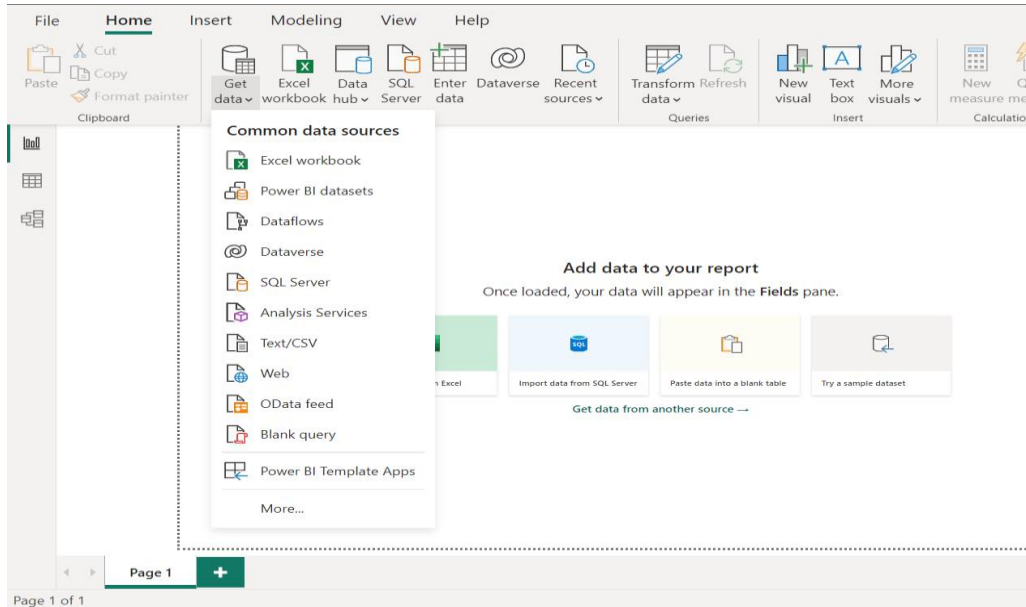


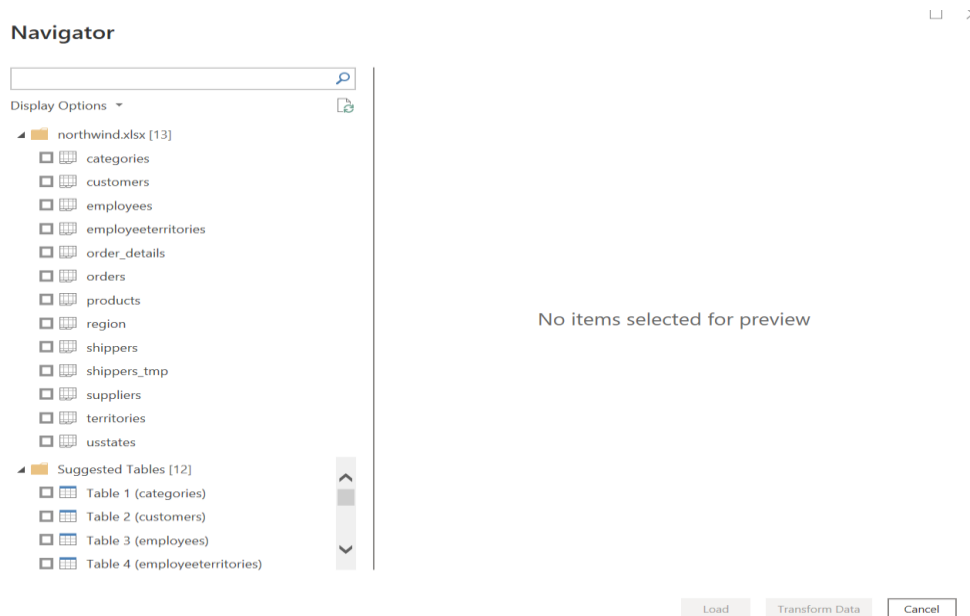
## Practical 01:

**Import the legacy data from different sources such as (SqlServer) and load in the target system. (sample database such as Adventureworks)**

**Step 1: Open Power BI & Click on Get data following list will be displayed → select Excel.**



**Step 2: Select required file and click on Open, Navigator screen appears, from which u can load the data.**



## Practical 02:

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

Step 1: Import the dataset and Select the necessary tables required.

The screenshot shows the Power BI Navigator window. On the left, under 'Display Options', the 'northwind.xlsx' file is listed. Below it, a tree view shows various tables: categories, customers, employees, employee territories, order\_details, orders, products, region, shippers, shippers\_tmp, suppliers, territories, and usstates. The 'products' table is selected. Below this, 'Suggested Tables' are listed: Table 1 (categories), Table 2 (customers), Table 3 (employees), and Table 4 (employee territories). On the right, the 'products' table is displayed with columns: productid, productname, supplierid, categoryid, and quantity. The table contains 22 rows of data. At the bottom, there are buttons for 'Load', 'Transform Data', and 'Cancel'.

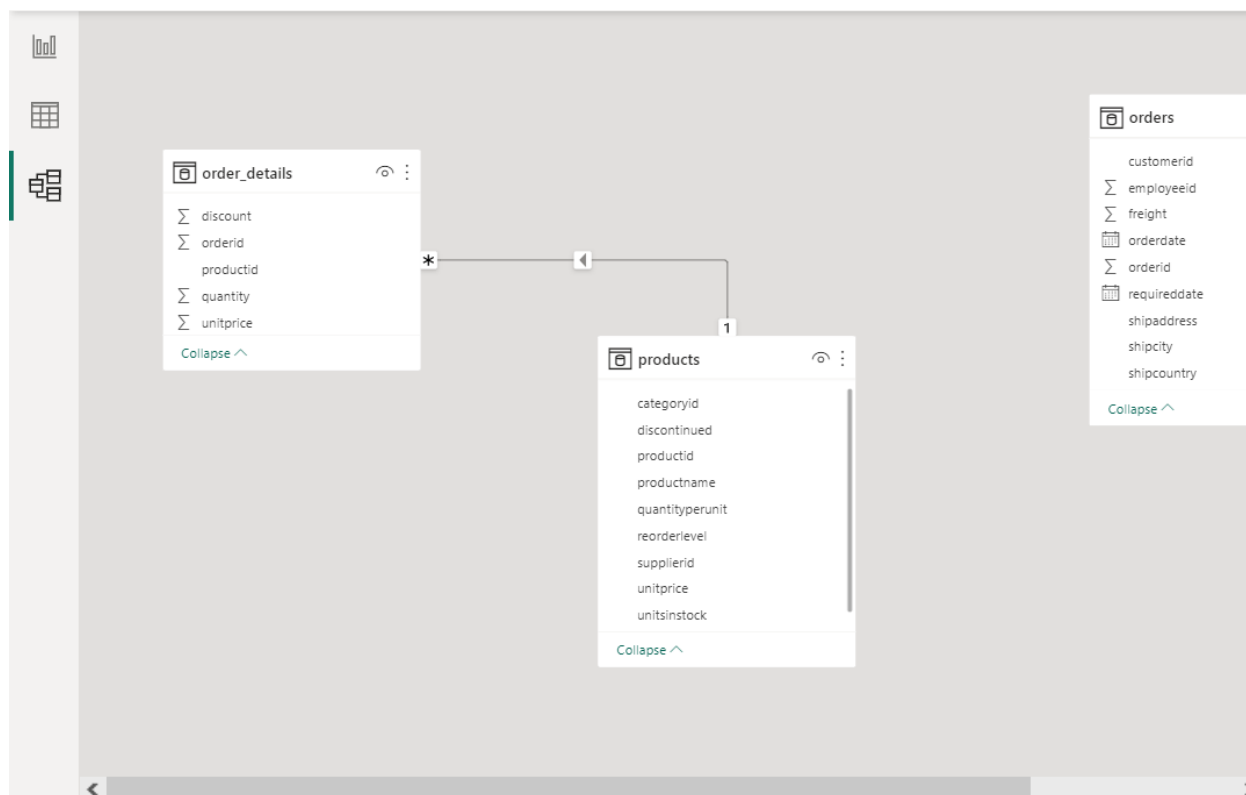
Step 2: Transform data and Managing relationship

### Create relationship

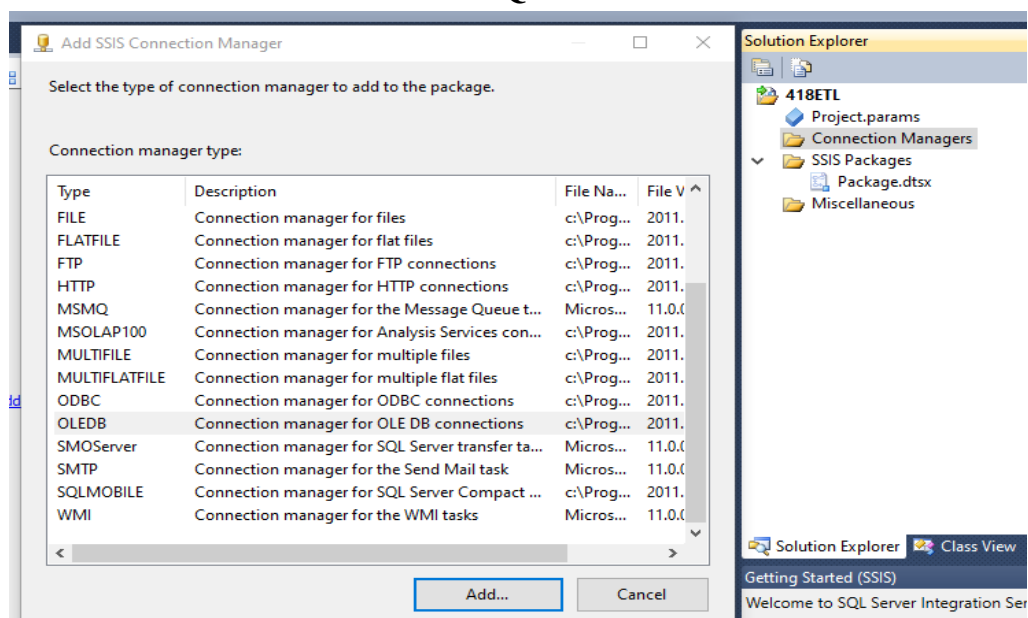
Select tables and columns that are related.

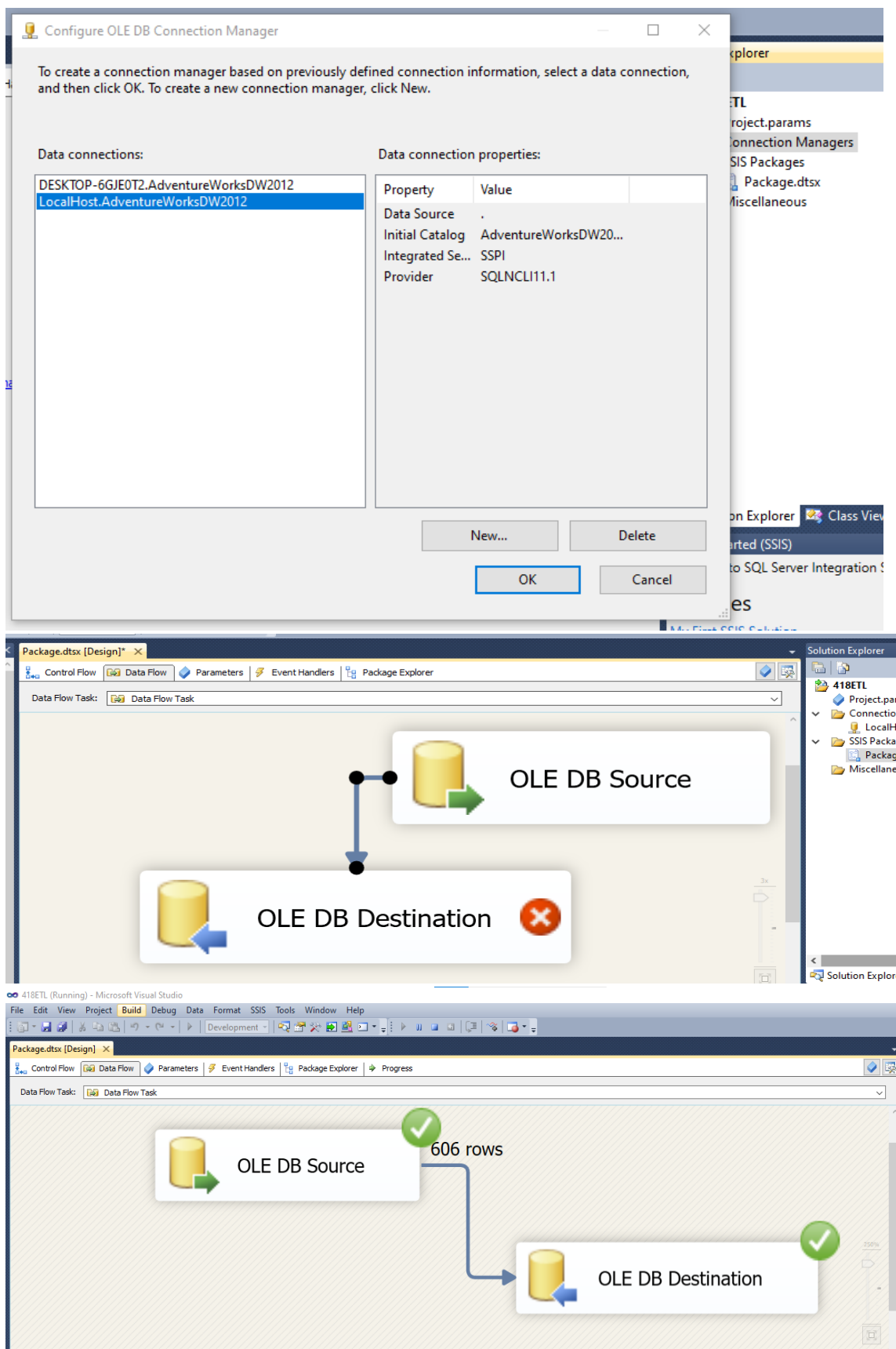
The screenshot shows the 'Create relationship' dialog in Power BI. It displays two tables: 'order\_details' and 'products'. The 'order\_details' table has columns: orderid, productid, unitprice, quantity, and discount. The 'products' table has columns: productid, productname, supplierid, categoryid, quantityperunit, unitprice, unitsinstock, and unitsonorder. The 'Cardinality' is set to 'Many to one (\*:1)' and the 'Cross filter direction' is set to 'Single'. The 'Make this relationship active' checkbox is checked. At the bottom, there are buttons for 'OK' and 'Cancel'.

Step 4: Table representation



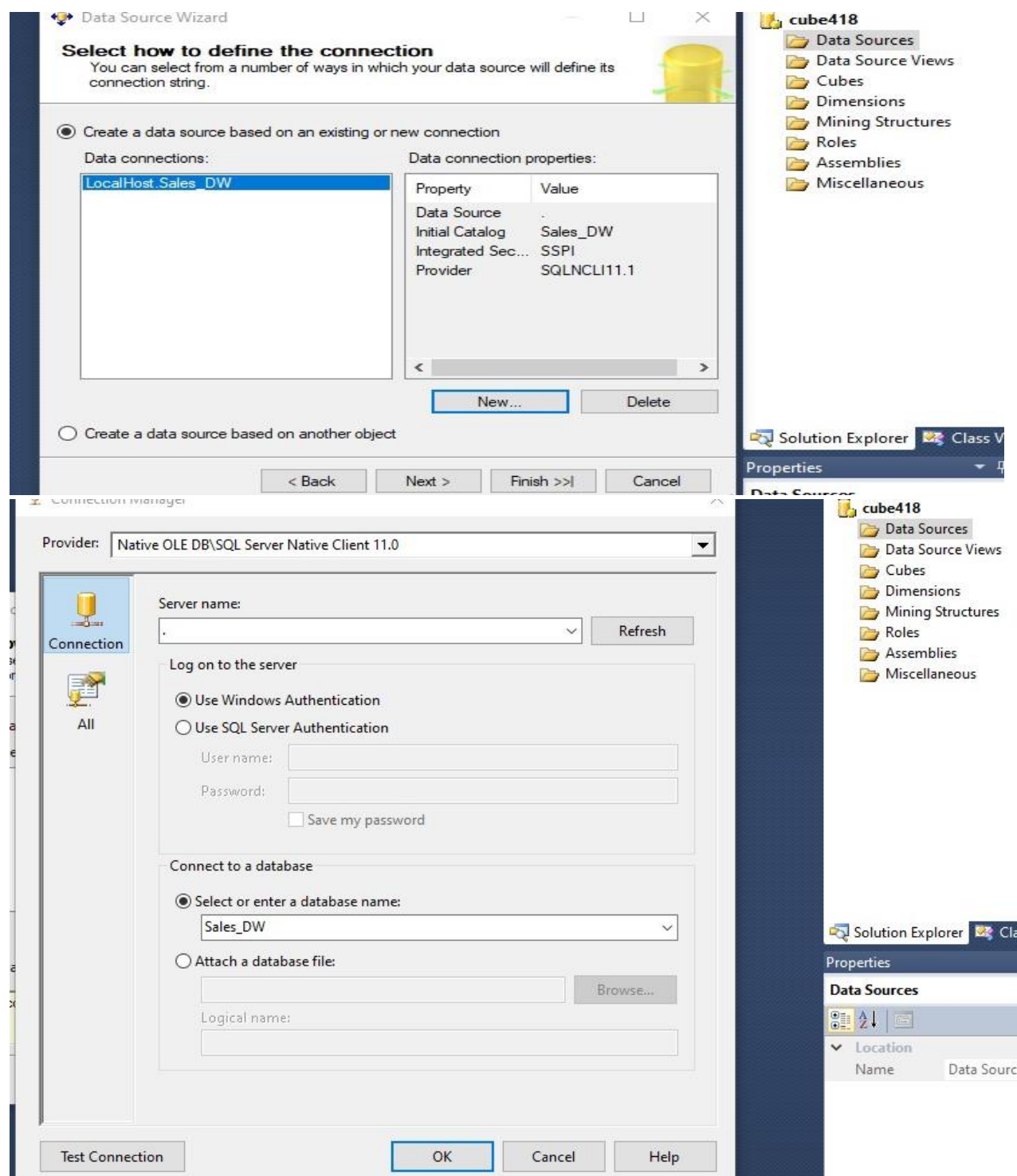
**B). Perform the Extraction Transformation and Loading (ETL) process to construct the database in the SQL Server.**





## Practical 03:

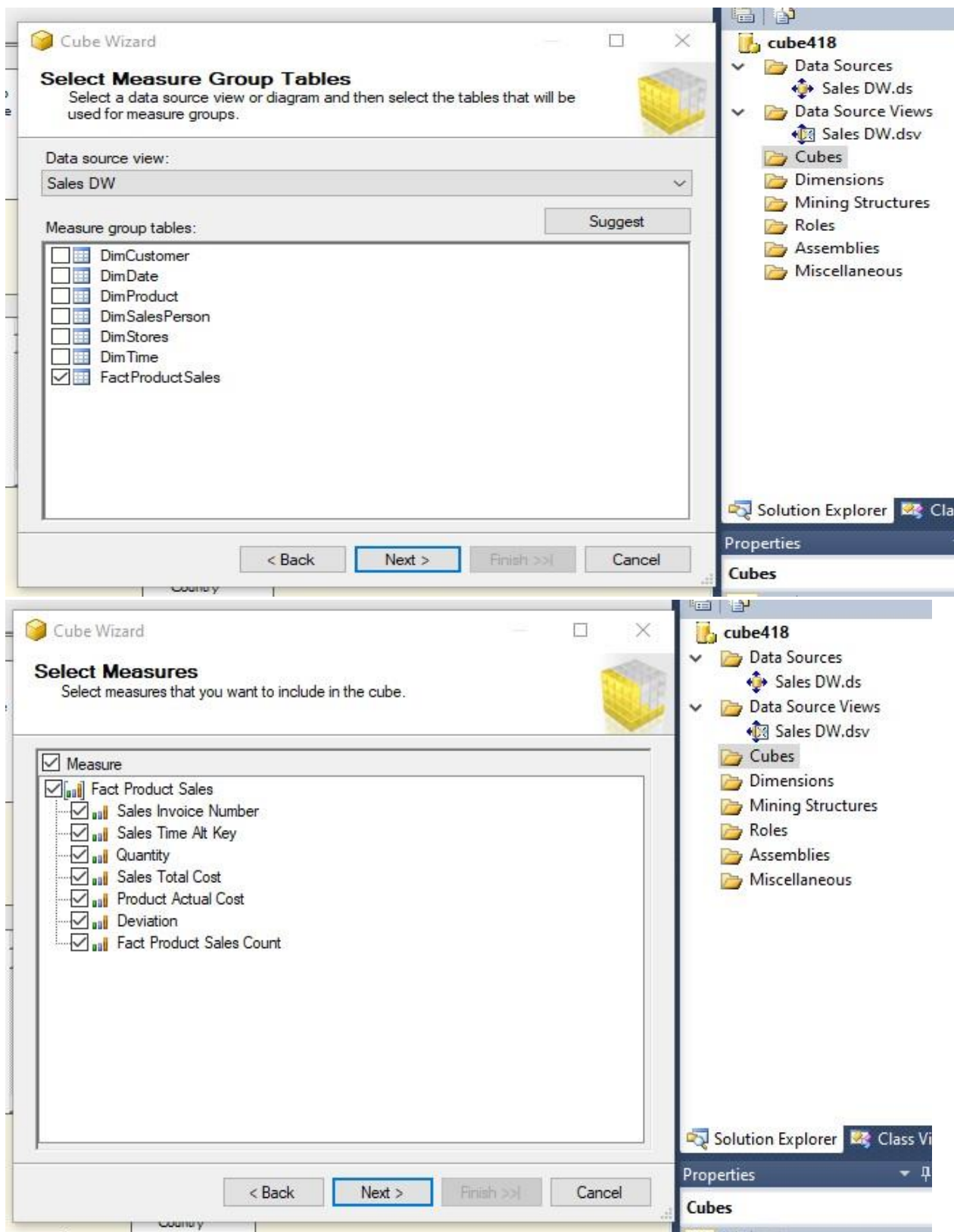
Create the cube with suitable dimension and fact tables based on OLAP.



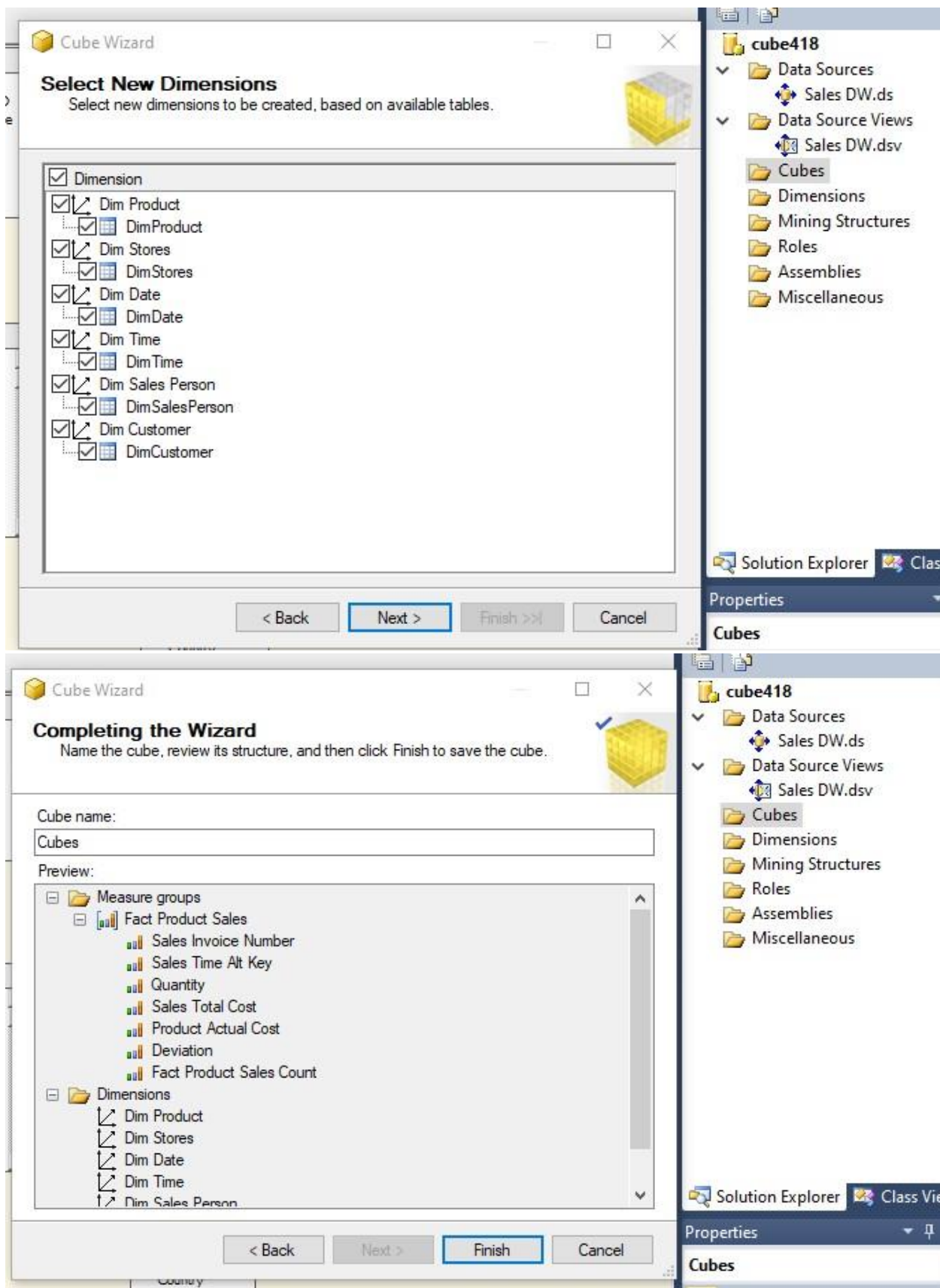
The top screenshot displays the 'Sales DW.dsv [Design]' window in SQL Server Data Tools. It shows a star schema diagram with a central fact table, **FactProductSales**, and six dimension tables: **DimSalesPerson**, **DimTime**, **DimStores**, **DimDate**, **DimCustomer**, and **DimProduct**. The diagram illustrates the relationships between these tables, with **FactProductSales** at the center and dimensions connected to it via foreign key relationships.

The bottom screenshot shows the 'Cube Wizard' dialog box, which is used to create a cube. The 'Select Creation Method' section offers three options: 'Use existing tables' (selected), 'Create an empty cube', and 'Generate tables in the data source'. The 'Template' dropdown is set to '(None)'. The 'Description' field contains the text: 'Create a cube based on one or more tables in a data source.' The 'Next >' button is highlighted, indicating the next step in the wizard.

On the right side of the bottom screenshot, the 'Solution Explorer' pane shows the project structure for 'cube418'. It includes folders for 'Data Sources', 'Data Source Views', 'Cubes', 'Dimensions', 'Mining Structures', 'Roles', 'Assemblies', and 'Miscellaneous'. The 'Sales DW.dsv' is listed under 'Data Source Views'.









The image displays three screenshots of the Microsoft SQL Server Data Tools (SSDT) interface, illustrating the design and deployment of a cube.

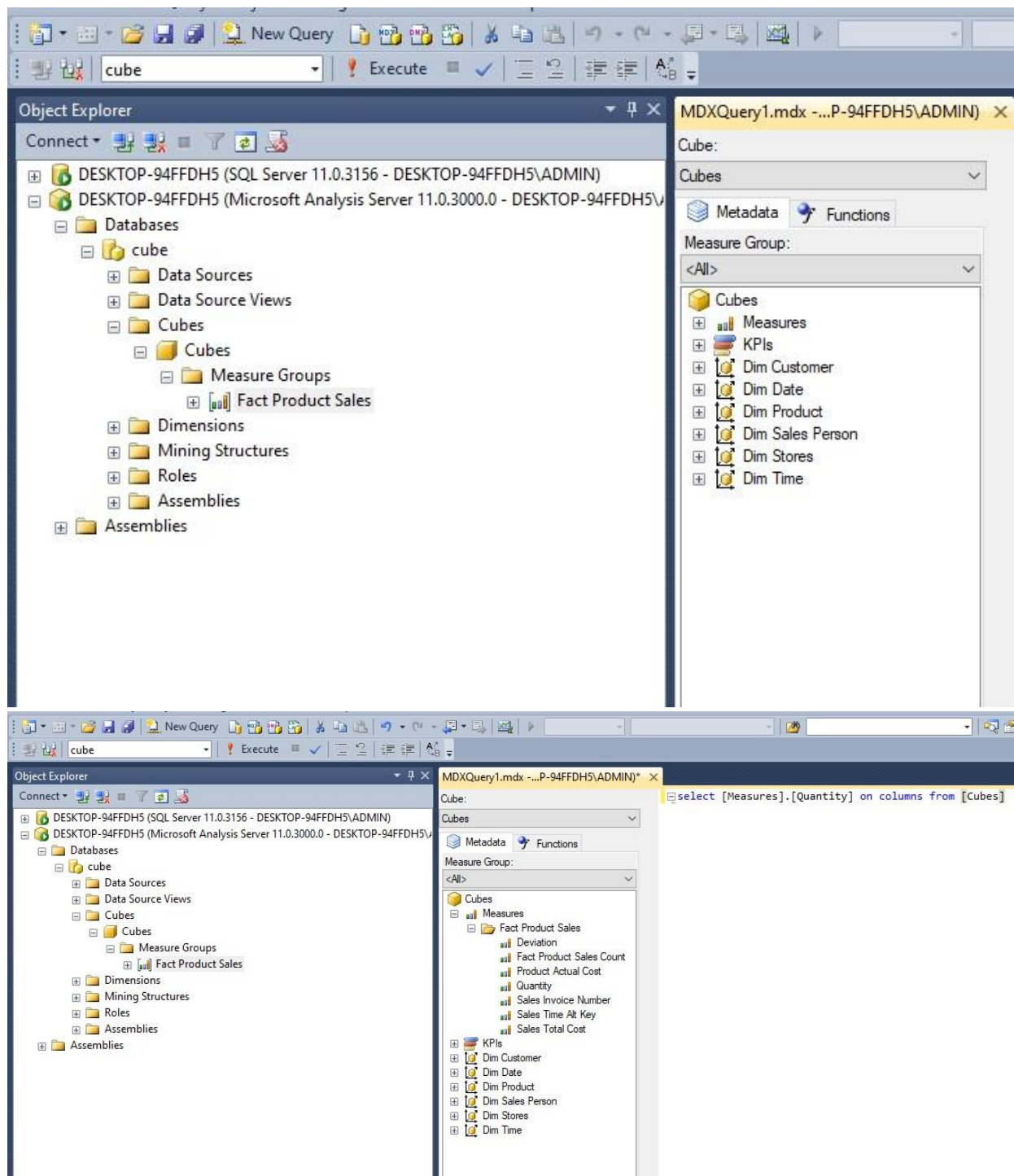
**Top Screenshot: Sales DW.dsv [Design]**  
 This screenshot shows the Data Source View (DSV) for the Sales DW data source. The central diagram illustrates the FactProductSales fact table, which is connected to several dimension tables: DimDate, DimSalesPerson, DimProduct, DimCustomer, DimStores, and DimTime. Each dimension table lists its attributes, such as DateKey, SalesPersonID, ProductKey, CustomerID, StoreID, and TimeKey. The left pane shows the 'Measures' list with 'Fact Product Sales' selected. The right pane shows the 'Solution Explorer' with the project structure, including Data Sources, Data Source Views, Cubes, and Dimensions. The 'Properties' window on the far right shows the 'Sales DW DataSourceView' properties.

**Middle Screenshot: Dim Date.dim [Design]**  
 This screenshot shows the design of the Dim Date dimension. The 'Attributes' list on the left includes Date Key, Full Date UK, Month Name, Quarter, Quarter Name, Week Of Month, and Year. The 'Hierarchies' pane shows a hierarchy for 'Date' with levels: Year, Quarter Name, Month Name, Week Of Month, and Full Date UK. The 'Data Source View' pane on the right lists the attributes of the FactProductSales fact table. The right pane shows the 'Solution Explorer' and the 'Properties' window for the 'Dim Date.dim' dimension.

**Bottom Screenshot: Deployment Progress**  
 This screenshot shows the 'Deployment Progress' dialog box, which displays the status of the deployment process. The 'Command' tab shows the following steps: Processing Cube 'Cubes' completed, Processing Measure Group 'Fact Product Sales' completed, Processing Dimension 'Dim Customer' completed, Processing Dimension 'Dim Date' completed, Processing Dimension 'Dim Product' completed, Processing Dimension 'Dim Sales Person' completed, Processing Dimension 'Dim Stores' completed, and Processing Dimension 'Dim Time' completed. The 'Status' section indicates 'Process succeeded.' The 'Stop', 'Reprocess', 'View Details...', and 'Copy' buttons are visible. The 'Close' and 'Help' buttons are also present. The 'Change Settings...' button is located at the bottom. The right pane shows the 'Solution Explorer' and the 'Properties' window for the 'Dim Date.dim' dimension.

## Practical 04:

Execute the MDX queries to extract the data from the data warehouse.



MDXQuery1.mdx -...P-94FFDH5\ADMIN)\* X

Cube: `select [Measures].[Quantity] on columns from [Cubes];`

Cubes

Metadata Functions

Measure Group: <All>

Cubes

- Measures
  - Fact Product Sales
    - Deviation
    - Fact Product Sales Count
    - Product Actual Cost
    - Quantity
    - Sales Invoice Number
    - Sales Time Alt Key
    - Sales Total Cost
- KPIs
- Dim Customer
- Dim Date
- Dim Product
- Dim Sales Person
- Dim Stores
- Dim Time

100 %

Messages Results

Quantity
86

MDXQuery1.mdx -...P-94FFDH5\ADMIN)\* X

Cube: `select [Measures].[Sales Total Cost] on columns from [Cubes];`

Cubes

Metadata Functions

Measure Group: <All>

Cubes

- Measures
  - Fact Product Sales
    - Deviation
    - Fact Product Sales Count
    - Product Actual Cost
    - Quantity
    - Sales Invoice Number
    - Sales Time Alt Key
    - Sales Total Cost
- KPIs
- Dim Customer
- Dim Date
- Dim Product
- Dim Sales Person
- Dim Stores
- Dim Time

100 %

Messages Results

Sales Total Cost
2463

MDXQuery1.mdx -...P-94FFDH5\ADMIN)\*

Cube: Cubes

Measure Group: <All>

Metadata Functions

Cubes

- Measures
  - Fact Product Sales
    - Deviation
    - Fact Product Sales Count
    - Product Actual Cost
    - Quantity
    - Sales Invoice Number
    - Sales Time Alt Key
    - Sales Total Cost
- KPIs
- Dim Customer
- Dim Date
- Dim Product
  - Product Key
    - Members
      - Product Key
- Dim Sales Person
- Dim Stores
- Dim Time

```
select [Measures].[Quantity] on columns,
[Dim Product].[Product Key].[Product Key] on rows
from [Cubes];
```

100 %

Messages Results

	Quantity
1	14
2	16
3	14
4	36
5	6

MDXQuery1.mdx -...P-94FFDH5\ADMIN)\*

Cube: Cubes

Measure Group: <All>

Metadata Functions

Cubes

- Measures
  - Fact Product Sales
    - Deviation
    - Fact Product Sales Count
    - Product Actual Cost
    - Quantity
    - Sales Invoice Number
    - Sales Time Alt Key
    - Sales Total Cost
- KPIs
- Dim Customer
  - Customer ID
    - Members
      - Customer ID
- Dim Date
- Dim Product
  - Product Key
    - Members
      - Product Key
- Dim Sales Person
- Dim Stores
- Dim Time

```
select {[Measures].[Quantity],[Measures].[Sales Total Cost]} on columns,
[Dim Customer].[Customer ID].[Customer ID] on rows
from [Cubes];
```

100 %

Messages Results

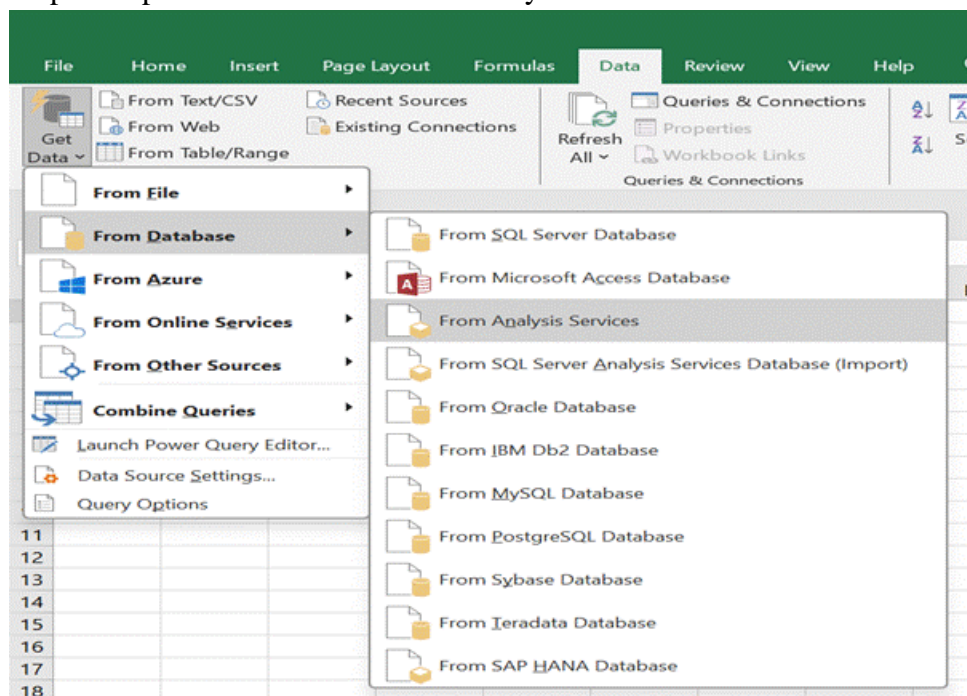
	Quantity	Sales Total Cost
1	36	1027
2	20	475
3	30	961
4	(null)	(null)
5	(null)	(null)



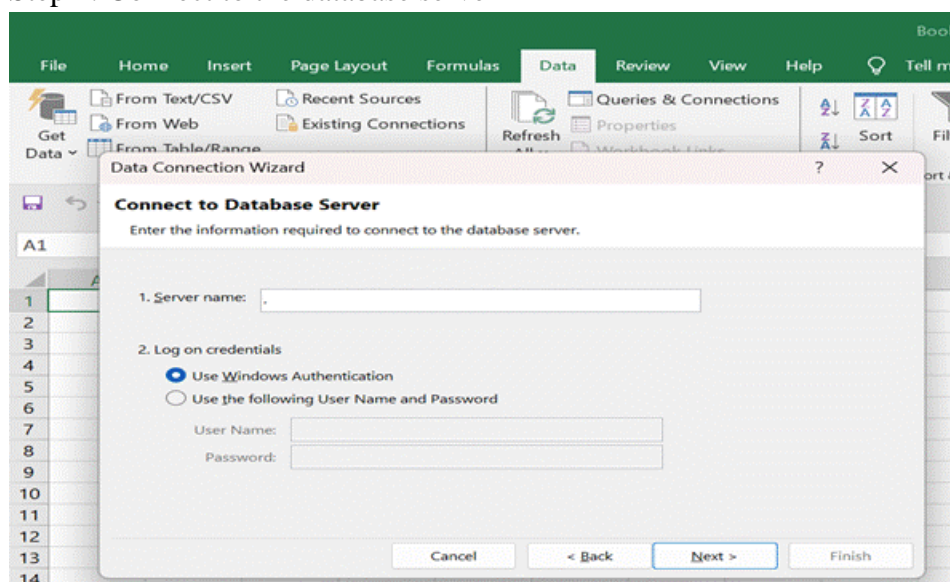
## Practical 05:

Import the cube in Microsoft Excel and create the Pivot table and Pivot Chart to perform data analysis.

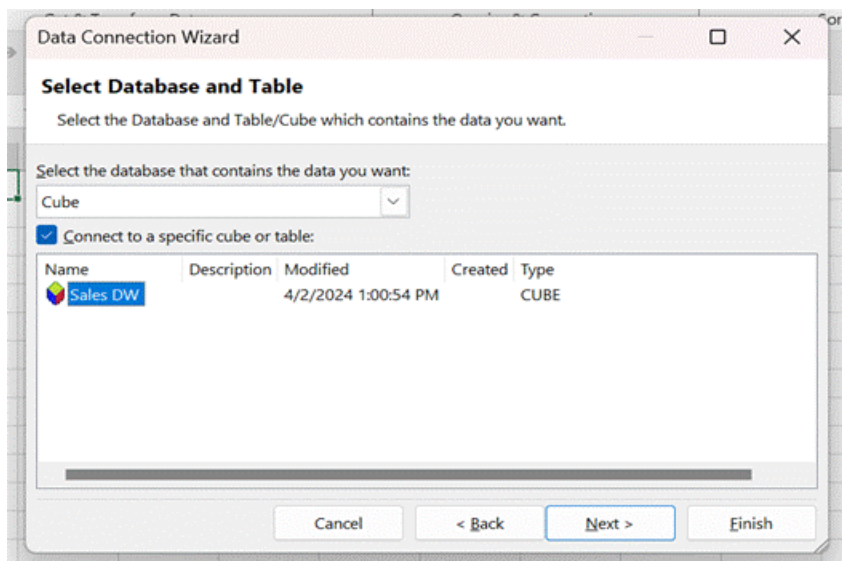
Step 1: Import the Cube from Data Analysis Services.



Step 2: Connect to the database server



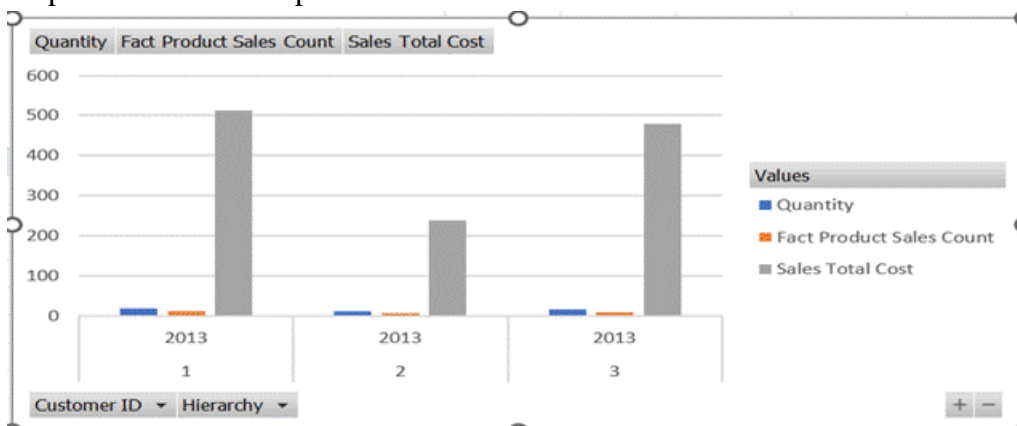
Step 3: Select the Cube created before



#### Step 4: Pivot Table report

	Fact Product Sales Count	Quantity	Sales Total Cost
<b>2013</b>			
1	4	7	45.5
2	6	8	104.5
3	6	7	304.5
4	6	18	360
5	3	3	417
<b>Grand Total</b>	<b>25</b>	<b>43</b>	<b>1231.5</b>

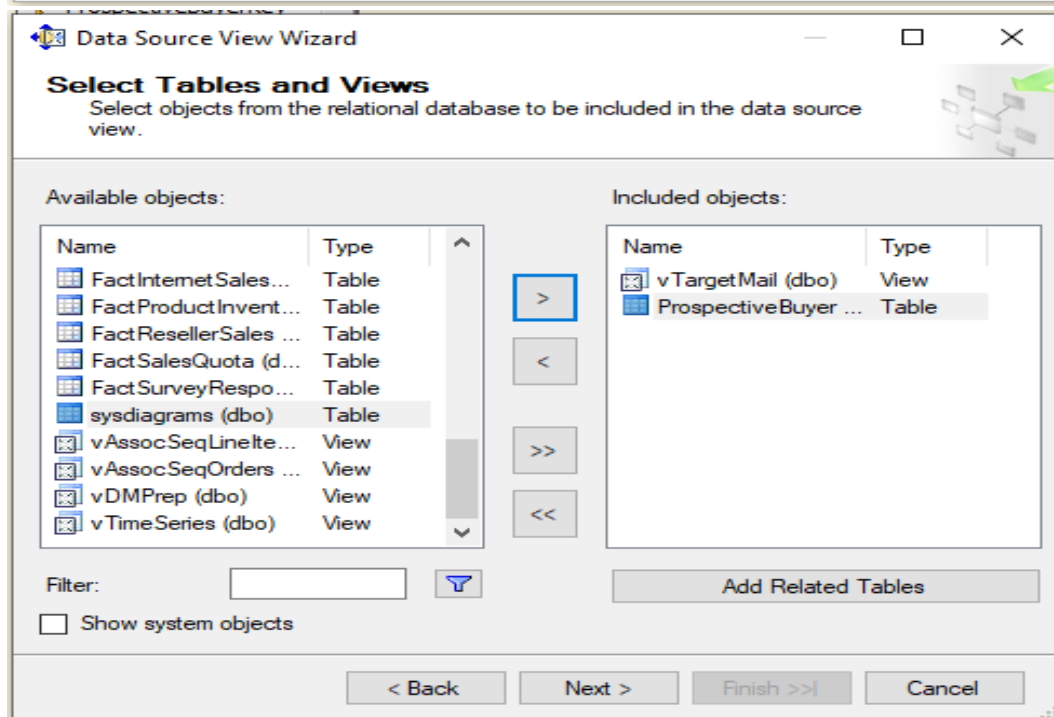
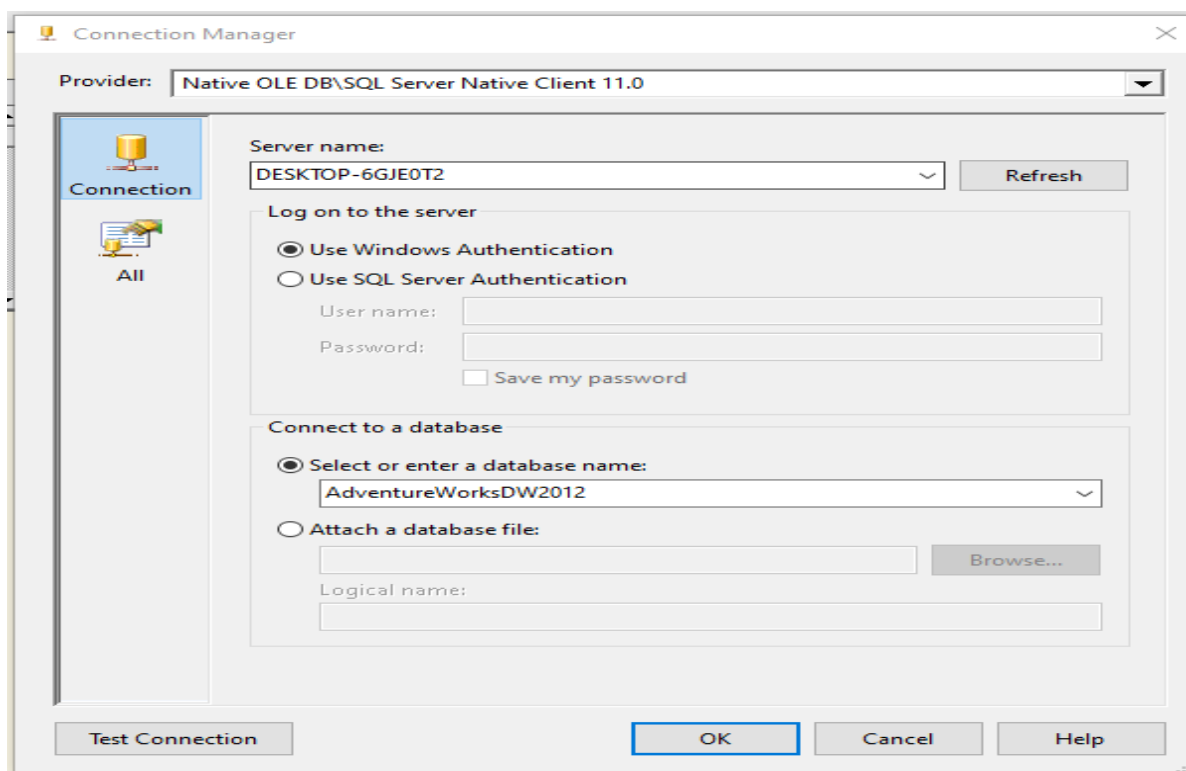
#### Step 5: Pivot Chart Report

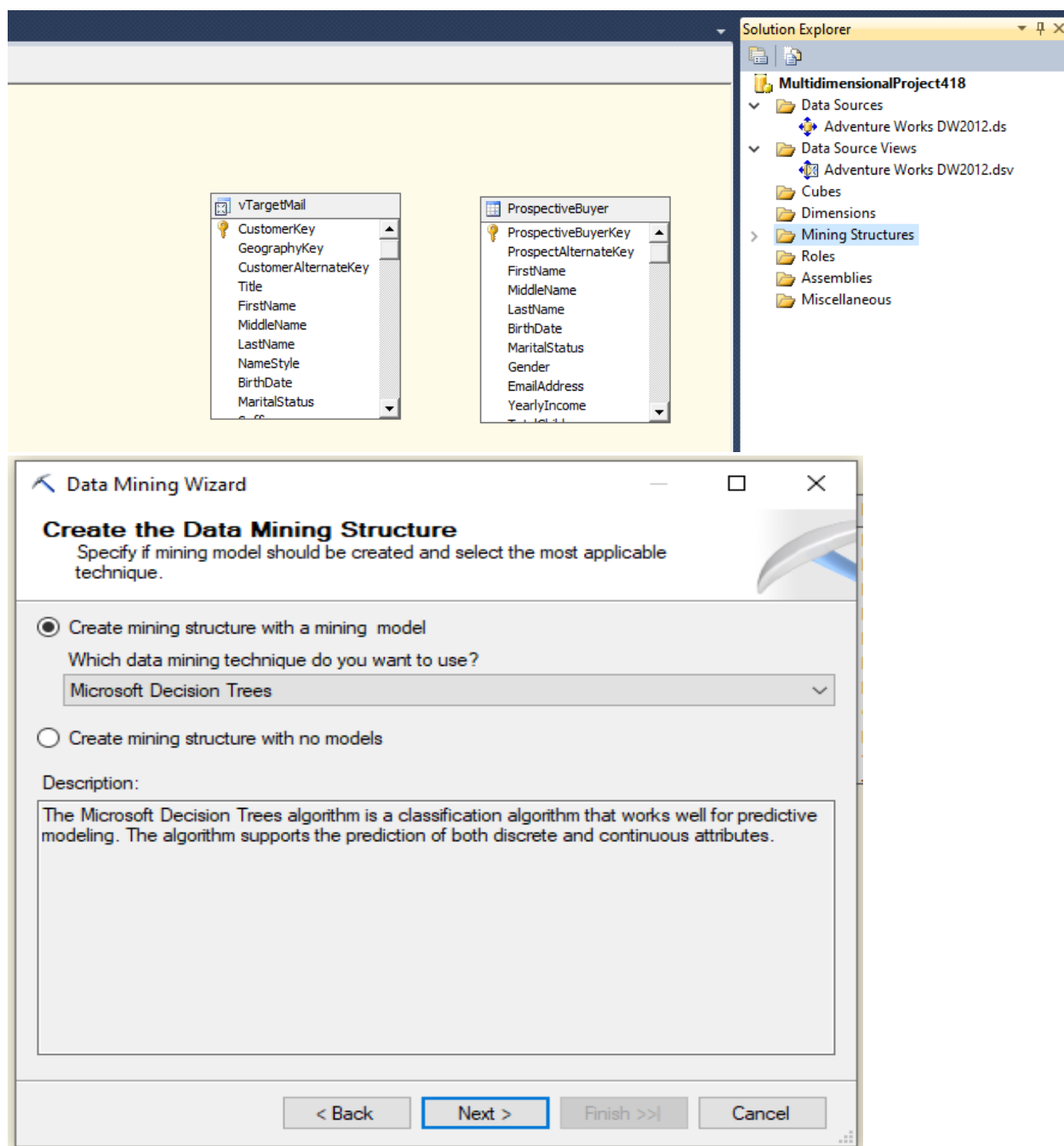




## Practical 07:

Perform the data classification using classification algorithm.





**Data Mining Wizard**

**Specify Table Types**  
Specify the type of tables to use for your analysis.

Input tables:

Tables	Case	Nested
Prospective Buyer	<input type="checkbox"/>	<input type="checkbox"/>
v Target Mail	<input checked="" type="checkbox"/>	<input type="checkbox"/>

< Back   **Next >**   Finish >>   Cancel

**Data Mining Wizard**


**Completing the Wizard**  
Completing the Data Mining Wizard by providing a name for the mining structure.

Mining structure name:

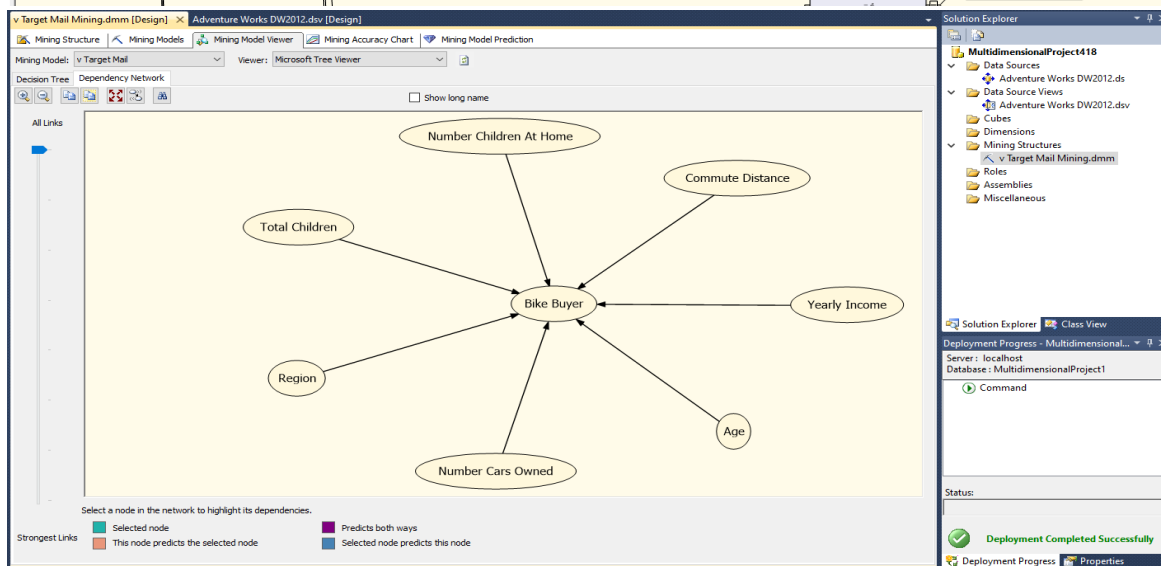
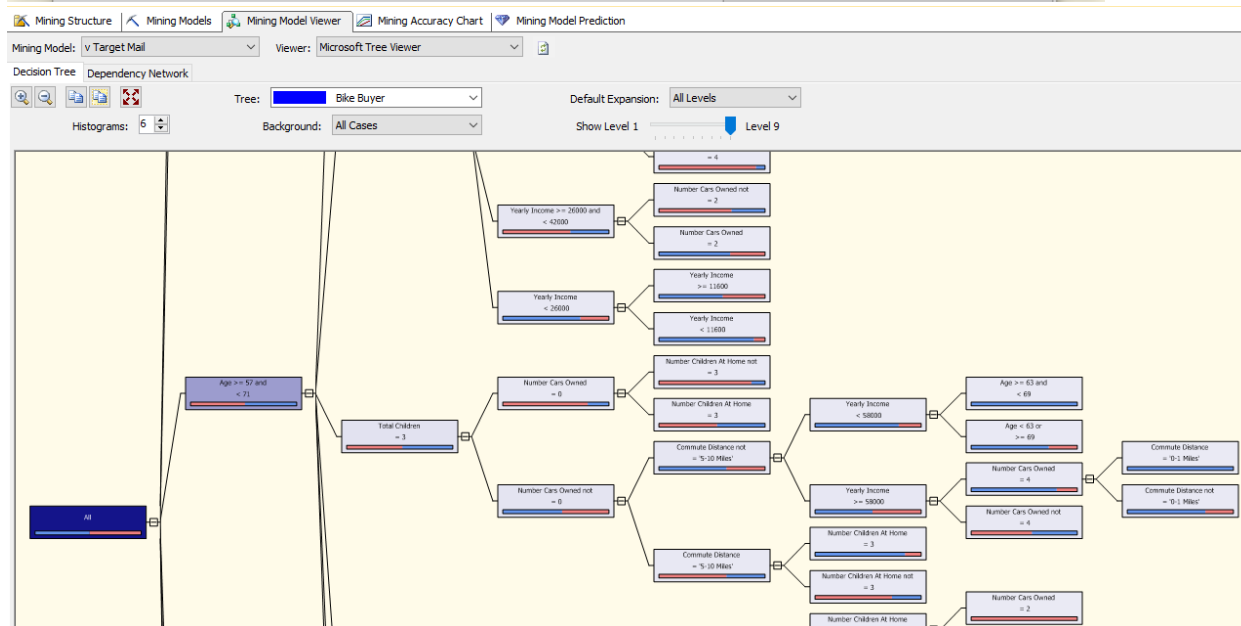
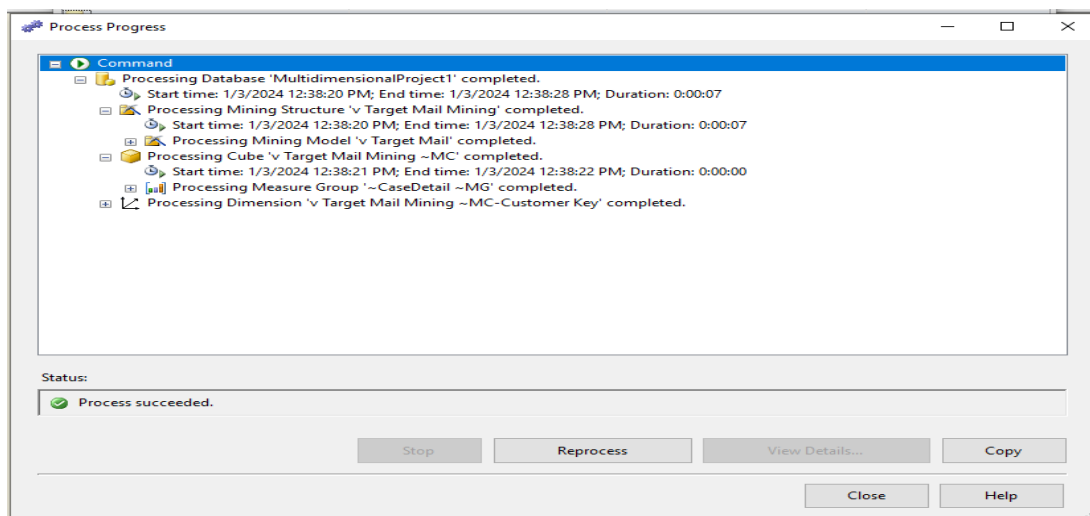
Mining model name:  
 ☒ Allow drill through

Preview:

- v Target Mail Mining
  - Columns
    - Age
    - Bike Buyer
    - Commute Distance
    - Customer Key

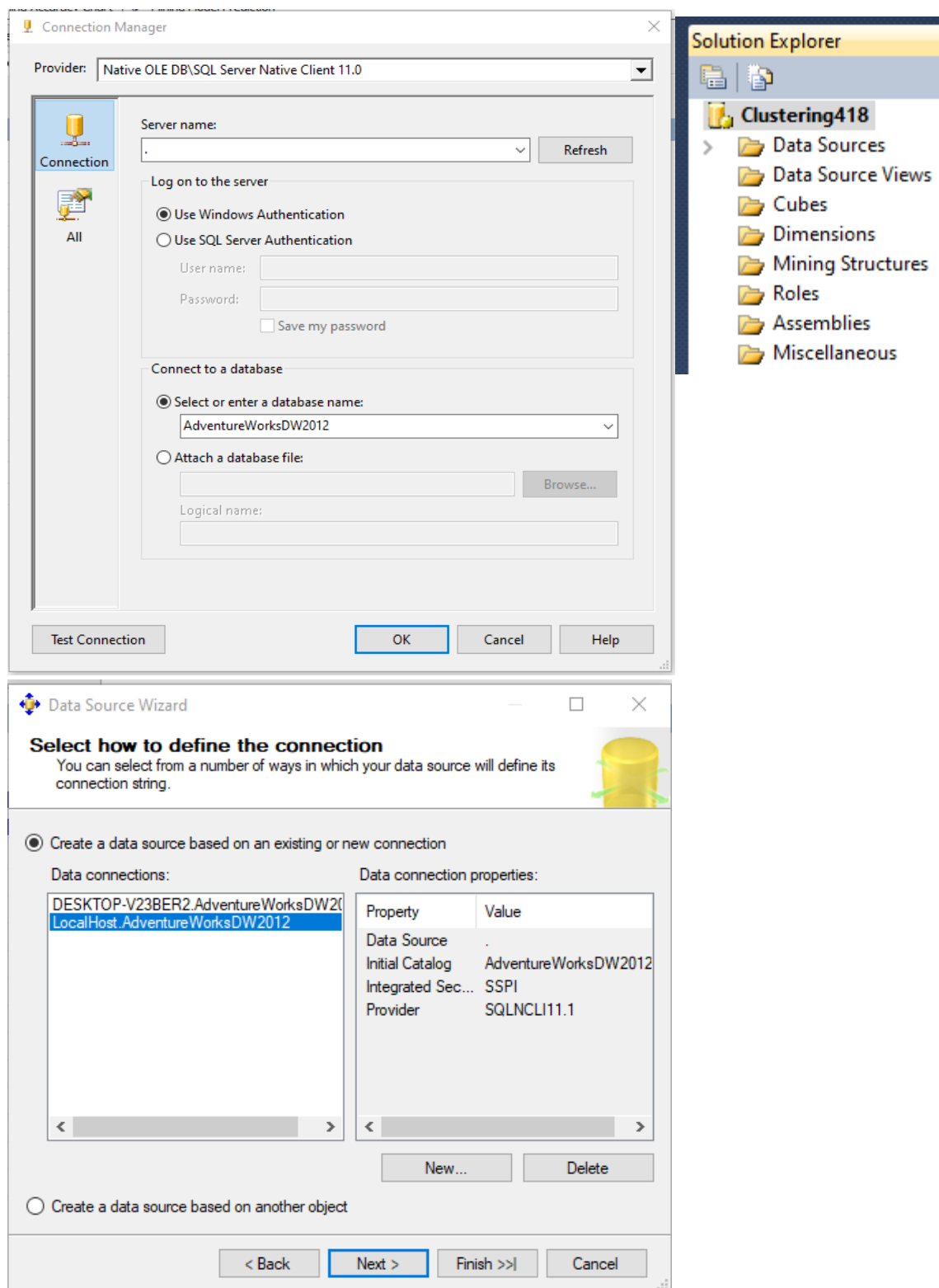
 Mining structure name error: Another 'MiningStructure' object has the 'v Target Mail Mining' name.

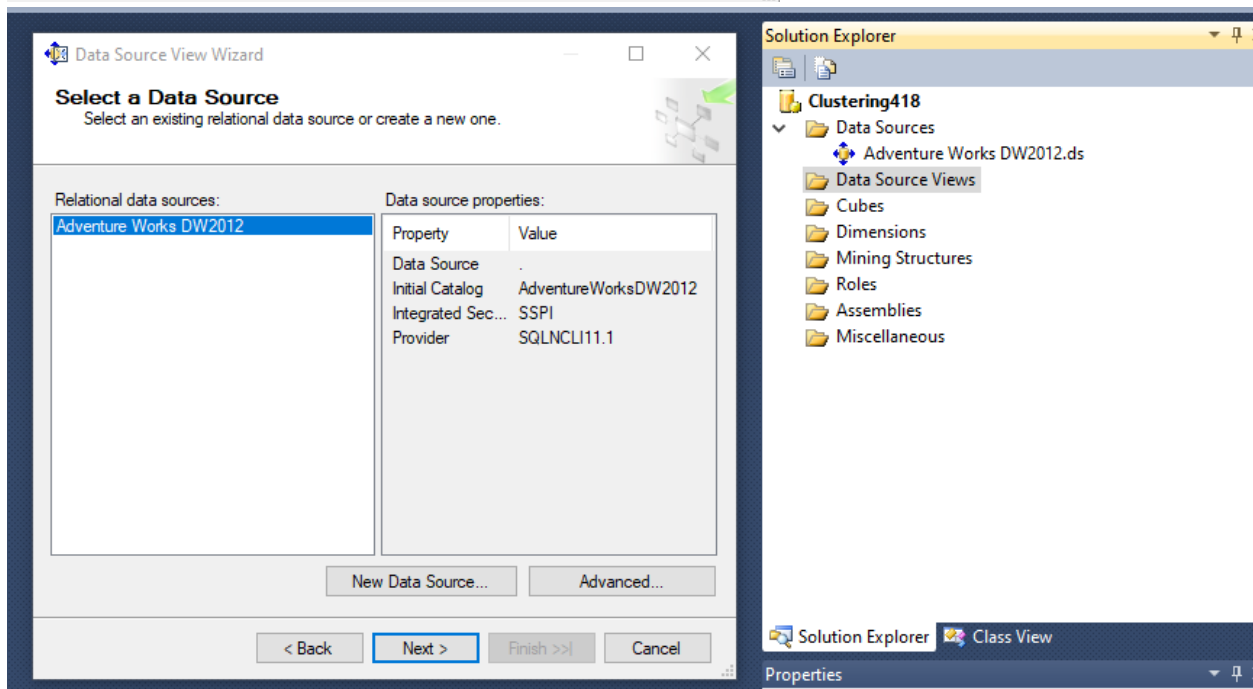
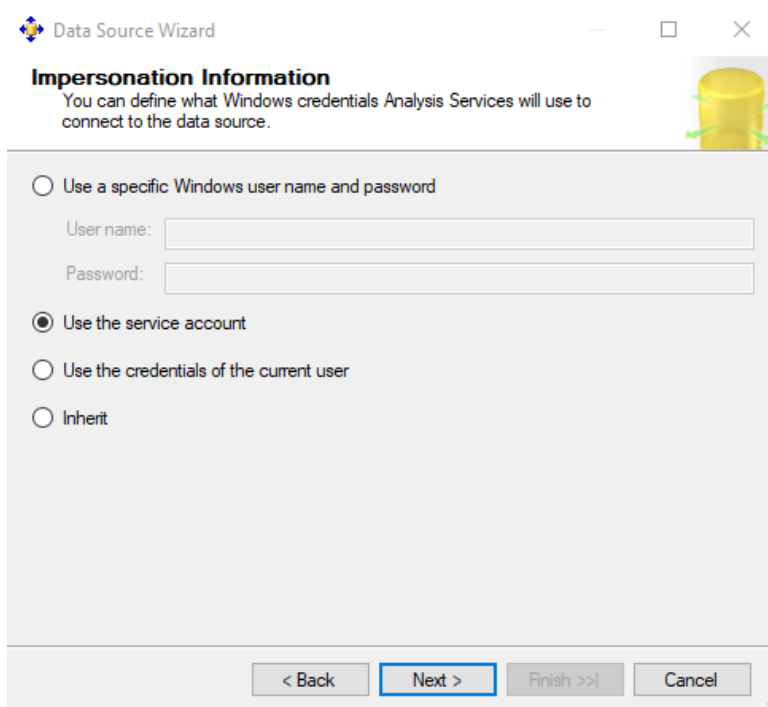
< Back   Next >   Finish   Cancel



## Practical 08:

Perform the data clustering using clustering algorithm.







The screenshot displays the SQL Server Data Source View Wizard and the Solution Explorer. The wizard is in the 'Select Tables and Views' step, showing available objects and included objects. The included objects are vTargetMail (dbo) View and ProspectiveBuyer (dbo) Table. The Solution Explorer shows the project structure for Clustering418, including Data Sources, Data Source Views, Cubes, Dimensions, Mining Structures, Roles, Assemblies, and Miscellaneous. The Data Source Views folder is expanded, showing the newly created Adventure Works DW2012.dsv.

**Data Source View Wizard - Select Tables and Views**

Select objects from the relational database to be included in the data source view.

**Available objects:**

Name	Type
DimScenario (dbo)	Table
FactAdditionalIntemati...	Table
FactCallCenter (dbo)	Table
FactCurrencyRate (dbo)	Table
FactFinance (dbo)	Table
FactInternetSales (dbo)	Table
FactInternetSalesReas...	Table
FactProductInventory (...)	Table
FactResellerSales (dbo)	Table
FactSalesQuota (dbo)	Table
FactSurveyResponse (...)	Table
sysdiagrams (dbo)	Table
vAssocSeqLineItems (...)	View
vAssocSeqOrders (dbo)	View
vDMPrep (dbo)	View
vTimeSeries (dbo)	View

**Included objects:**

Name	Type
vTargetMail (dbo)	View
ProspectiveBuyer (dbo)	Table

Filter:

☐ Show system objects

**Solution Explorer**

Clustering418

- Data Sources
  - Adventure Works DW2012.ds
- Data Source Views
  - Adventure Works DW2012.dsv
- Cubes
- Dimensions
- Mining Structures
- Roles
- Assemblies
- Miscellaneous

**Properties**

**Data Source Views**

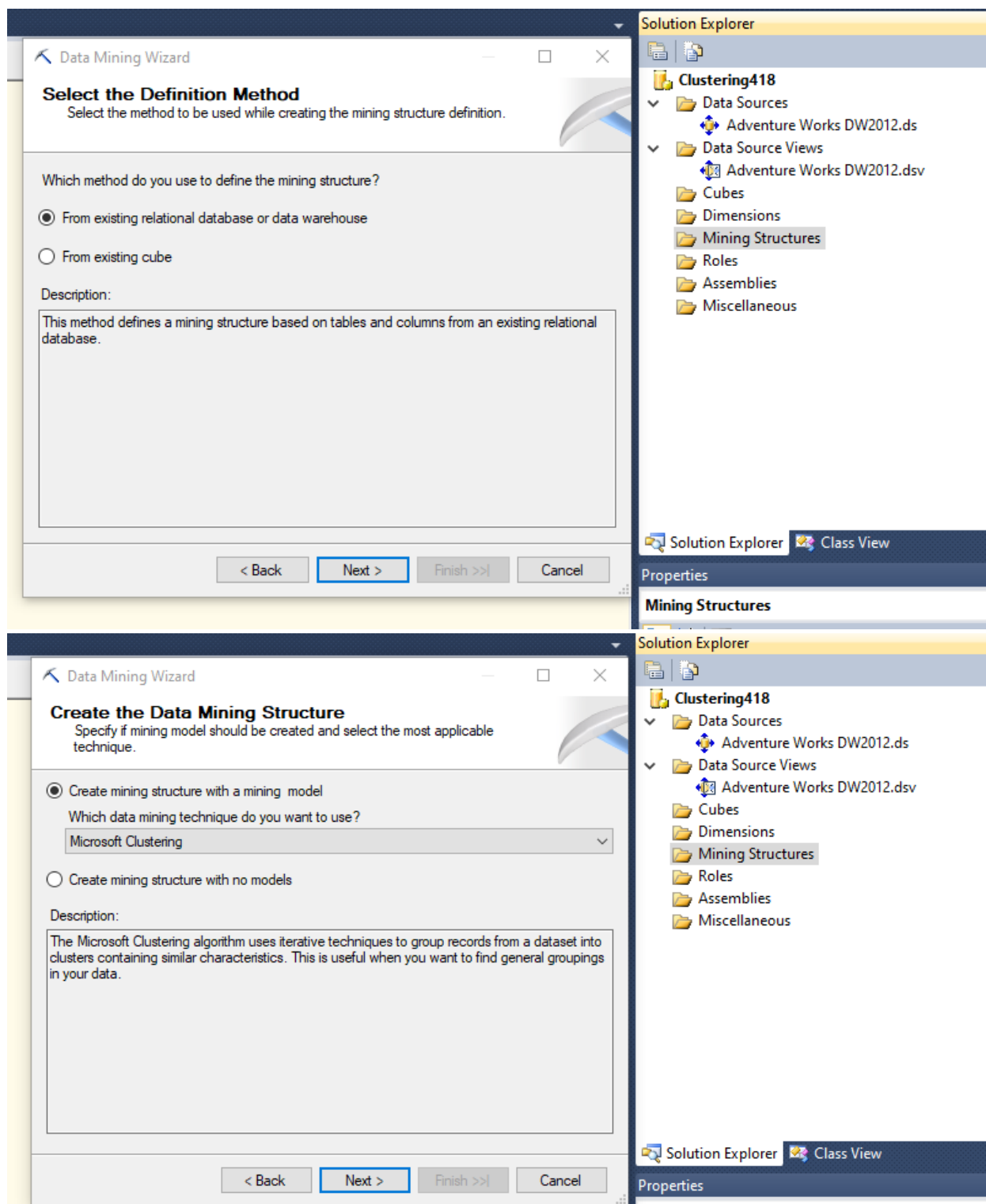
Location: Name Data Source Views

**ProspectiveBuyer**

ProspectiveBuyerKey
ProspectAlternateKey
FirstName
MiddleName
LastName
BirthDate
MaritalStatus
Gender
EmailAddress
YearlyIncome

**vTargetMail**

CustomerKey
GeographyKey
CustomerAlternateKey
Title
FirstName
MiddleName
LastName
NameStyle
BirthDate
MaritalStatus



**Data Mining Wizard**

**Select Data Source View**  
Select the data source view to provide the data for the mining structure.

Available data source views:

Adventure Works DW2012

Tables:

ProspectiveBuyer  
vTargetMail

Browse...

< Back   Next >   Finish >>   Cancel

**Solution Explorer**

**Clustering418**

- Data Sources
  - Adventure Works DW2012.ds
- Data Source Views
  - Adventure Works DW2012.dsv
- Cubes
- Dimensions
- Mining Structures
- Roles
- Assemblies
- Miscellaneous

**Properties**

**Mining Structures**

**Data Mining Wizard**

**Specify Table Types**  
Specify the type of tables to use for your analysis.

Input tables:

Tables	Case	Nested
ProspectiveBuyer	<input type="checkbox"/>	<input type="checkbox"/>
vTargetMail	<input checked="" type="checkbox"/>	<input type="checkbox"/>

< Back   Next >   Finish >>   Cancel

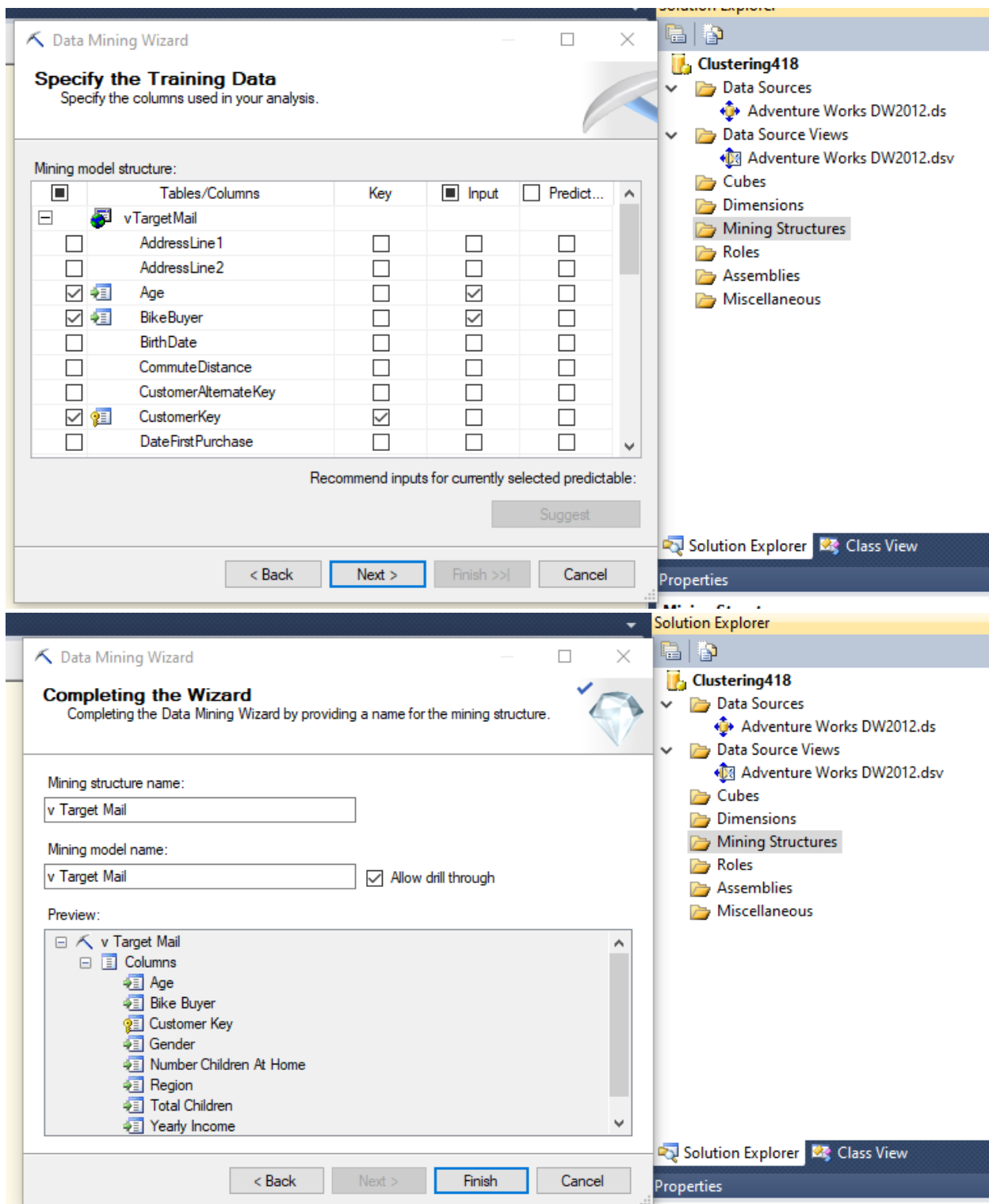
**Solution Explorer**

**Clustering418**

- Data Sources
  - Adventure Works DW2012.ds
- Data Source Views
  - Adventure Works DW2012.dsv
- Cubes
- Dimensions
- Mining Structures
- Roles
- Assemblies
- Miscellaneous

**Properties**

**Mining Structures**



**Process Progress**

Command

- Processing Database 'Clustering418' completed.
  - Start time: 1/6/2024 8:45:08 AM; End time: 1/6/2024 8:45:12 AM; Duration: 0:00:04
- Processing Mining Structure 'v Target Mail' completed.
  - Start time: 1/6/2024 8:45:08 AM; End time: 1/6/2024 8:45:12 AM; Duration: 0:00:04
- Processing Mining Model 'v Target Mail' completed.
- Processing Cube 'v Target Mail ~MC' completed.
  - Start time: 1/6/2024 8:45:09 AM; End time: 1/6/2024 8:45:10 AM; Duration: 0:00:00
- Processing Measure Group '~CaseDetail ~MG' completed.
- Processing Dimension 'v Target Mail ~MC-Customer Key' completed.

Status:

Process succeeded.

**Clustering418**

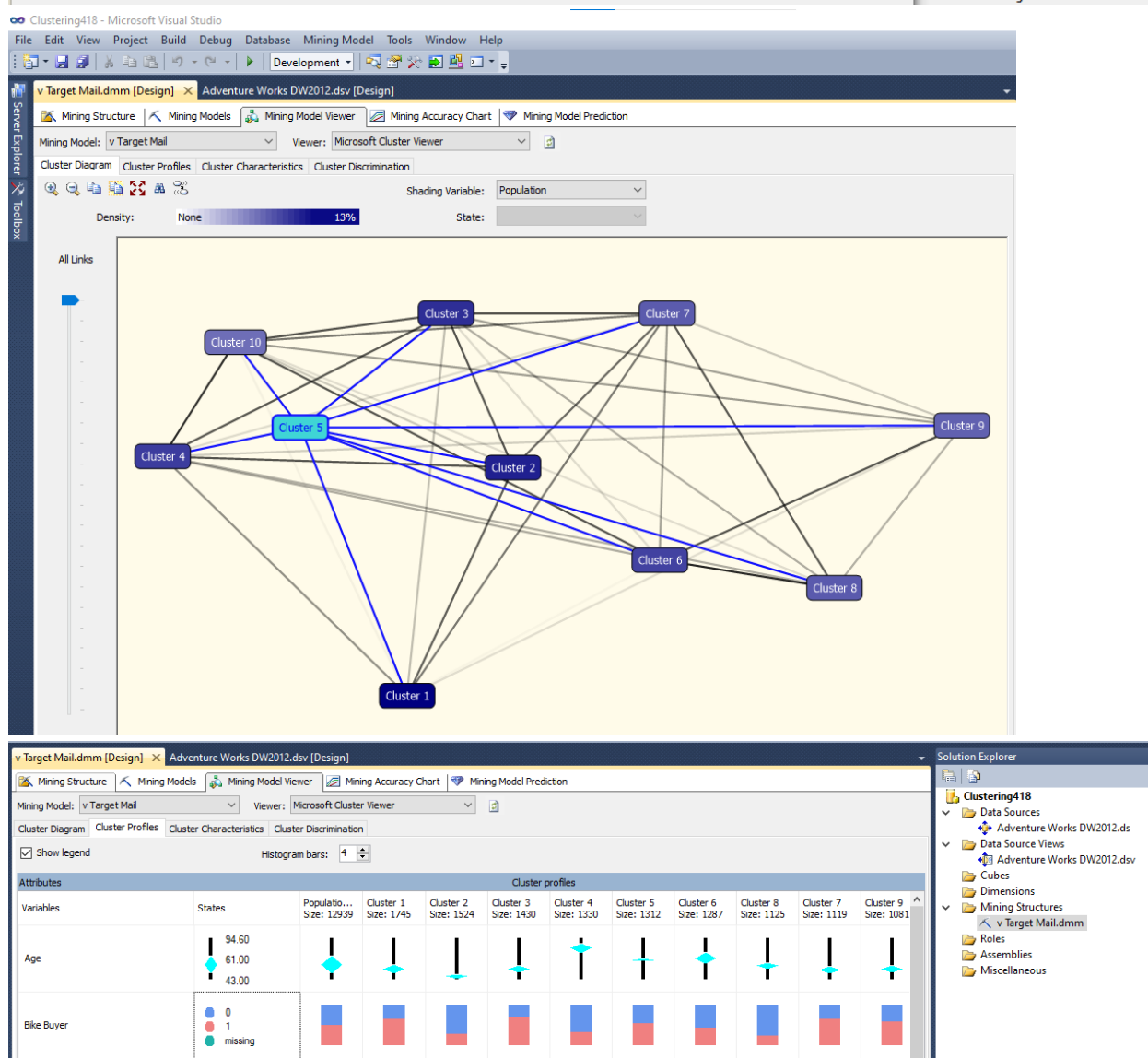
- Data Sources
  - Adventure Works DW2012.ds
- Data Source Views
  - Adventure Works DW2012.dsv
- Cubes
  - Dimensions
  - Mining Structures
    - v Target Mail.dmm
- Roles
- Assemblies
- Miscellaneous

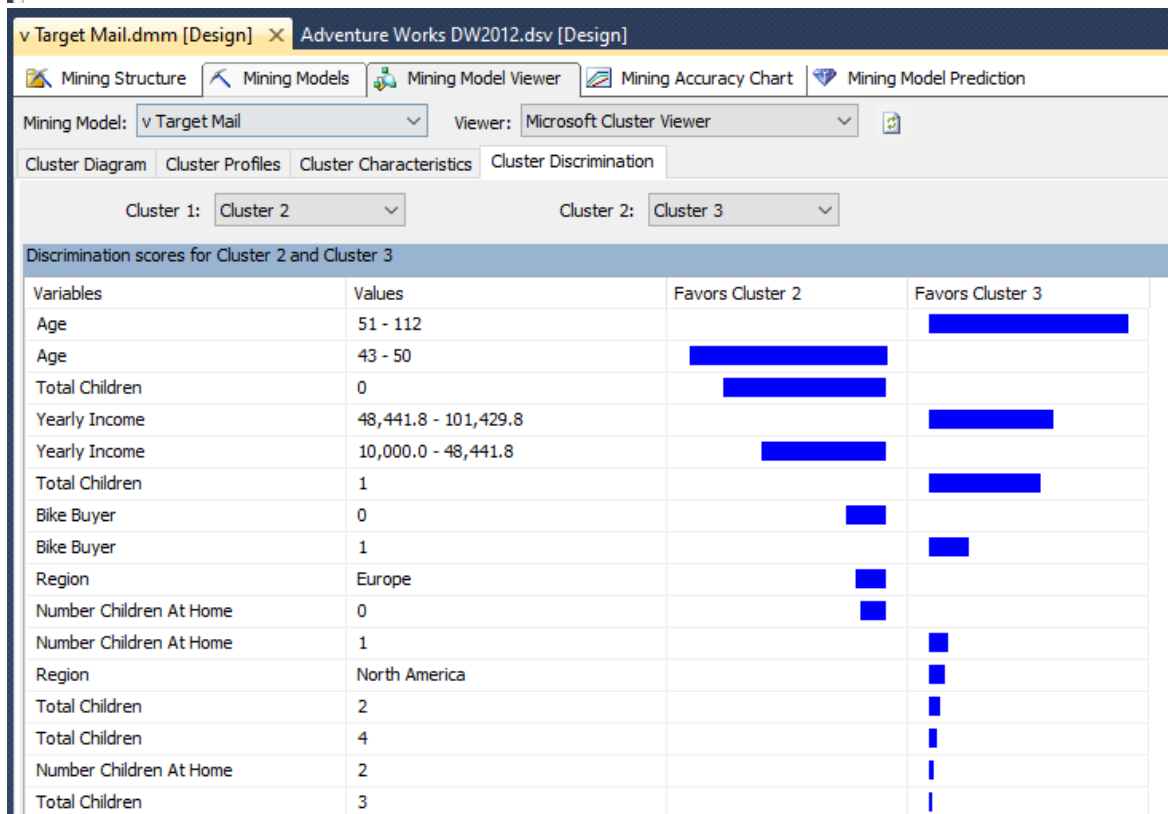
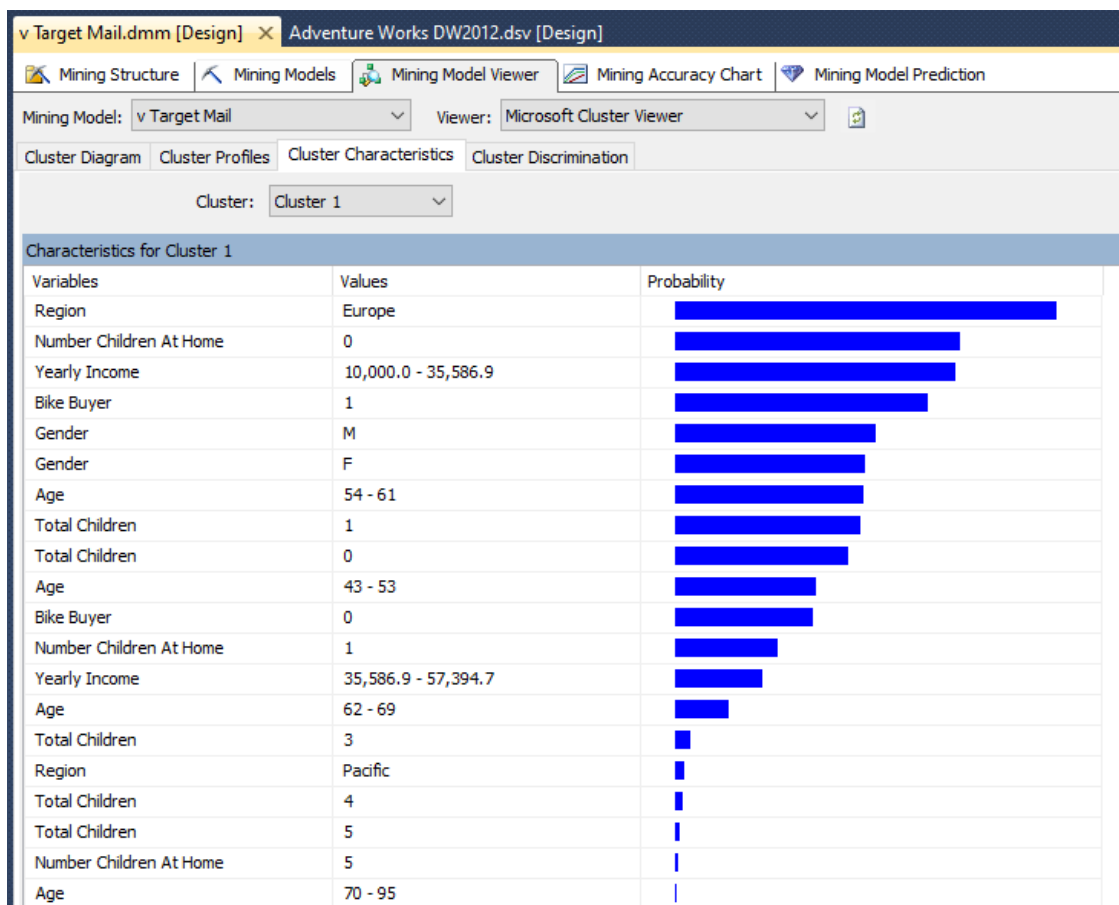
Solution Explorer

Deployment Progress - Clustering418

Server: localhost

Database: Clustering418

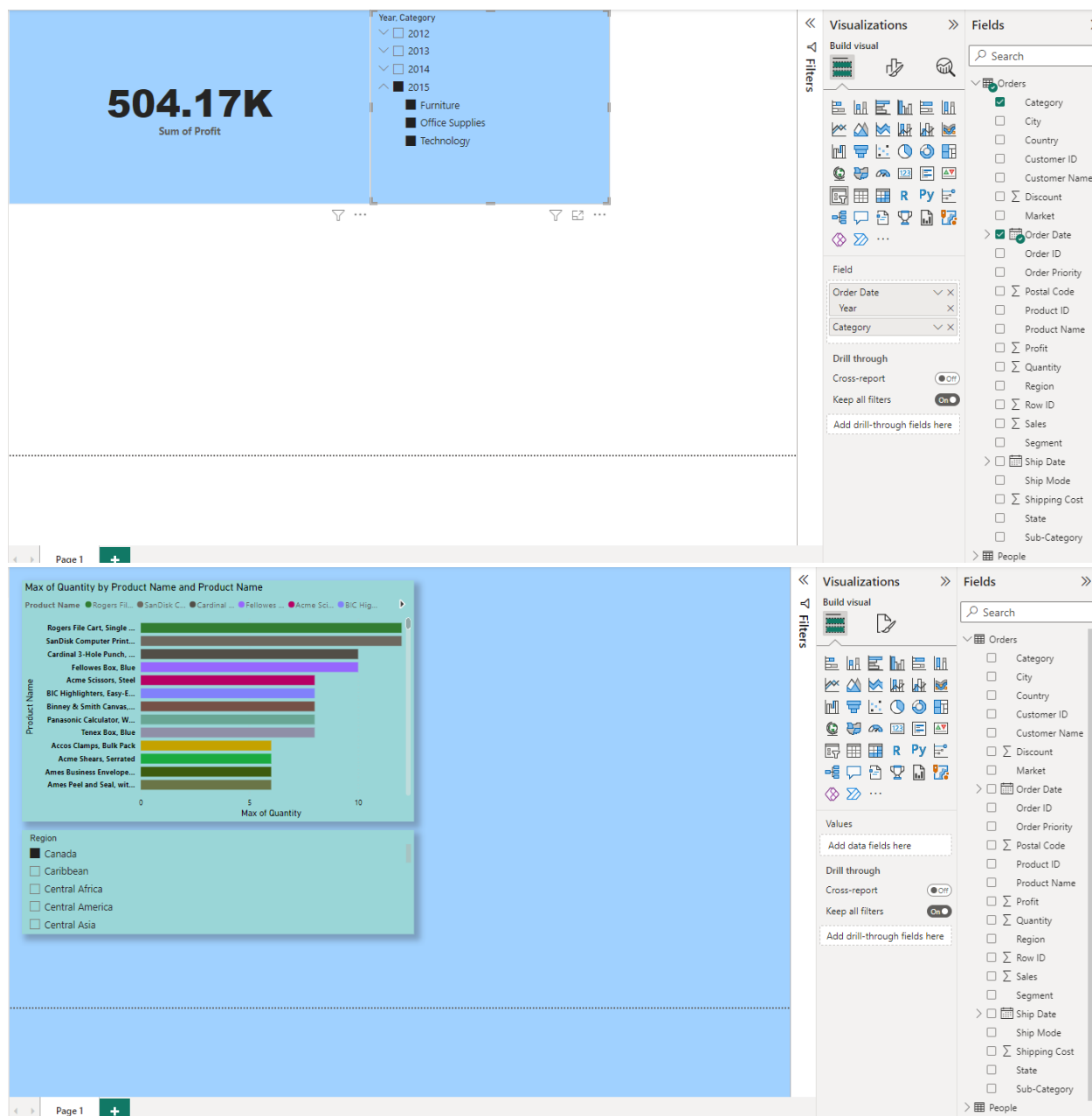






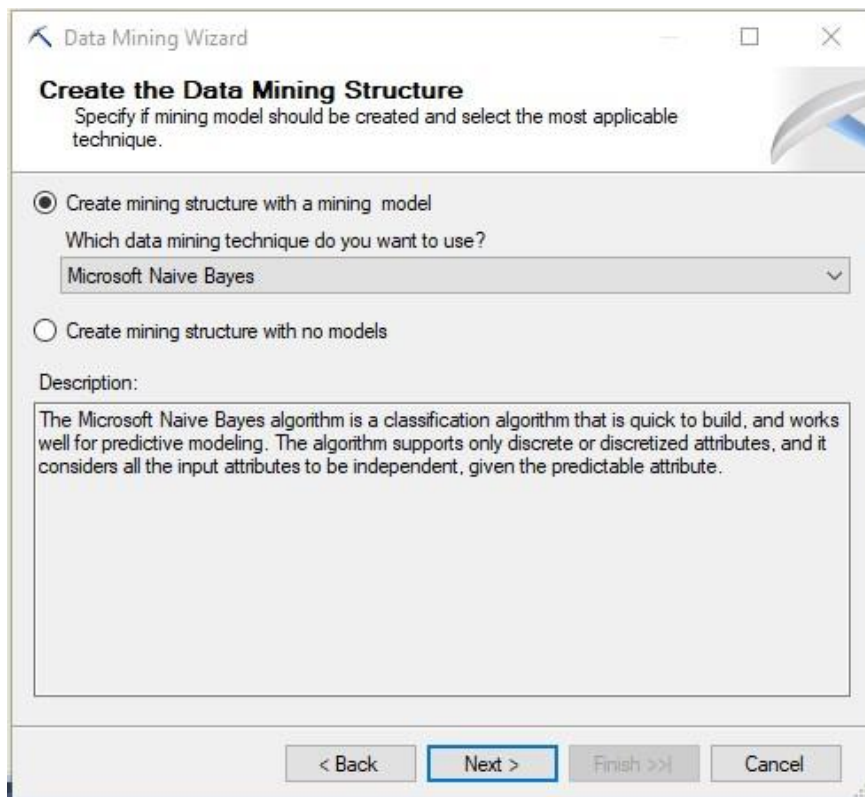
## Practical 09:

Data visualization using power BI (Use of visualization tools like Card, Donut, chart, stacked chart, tree map etc.)



## Practical 10:

### Perform the Naive Bayes Algorithm.



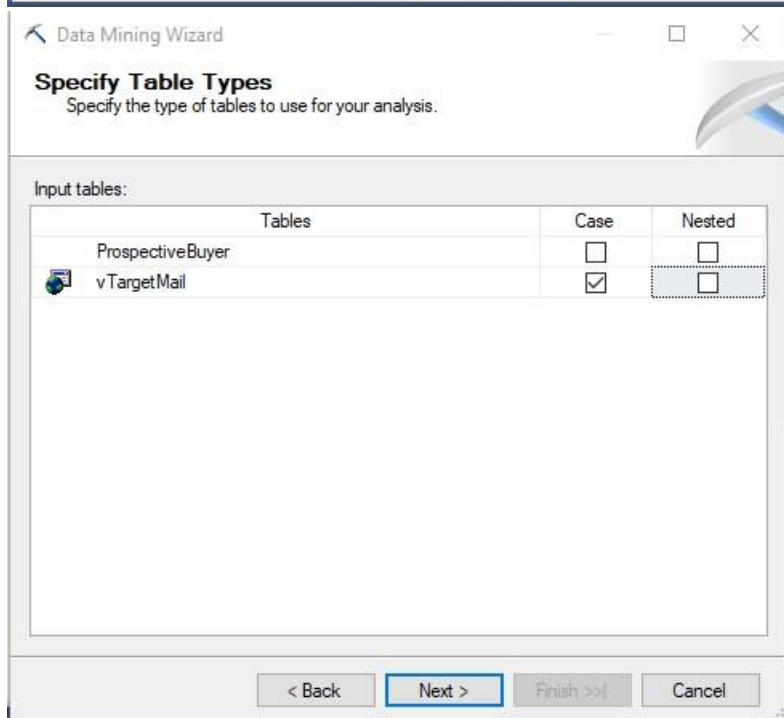
**Create the Data Mining Structure**  
Specify if mining model should be created and select the most applicable technique.

☒ Create mining structure with a mining model  
Which data mining technique do you want to use?  
Microsoft Naive Bayes

☐ Create mining structure with no models

Description:  
The Microsoft Naive Bayes algorithm is a classification algorithm that is quick to build, and works well for predictive modeling. The algorithm supports only discrete or discretized attributes, and it considers all the input attributes to be independent, given the predictable attribute.

< Back   Next >   Finish >>   Cancel



**Specify Table Types**  
Specify the type of tables to use for your analysis.

Input tables:

Tables	Case	Nested
ProspectiveBuyer	<input type="checkbox"/>	<input type="checkbox"/>
vTargetMail	<input checked="" type="checkbox"/>	<input type="checkbox"/>

< Back   Next >   Finish >>   Cancel

The screenshot displays the Microsoft SQL Server Data Mining (SSDM) interface. The top window shows the 'v Target Mail Mining.dmm [Design]' and 'Adventure Works DW2012.dsv [Design]' tabs. The 'Mining Structure' pane on the left lists various attributes: Age, Bike Buyer, Commute Distance, Customer Key, Gender, Number Cars Owned, Number Children At Home, Region, Total Children, and Yearly Income. The 'Mining Models' pane on the right shows a list of models, including 'Microsoft\_Ddecision\_Trees', 'Input', 'PredictOnly', and several 'Input' models.

A 'New Mining Model' dialog box is open in the foreground. It contains the following fields:

- Model name:** nb\_418
- Algorithm name:** Microsoft Naive Bayes

At the bottom of the dialog are three buttons: 'OK', 'Cancel', and 'Help'.

The bottom window shows the 'Mining Model: nb\_418' and 'Viewer: Microsoft Naive Bayes Viewer'. The 'Dependency Network' tab is selected, displaying a network diagram. The diagram shows a central node 'Bike Buyer' with five incoming arrows from nodes 'Region', 'Total Children', 'Commute Distance', 'Number Cars Owned', and 'Number Children At Home'. A legend at the bottom explains the network components:

- Strongest Links:**
  - Selected node (blue square)
  - This node predicts the selected node (orange square)
- Predicts both ways:**
  - Selected node predicts this node (purple square)

The 'Variables' pane at the bottom right is empty.

