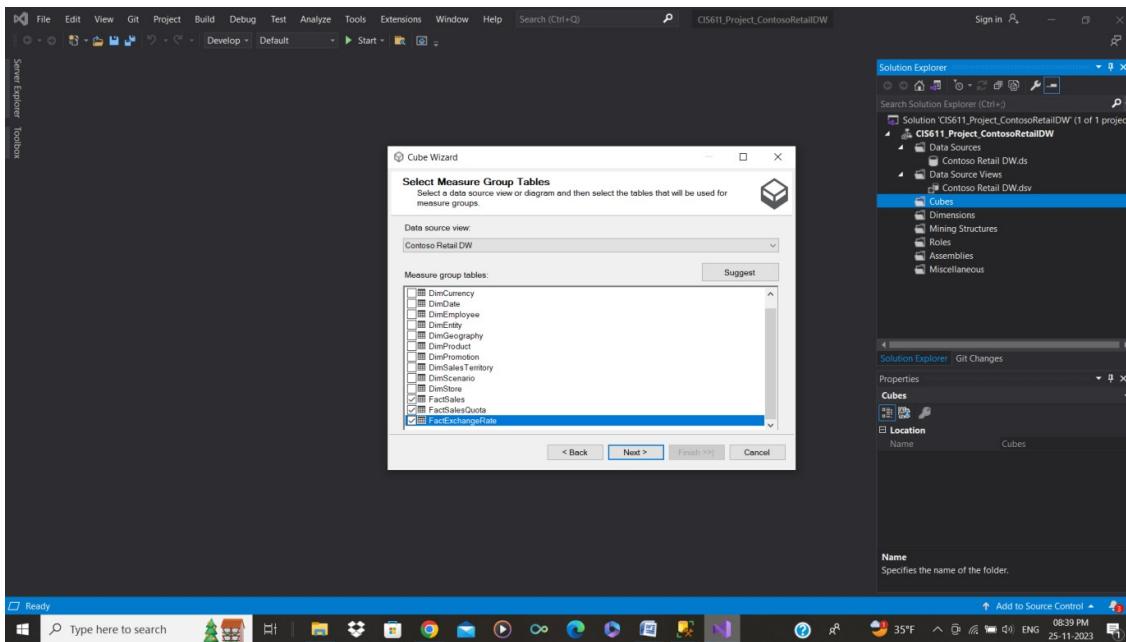
- Below, it shows that which Tables and views (Selecting Features) to be selected from the database in the left side and then after clicking next the selected tables and views will be pop-up in the right side of the area.
- After that a window will pop-up, that's shows a confirmation to complete the whole process. Click FINISH.



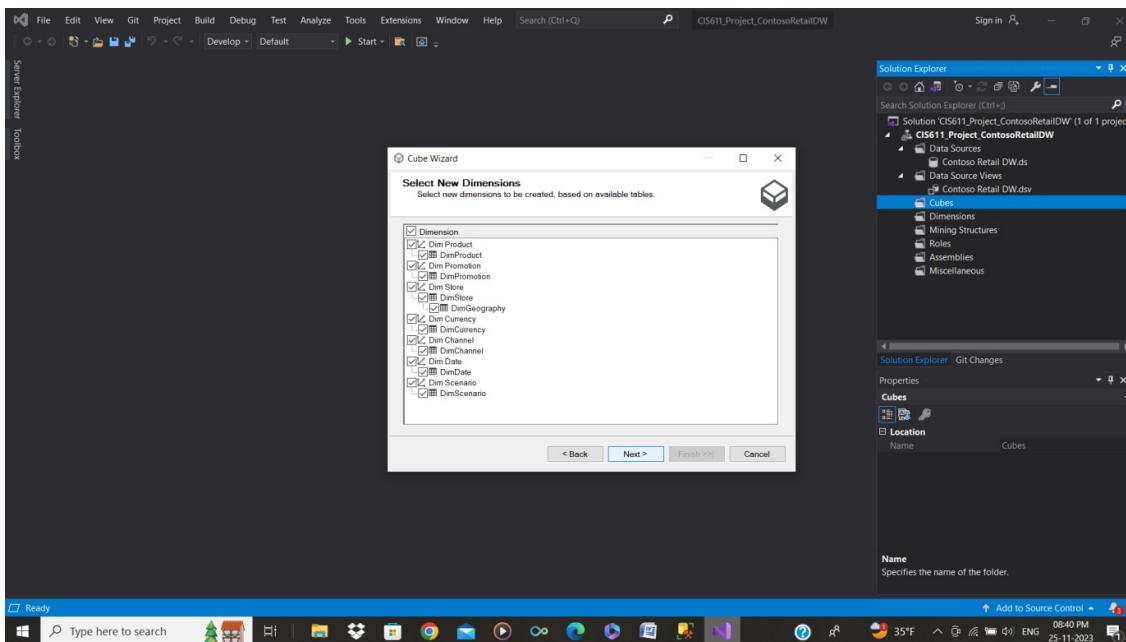
7. Creating the Cube



- After that this window will for Select Measure Group Tables, it means you have select which tables should be group in it. Like: **FACT SALES**, **FACT SALES QUOTA**, **FACT EXCHANGE RATE**.

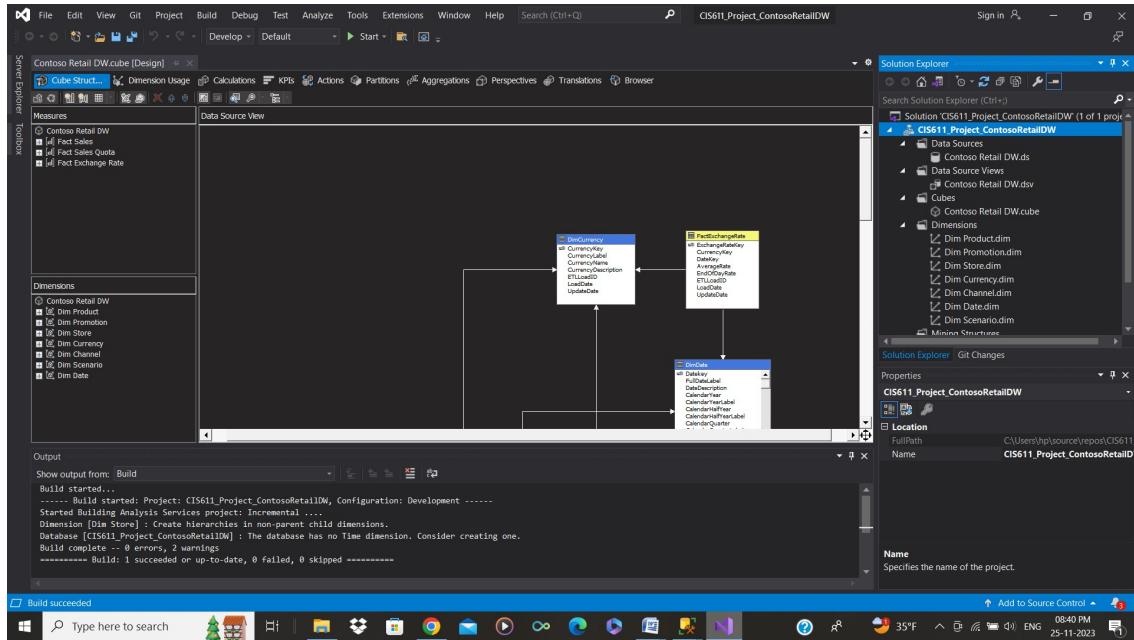


- After that window for selected New Dimensions will appear, the dimensions for the Cube will be created automatically. Click NEXT.



8. Building the Cube:

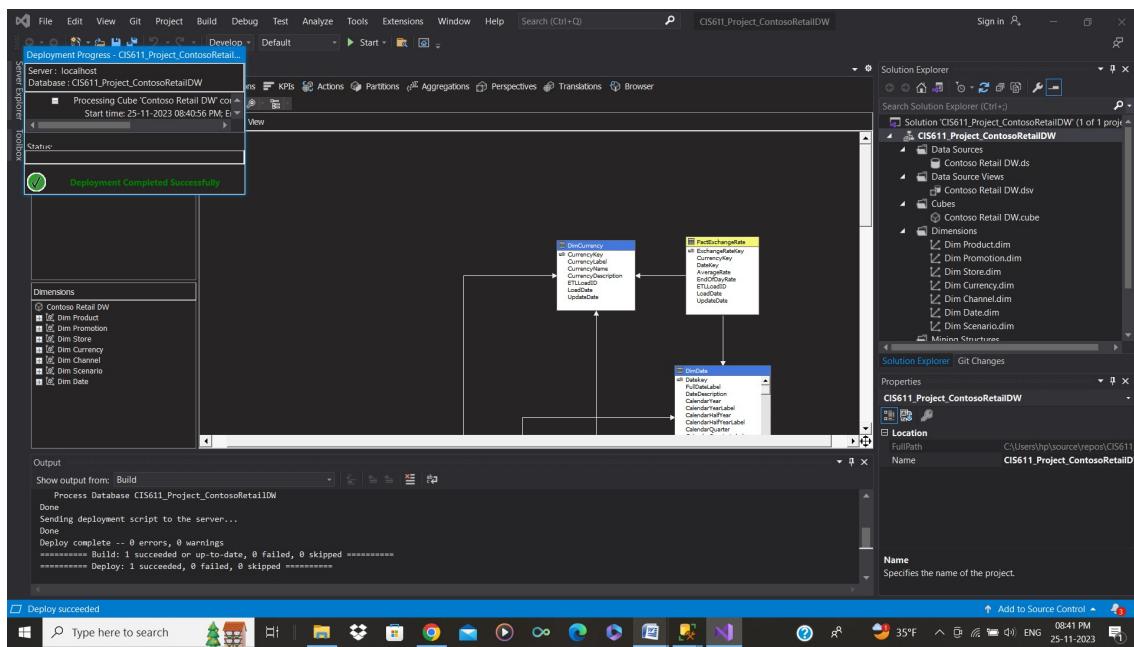
Build Successful



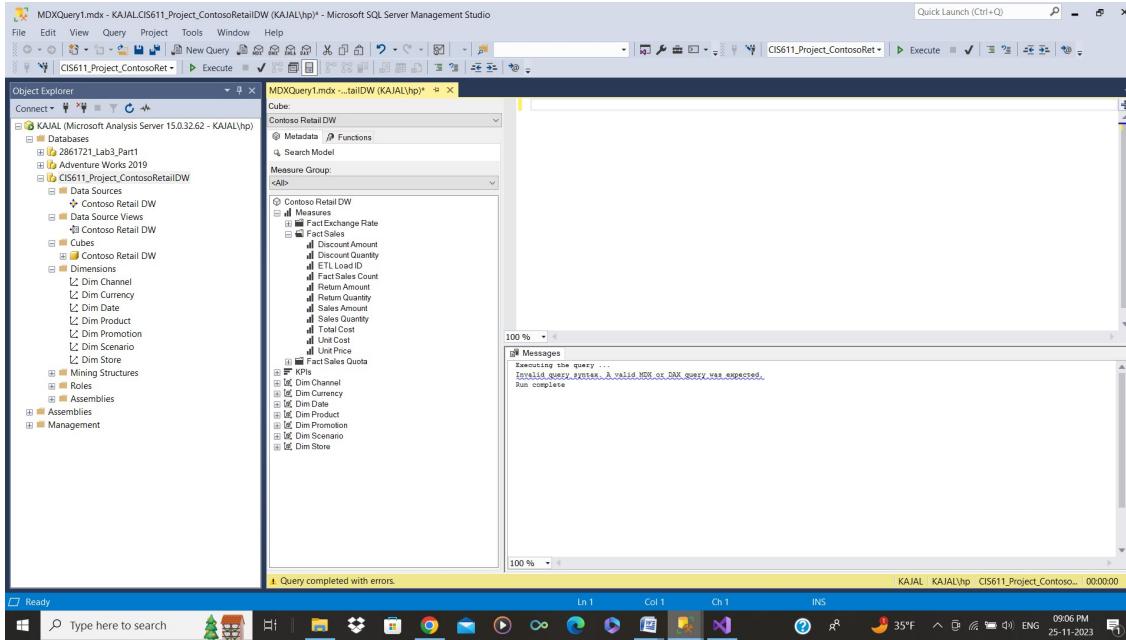
9. CUBE DEPLOYMENT:

- Deploying the cube on the Analysis Server

Deployment Successful



Checking the Analysis Server to Check if the CUBE is DEPLOYED SUCCESSFULLY



4 . MDX Query for analysis: -

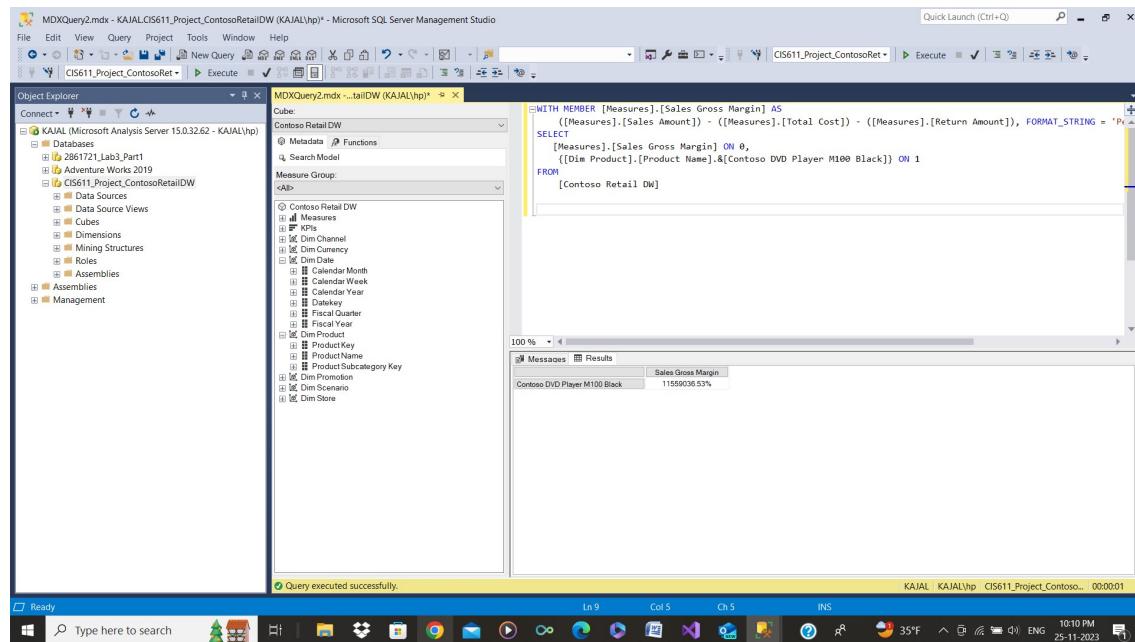
- Executing MDX (Multidimensional Expressions) over the cube.
- MDX is specifically designed for OLAP systems, allows users to formulate complex queries and retrieve data from the cube.

Here, we are analyzing the sales of Contoso DVD Player M100 Black.

• Code Script:

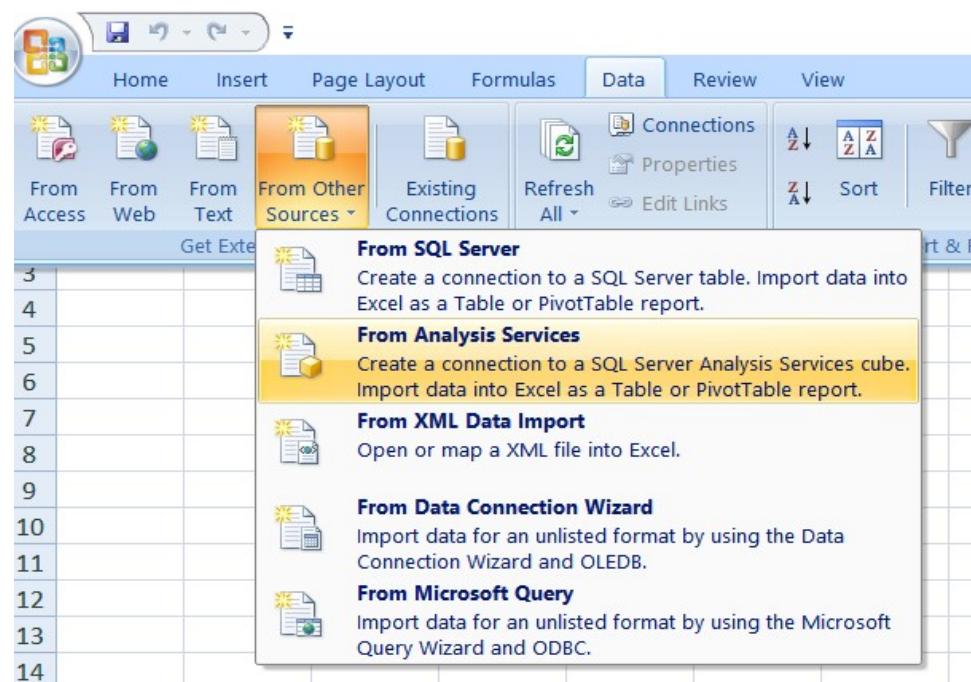
```
WITH MEMBER [Measures].[Sales Gross Margin] AS
    ([Measures].[Sales Amount]) - ([Measures].[Total Cost]) - ([Measures].[Return Amount]),
FORMAT_STRING = 'Percent'
SELECT
    [Measures].[Sales Gross Margin] ON 0,
    {[Dim Product].[Product Name].&[Contoso DVD Player M100 Black]} ON 1
FROM
    [Contoso Retail DW]
```

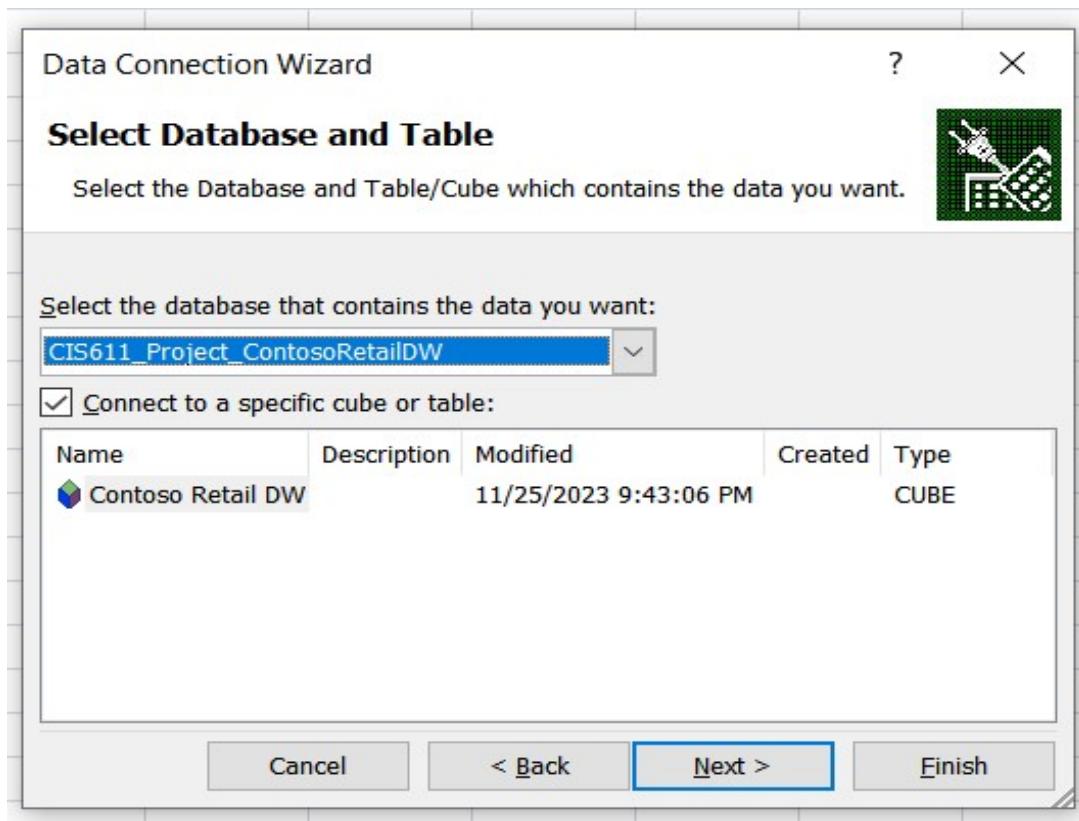
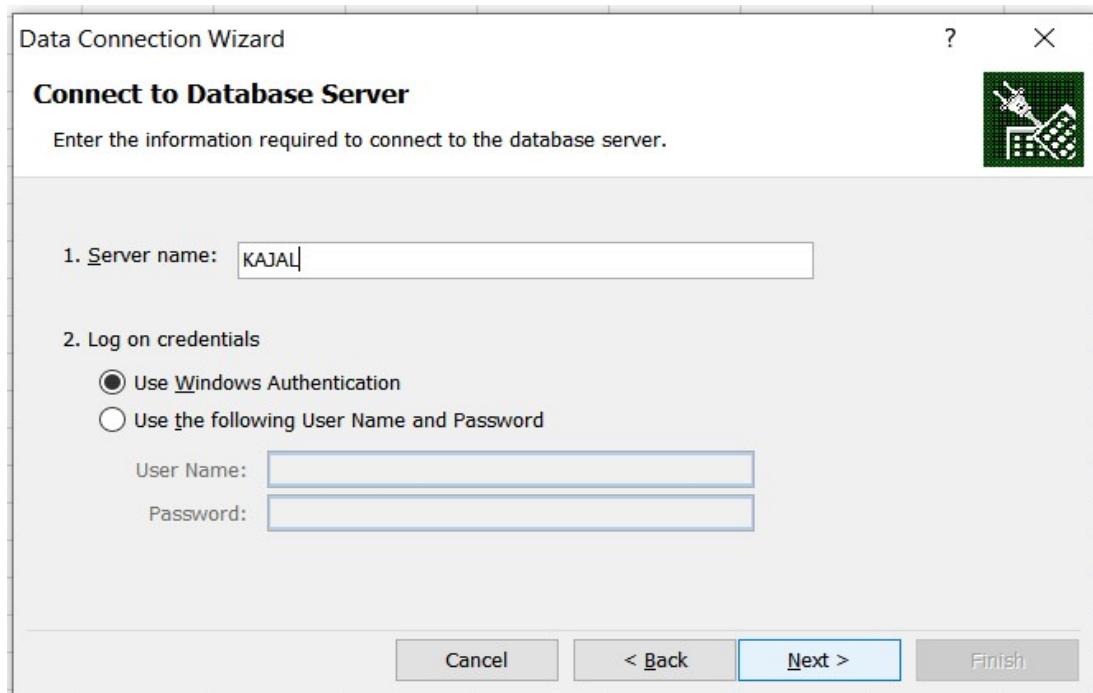
- In this we have calculated the Sales Gross Margin for the Contoso DVD Player M100 Black



For performing analysis on the developed CUBE, we have used the Excel Platform.

In this screenshot, we are importing the Datawarehouse CUBE which we have deployed in the analysis server.





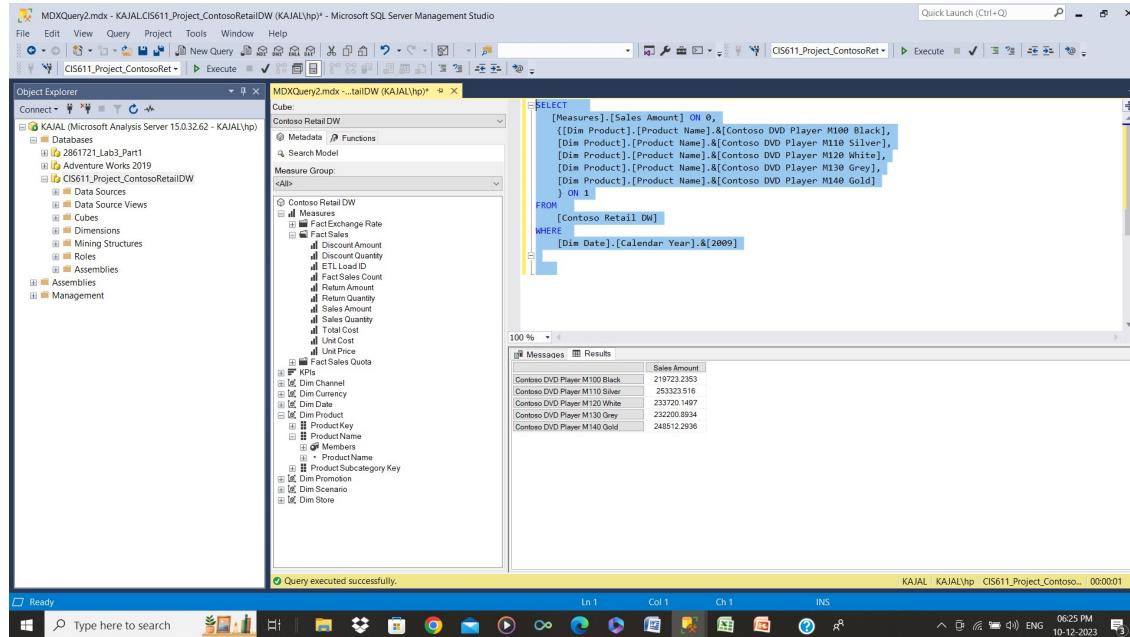
5. Analysis and Visualization over the CUBE: -

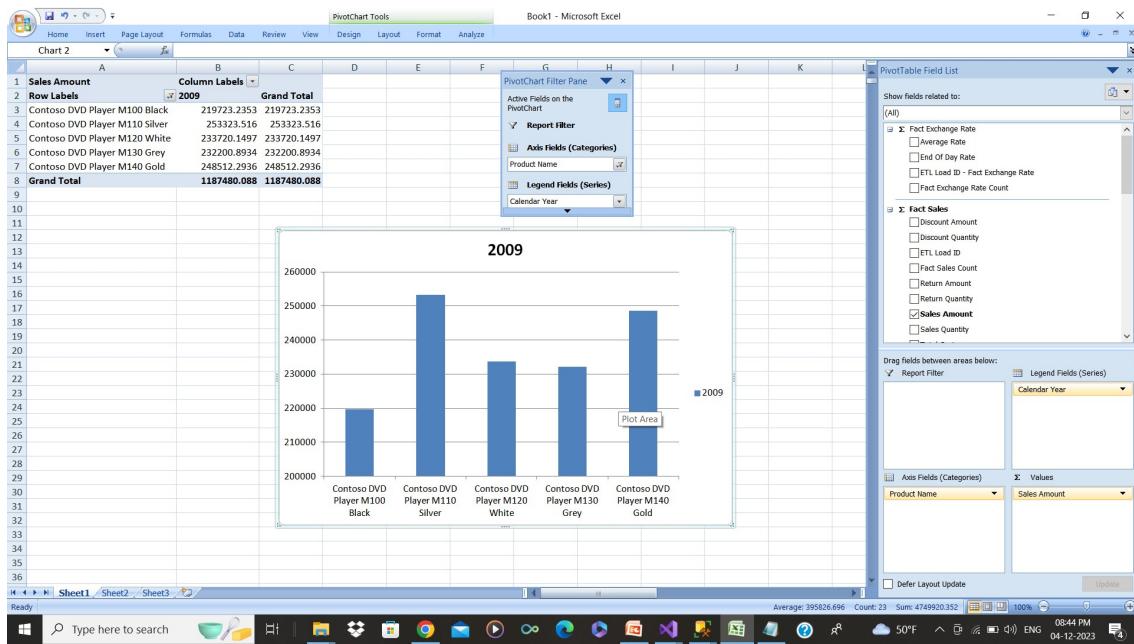
We have implemented a robust Business Intelligence (BI) and visualization solution by utilizing a Multidimensional OLAP Cube and Microsoft Data Tools using a selected Data Warehouse. It will provide stakeholders with powerful analytical capabilities and intuitive visualizations to enhance decision-making processes.

Query 1: Here, we are analyzing the sales of Contoso DVD Player M series in different colours for the year 2009.

Code Script:

```
SELECT  
    [Measures].[Sales Amount] ON 0,  
    {[Dim Product].[Product Name].&[Contoso DVD Player M100 Black],  
     [Dim Product].[Product Name].&[Contoso DVD Player M110 Silver],  
     [Dim Product].[Product Name].&[Contoso DVD Player M120 White],  
     [Dim Product].[Product Name].&[Contoso DVD Player M130 Grey],  
     [Dim Product].[Product Name].&[Contoso DVD Player M140 Gold]  
    } ON 1  
FROM  
    [Contoso Retail DW]  
WHERE  
    [Dim Date].[Calendar Year].&[2009]
```





Result:

From this we can analyse that, In 2009 the sales of Contoso DVD Player M110 Silver are the Highest.

Query 2: Here, we are analyzing the sales of SV Keyboard E series in different colours for different years 2007, 2008, 2009.

Code Script:

```

SELECT
    [Measures].[Sales Amount] ON 0,
    {[Dim Product].[Product Name].&[SV Keyboard E10 Black],
     [Dim Product].[Product Name].&[SV Keyboard E10 Grey],
     [Dim Product].[Product Name].&[SV Keyboard E10 Silver],
     [Dim Product].[Product Name].&[SV Keyboard E10 White],
     [Dim Product].[Product Name].&[SV Keyboard E90 Black],
     [Dim Product].[Product Name].&[SV Keyboard E90 Grey],
     [Dim Product].[Product Name].&[SV Keyboard E90 Silver],
     [Dim Product].[Product Name].&[SV Keyboard E90 White]
    } ON 1
FROM
    [Contoso Retail DW]
WHERE
    [Dim Date].[Calendar Year].members
  
```

MDXQuery2.mdx - KAJALCIS611_Project_ContosoRetailDW (KAJAL\hp) - Microsoft SQL Server Management Studio

File Edit View Query Project Tools Window Help

Execute

Object Explorer

Contoso Retail DW

Measures Group:

Cube: Contoso Retail DW

```

SELECT
    [Measures].[Sales Amount] ON 0,
    {[Dim Product].[Product Name].&[SV Keyboard E10 Black],
     [Dim Product].[Product Name].&[SV Keyboard E10 Grey],
     [Dim Product].[Product Name].&[SV Keyboard E10 Silver],
     [Dim Product].[Product Name].&[SV Keyboard E10 White],
     [Dim Product].[Product Name].&[SV Keyboard E90 Black],
     [Dim Product].[Product Name].&[SV Keyboard E90 Grey],
     [Dim Product].[Product Name].&[SV Keyboard E90 Silver],
     [Dim Product].[Product Name].&[SV Keyboard E90 White]
    } ON 1
FROM
    [Contoso Retail DW]
WHERE
    [Dim Date].[Calendar Year].members
  
```

100 %

g Messages Results

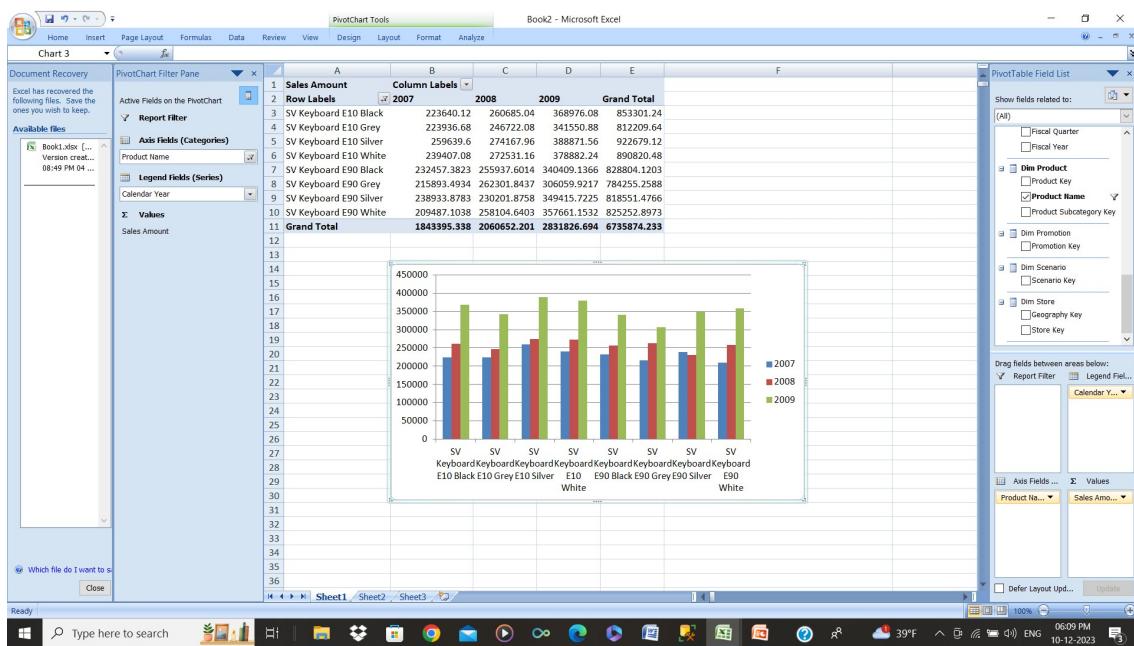
	Sales Amount
BV Keyboard E10 Black	85330124
BV Keyboard E10 Grey	85330944
BV Keyboard E10 Silver	82912042
BV Keyboard E10 White	80082048
BV Keyboard E90 Black	828041203
BV Keyboard E90 Grey	7642552588
BV Keyboard E90 Silver	8185514766
BV Keyboard E90 White	8252528973

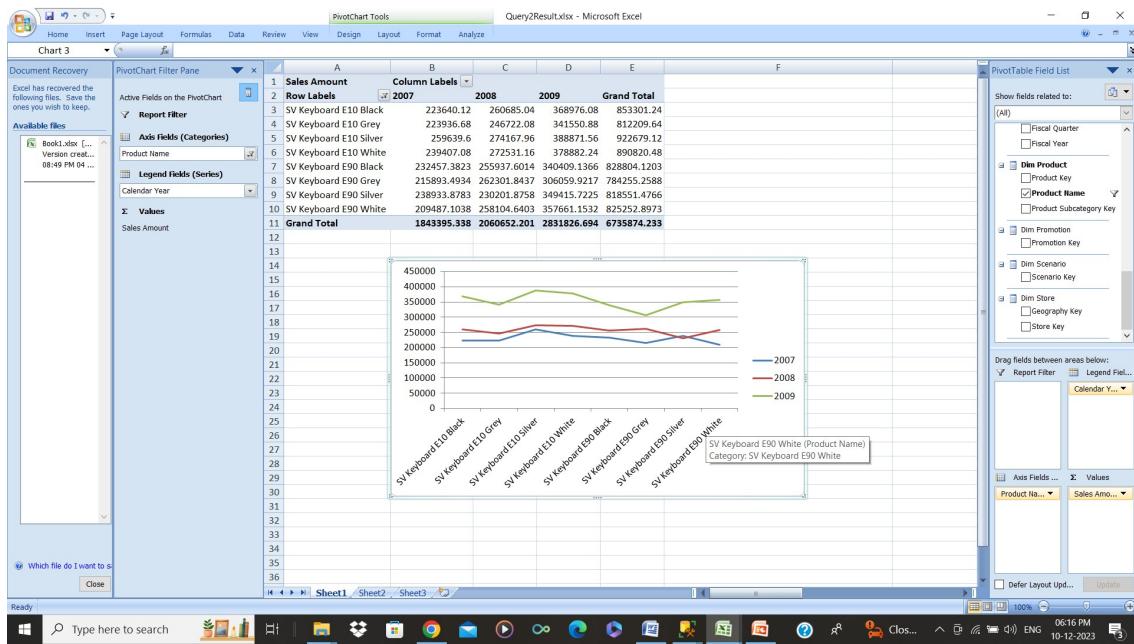
Query executed successfully.

KAJAL KAJAL\hp CIS611_Project_Contoso... 00:00:01

Ready Type here to search

06:20 PM 10-12-2023





From the above analysis, we can conclude that the sales of each of the E series Keyboard is lesser in 2007, and 2008 comparing to the sales in 2009.

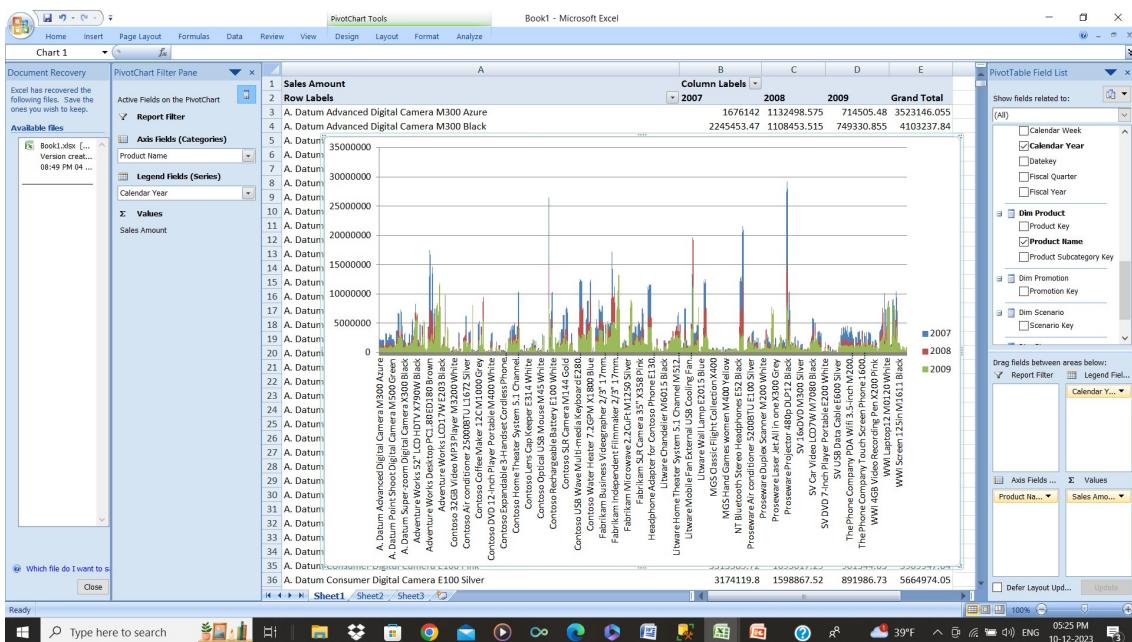
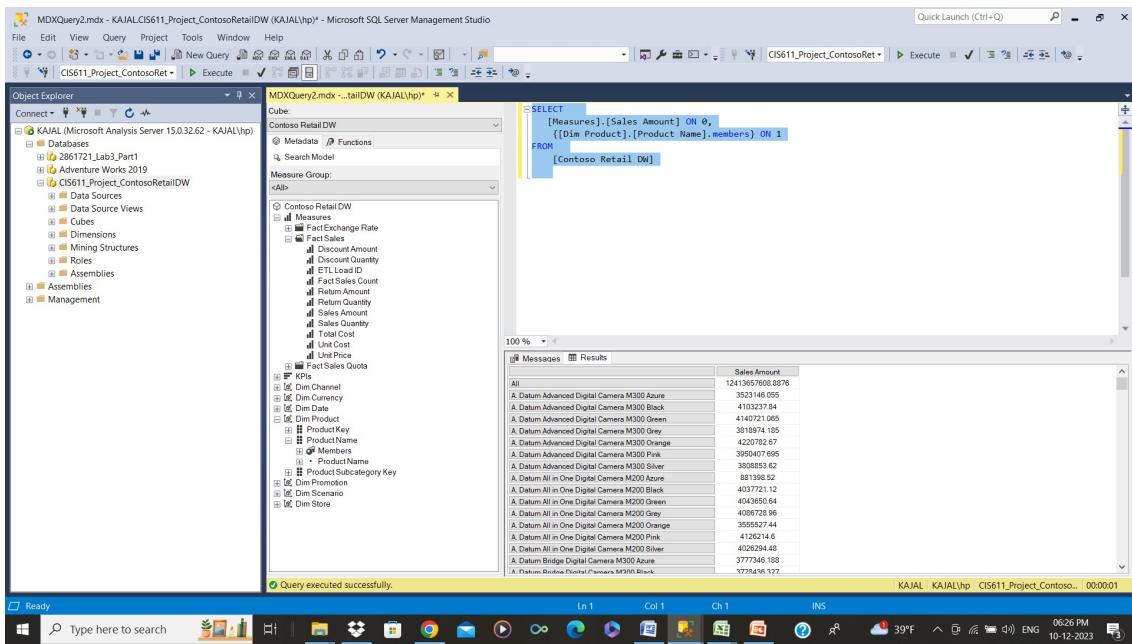
Query 3: We are analyzing the sales of different products in different years 2007, 2008 and 2009.

Code Script:

```

SELECT
    [Measures].[Sales Amount] ON 0,
    {[Dim Product].[Product Name].members} ON 1
FROM
    [Contoso Retail DW]

```

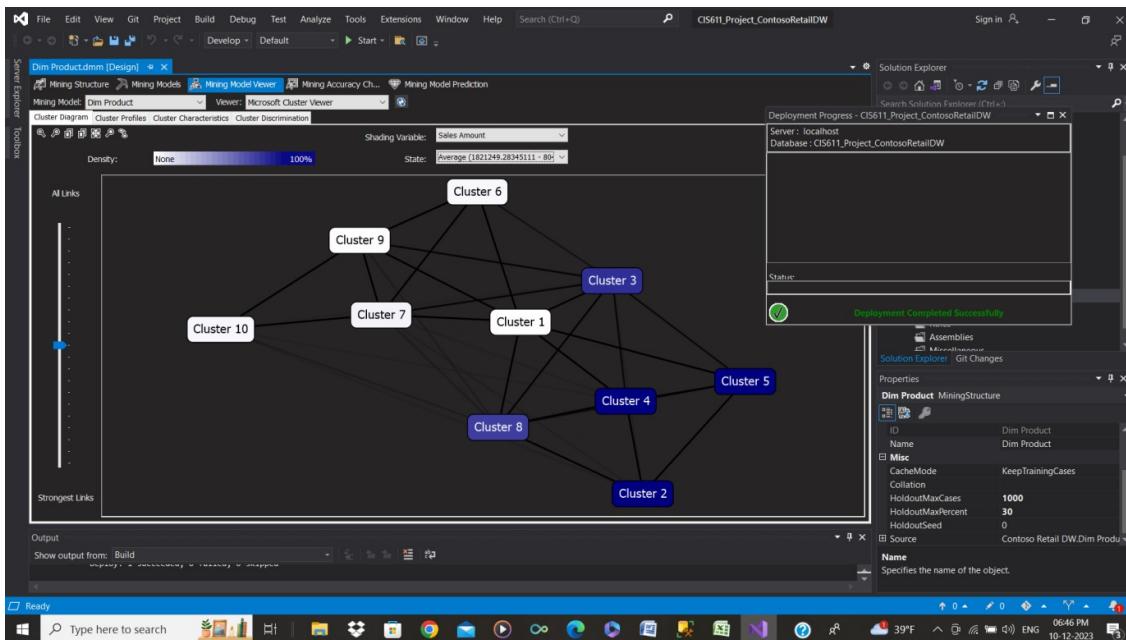
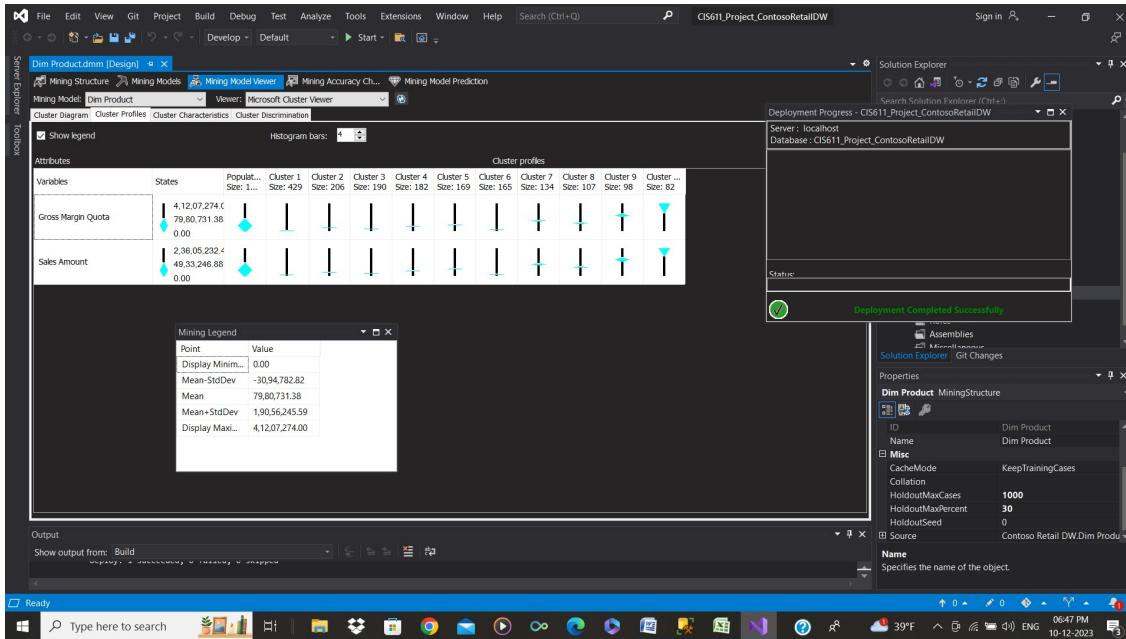


Result:

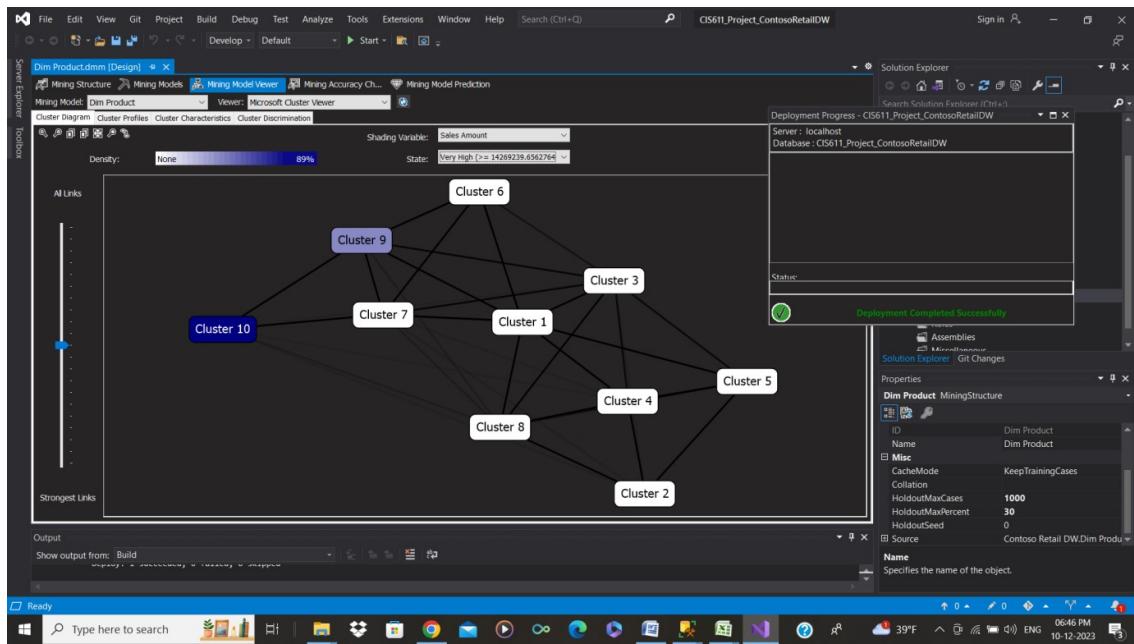
For this we can conclude that, the sales amount of Proseware Projector 480P DLP 12 Black is highest for the year 2007.

6. Data Mining Performed:

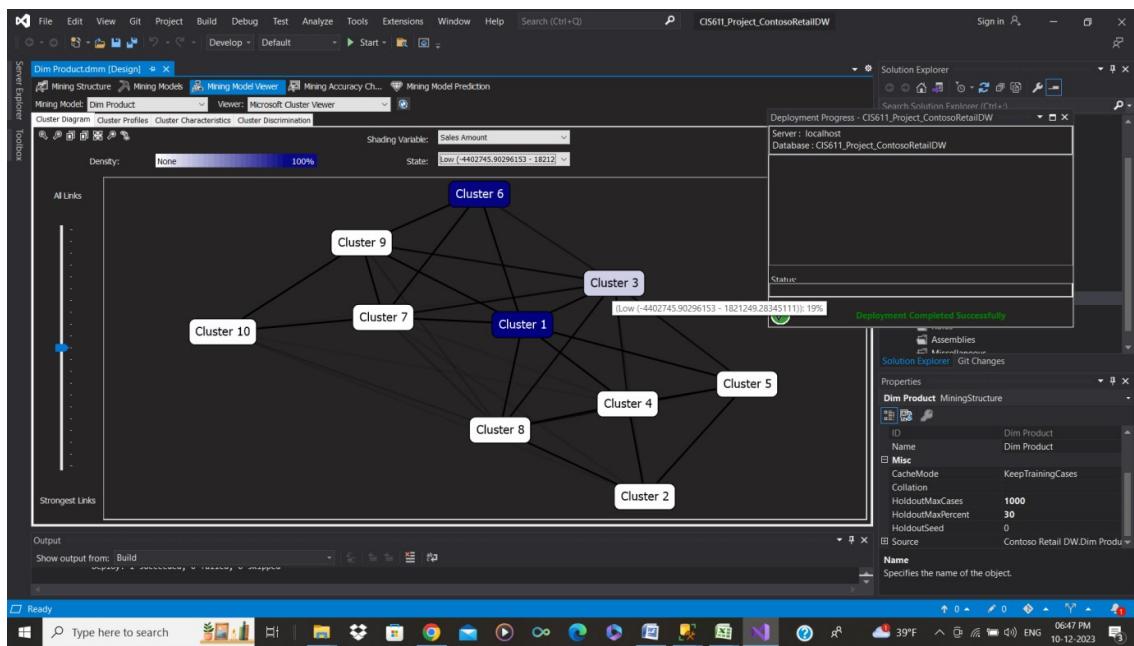
→ Here, we tried to perform data mining on the CUBE using the clustering techniques.



In this screenshot, we can observe the mining model for Dim_Product, we have 5 clusters that shows the average sales amount.



In this screenshot, we can observe the mining model for Dim_Product, we have 2 clusters that shows the very high sales amount.



In this screenshot, we can observe the mining model for Dim_Product, we have 2 clusters that shows the Low sales amount.

7 . References:

SQL Server 2012 Tutorials - Analysis Services Data Mining

<https://learn.microsoft.com/en-us/analysis-services/multidimensional-tutorial/lesson-2-2-defining-a-cube?view=asallproducts-all>

<https://learn.microsoft.com/en-us/sql/dmx/dmx-tutorials-analysis-services-data-mining?view=sql-server-ver16&version=16.0.0>