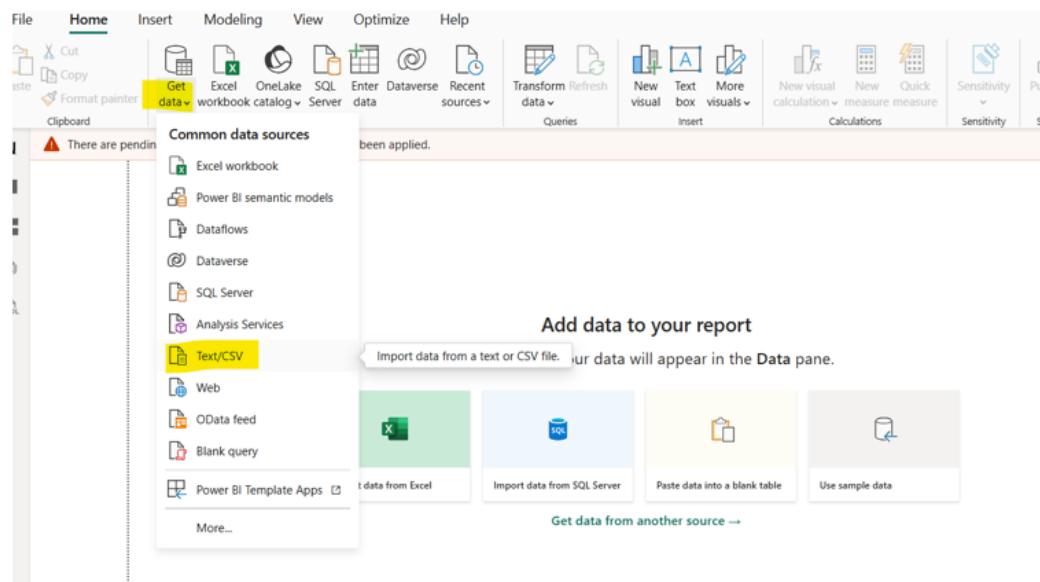


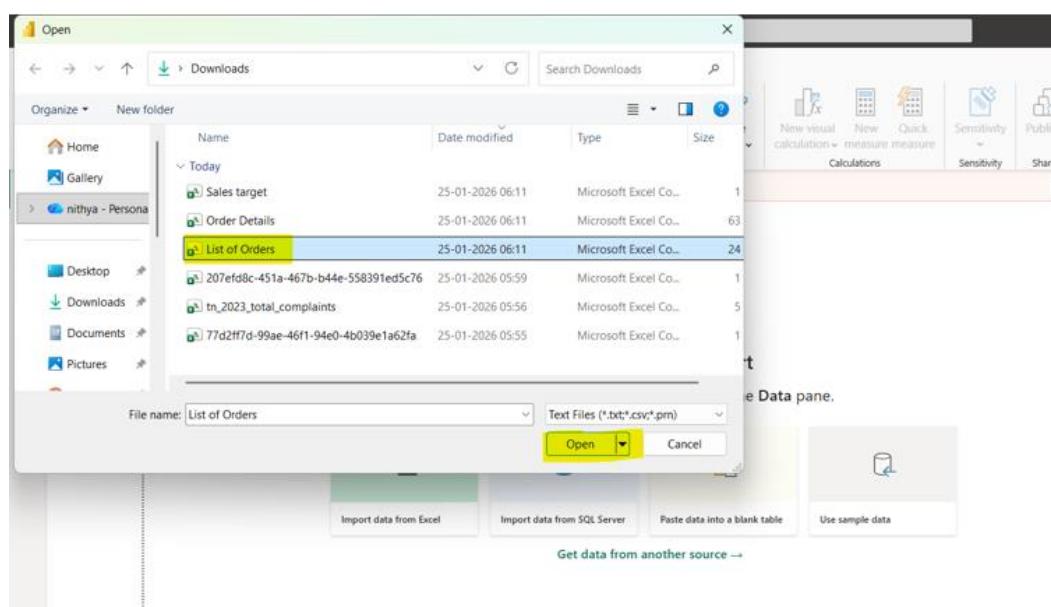
Power BI Assignment 1 – Data Transformation & Data Modelling

Import Data

- Import “List of Orders.csv” file into Power BI. In Home tab, click “Get Data” and select “Text/CSV”.



- Select the file and click “open”



- Open “List of Orders” in Power Query Editor by clicking on ‘Transform’.

The screenshot shows the Power Query Editor interface with the following details:

- File Origin:** 1252: Western European (Windows)
- Delimiter:** Comma
- Data Type Detection:** Based on first 200 rows
- Table Data:**

Order ID	Order Date	CustomerName	State	City
B-25601	01-04-2018	Bharat	Gujarat	Ahmedabad
B-25602	01-04-2018	Pearl	Maharashtra	Pune
B-25603	03-04-2018	Jahan	Madhya Pradesh	Bhopal
B-25604	03-04-2018	Divsha	Rajasthan	Jaipur
B-25605	05-04-2018	Kasheen	West Bengal	Kolkata
B-25606	06-04-2018	Hazel	Karnataka	Bangalore
B-25607	06-04-2018	Sonakshi	Jammu and Kashmir	Kashmir
B-25608	08-04-2018	Aarushi	Tamil Nadu	Chennai
B-25609	09-04-2018	Jitesh	Uttar Pradesh	Lucknow
B-25610	09-04-2018	Yogesh	Bihar	Patna
B-25611	11-04-2018	Anita	Kerala	Thiruvananthapuram
B-25612	12-04-2018	Shrichand	Punjab	Chandigarh
B-25613	12-04-2018	Mukesh	Haryana	Chandigarh
B-25614	13-04-2018	Vandana	Himachal Pradesh	Simla
B-25615	15-04-2018	Bhavna	Sikkim	Gangtok
B-25616	15-04-2018	Kanak	Goa	Goa
B-25617	17-04-2018	Sagar	Nagaland	Kohima
B-25618	18-04-2018	Manju	Andhra Pradesh	Hyderabad
B-25619	18-04-2018	Ramesh	Gujarat	Ahmedabad
B-25620	20-04-2018	Sarita	Maharashtra	Pune
- Note:** The data in the preview has been truncated due to size limits.
- Buttons:** Extract Table Using Examples, Load, Transform Data (highlighted), Cancel

- Table loaded into the Power Query Editor

The screenshot shows the Power Query Editor interface with the following details:

- File** | **Untitled - Power Query Editor**
- Queries [3]** (List of Orders, Order Details, Sales target)
- Transform** ribbon tab selected
- Table Data:**

Order ID	Order Date	CustomerName	State	City
B-25601	01-04-2018	Bharat	Gujarat	Ahmedabad
B-25602	01-04-2018	Pearl	Maharashtra	Pune
B-25603	03-04-2018	Jahan	Madhya Pradesh	Bhopal
B-25604	03-04-2018	Divsha	Rajasthan	Jaipur
B-25605	05-04-2018	Kasheen	West Bengal	Kolkata
B-25606	06-04-2018	Hazel	Karnataka	Bangalore
B-25607	06-04-2018	Sonakshi	Jammu and Kashmir	Kashmir
B-25608	08-04-2018	Aarushi	Tamil Nadu	Chennai
B-25609	09-04-2018	Jitesh	Uttar Pradesh	Lucknow
B-25610	09-04-2018	Yogesh	Bihar	Patna
B-25611	11-04-2018	Anita	Kerala	Thiruvananthapuram
B-25612	12-04-2018	Shrichand	Punjab	Chandigarh
B-25613	12-04-2018	Mukesh	Haryana	Chandigarh
B-25614	13-04-2018	Vandana	Himachal Pradesh	Simla
B-25615	15-04-2018	Bhavna	Sikkim	Gangtok
B-25616	15-04-2018	Kanak	Goa	Goa
B-25617	17-04-2018	Sagar	Nagaland	Kohima
B-25618	18-04-2018	Manju	Andhra Pradesh	Hyderabad
B-25619	18-04-2018	Ramesh	Gujarat	Ahmedabad
B-25620	20-04-2018	Sarita	Maharashtra	Pune

- Import “Order Details.csv” and “Sales target.csv” into Power Query Editor by clicking Home tab, then click “New Source”, select “Text/CSV”

The screenshot shows the Power Query Editor interface with the 'Home' tab selected. In the 'New Source' dropdown, the 'Text/CSV' option is being selected, indicated by a context menu. The menu shows the sub-option 'Import data from a text or CSV file.' highlighted.

- Select the file and click “open”.

The screenshot shows the Power Query Editor with the 'File' tab selected. An 'Open' dialog box is displayed, showing the file 'Sales target' selected in the 'File name:' field. The 'Downloads' folder is selected in the 'Organize' dropdown. The 'Text Files (*.txt;*.csv;*.prn)' file type is selected in the 'Text Files (*.txt;*.csv;*.prn)' dropdown.

- Open “Sales Target” in Power Query Editor by clicking on ‘OK’.

The screenshot shows the Power Query Editor interface with the following details:

- File Origin:** 1252: Western European (Windows)
- Delimiter:** Comma
- Data Type Detection:** Based on first 200 rows
- Table Preview:** Shows data from January 2018 to November 2018, with categories Furniture and Clothing and target values ranging from 10400 to 16000.
- Message:** A note states: "The data in the preview has been truncated due to size limits."
- Buttons:** OK (highlighted), Cancel, Extract Table Using Examples.

Transform Data

Restricts Row Count:

- "List of Orders" table restricts to only the first 500 rows to remove the null values rows.

The screenshot shows the Power Query Editor interface with the following details:

- Table:** List of Orders
- Context Menu:** Keep Top Rows...
- Dialog Box:** Keep Top Rows
 - Specify how many rows to keep.
 - Number of rows: 500
- Applied Steps:** Kept First Rows

Data Type Change:

- “Order Date” column in the “List of Orders” table was changed to data type ‘Date’.

The screenshot shows the Power BI Data Editor interface. The 'List of Orders' table is open. The 'Order Date' column is selected and highlighted in yellow. The 'Data Type' dropdown menu is open, showing various options: Decimal Number, Fixed decimal number, Whole Number, Percentage, Date/Time, Text, True/False, and Binary. The 'Fixed decimal number' option is also highlighted in yellow. The table data includes columns for Order ID, Customer Name, State, and City, with rows showing dates from April 1, 2018, to April 18, 2018, and locations like Bharat, Maharashtra, etc.

- “Amount” column in the “Order Details” table data type changed to ‘Fixed Decimal Number’.

The screenshot shows the Power BI Data Editor interface. The 'Order Details' table is open. The 'Amount' column is selected and highlighted in yellow. The 'Data Type' dropdown menu is open, showing various options: Decimal Number, Fixed decimal number, Whole Number, Percentage, Date/Time, Text, True/False, and Binary. The 'Fixed decimal number' option is also highlighted in yellow. The table data includes columns for Order ID, Profit, Quantity, Category, and Sub-Category, with rows showing amounts like 1,275.00, 66.00, etc., and categories like Furniture, Clothing, Electronics, etc.

- “Target” column in the “Sales Target” table data type changed to ‘Fixed Decimal Number’.

The screenshot shows the Power BI Data Editor interface. The 'Sales Target' table is open. The 'Target' column is selected and highlighted in yellow. The 'Data Type' dropdown menu is open, showing various options: Decimal Number, Fixed decimal number, Whole Number, Percentage, Date/Time, Text, True/False, and Binary. The 'Fixed decimal number' option is also highlighted in yellow. The table data includes columns for Month of Order Date, Category, and Target, with rows showing target values like 10,400.00, 10,500.00, etc., for Furniture and Clothing categories.

Text Formatting:

- "Customer Name" column Formatted into proper case, ensuring consistent capitalization for each word.

The screenshot shows the Power BI desktop interface with the 'Text' ribbon tab selected. A context menu is open over the 'CustomerName' column, with 'Format' selected. Under 'Format', 'Text' is chosen, and 'Proper, type text' is highlighted. The 'Applied Steps' pane shows the step 'Capitalized Each Word'.

Order ID	Order Date	CustomerName	State	City
B-25601	01-04-2018	Bharat	Rajasthan	Jaipur
B-25602	01-04-2018	Pearl	West Bengal	Kolkata
B-25603	03-04-2018	Jahan	Karnataka	Bangalore
B-25604	03-04-2018	Divsha	Jammu and Kashmir	Kashmir
B-25605	05-04-2018	Kasheen	Tamil Nadu	Chennai
B-25606	06-04-2018	Hazel	Uttar Pradesh	Lucknow
B-25607	06-04-2018	Sonakshi	Bihar	Patna
B-25608	08-04-2018	Aarushi	Kerala	Thiruvananthapuram
B-25609	09-04-2018	Jitesh	Punjab	Chandigarh
B-25610	09-04-2018	Yogesh	Haryana	Chandigarh
B-25611	11-04-2018	Anita	Himachal Pradesh	Simla
B-25612	12-04-2018	Shrichand	Sikkim	Gangtok
B-25613	12-04-2018	Mukesh	Goa	Goa
B-25614	13-04-2018	Vandana	Nagaland	Kohima
B-25615	15-04-2018	Bhavna		
B-25616	15-04-2018	Kanak		
B-25617	17-04-2018	Sagar		

Merge Columns:

- Merged the "State" and "City" columns to create a new column named "Location" in the format 'City, State'.

The screenshot shows the Power BI desktop interface with the 'Text' ribbon tab selected. A 'Merge Columns' dialog box is open, prompting to choose how to merge the selected columns. The 'Separator' dropdown is set to 'Comma', and the 'New column name (optional)' input field contains 'Location'. The 'OK' button is highlighted.

Order Date	CustomerName	State	City
01-04-2018	Bharat	Gujarat	Ahmedabad
01-04-2018	Pearl	Maharashtra	Pune

- New Column “Location” created in the form of “City, State”

The screenshot shows the Power Query Editor interface. On the left, there is a preview of a table with three columns: 'Order Date', 'CustomerName', and 'Location'. The 'Location' column contains values like 'Ahmedabad,Gujarat', 'Pune,Maharashtra', etc. Above the preview, a formula bar displays the M code: `= Table.CombineColumns(#"Capitalized Each Word", {"City", "State"}, Combiner.CombineTextByDelimiter(", ", true))`. To the right of the preview, the 'Properties' pane shows the 'Name' as 'List of Orders' and the 'Applied Steps' pane shows the steps taken: Source, Use First Row as Header, Kept First Rows, Changed Type, Capitalized Each Word, and Merged Columns.

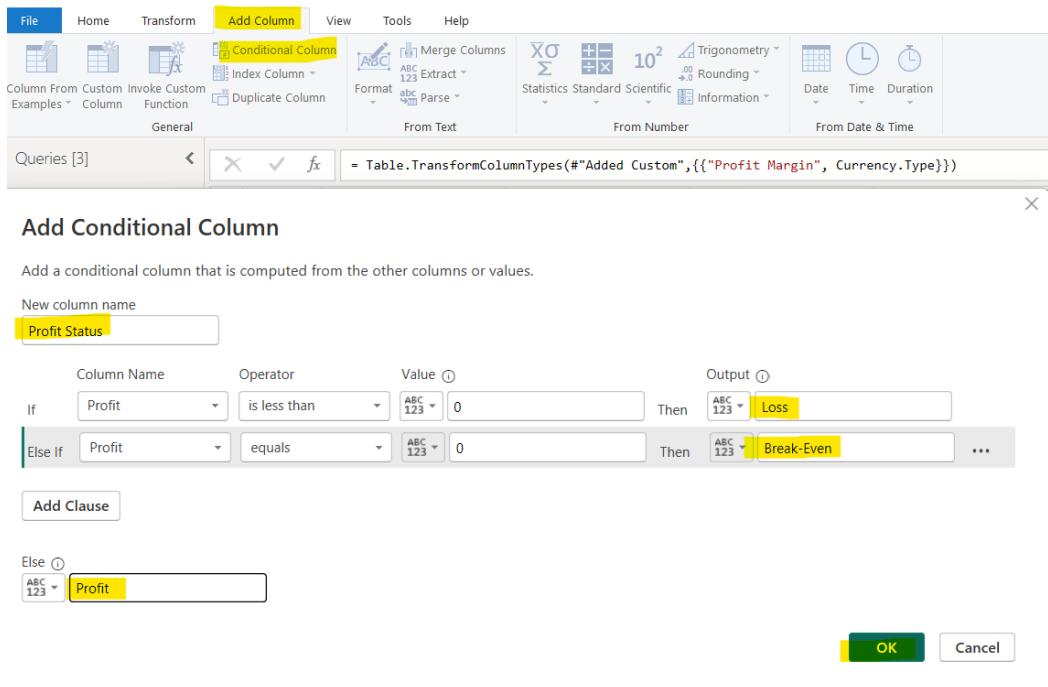
Create Custom Column:

- New custom column created named "Profit Margin" as the percentage of "Profit" divided by "Amount".

The screenshot shows the Power Query Editor with the 'Transform' tab selected. A 'Custom Column' dialog box is open over a table preview. The dialog box has a 'New column name' field containing 'Profit Margin' and a 'Custom column formula' field containing the formula `= [Profit]/[Amount]`. To the right of the formula is a list of 'Available columns' including 'Order ID', 'Amount', 'Profit', 'Quantity', 'Category', and 'Sub-Category'. At the bottom of the dialog box, a message says 'No syntax errors have been detected.' and there are 'OK' and 'Cancel' buttons. The background table preview shows columns for Order ID, Amount, Profit, Quantity, Category, and Sub-Category.

Create Conditional Column:

- New conditional column created named "Profit Status" based on the values in the "Profit" column. The conditions are as follows: if the profit is less than 0, the label should be "Loss"; if the profit equals 0, the label should be "Break-Even"; and if the profit is greater than 0, the label should be "Profit".



- Created new columns shown as follows,

```
= Table.AddColumn(#"Changed Type1", "Profit Status", each if [Profit] < 0 then "Loss" else if [Profit] = 0 then "Break-Even" else "Profit")
```

	Quantity	Category	Sub-Category	Profit Margin	Profit Status
1	-1,148.00	7 Furniture	Bookcases	-0.90	Loss
2	-12.00	5 Clothing	Stole	-0.18	Loss
3	-2.00	3 Clothing	Hankerchief	-0.25	Loss
4	-56.00	4 Electronics	Electronic Games	-0.70	Loss
5	-111.00	2 Electronics	Phones	-0.66	Loss
6	-272.00	5 Electronics	Phones	-0.64	Loss
7	1,151.00	4 Electronics	Phones	0.44	Profit
8	212.00	3 Clothing	Saree	0.38	Profit
9	-5.00	8 Clothing	Saree	-0.04	Loss
10	-60.00	5 Clothing	Trousers	-0.04	Loss
11	-30.00	1 Furniture	Chairs	-1.25	Loss
12	-166.00	3 Clothing	Saree	-0.86	Loss
13		2 Clothing	Trousers	0.03	Profit

PROPERTIES
Name
Order Details
All Properties

APPLIED STEPS
Source
Promoted Headers
Changed Type
Added Custom
Changed Type1
Added Conditional Column
Changed Type2

Merging Data (Joins):

- Merge the "List of Orders" and "Order Details" tables into a new single table named "Orders Data" based on the "Order ID" relationship, by clicking under "Home" tab, click the "Merge Queries" dropdown symbol, select "Merge Queries as New" and select the two table that needs to merge.

The screenshot shows the Microsoft Power BI desktop application. In the center, the 'Merge' dialog box is open, prompting the user to select tables and matching columns to create a merged table. Two tables are listed: 'Order Details' and 'List of Orders'. The 'Join Kind' is set to 'Left Outer (all from first, matching from second)'. The 'APPLIED STEPS' pane on the right indicates that 'Removed Duplicates' has been applied.

- After merging the tables, a new table "Orders Data" created and by expand the table, got the all column from secondary table.

The screenshot shows the Microsoft Power BI desktop application with the 'Orders Data' table displayed. The table includes columns: Order ID, Order Date, CustomerName, Location, Amount, Profit, and Quantity. The 'APPLIED STEPS' pane on the right highlights 'Expanded List of Orders' and 'Reordered Columns'.

Handling Missing Data & Duplicate Data:

- There are no missing values found in the data and determine a strategy to address them.
- There are no duplicate rows found.

Sorting and Filtering Data:

- In the ‘Orders Data’ table, Sort the orders by Order Date in descending order to analyse recent trends.

The screenshot shows the 'Orders Data' table in Power BI. The columns are Order ID, Order Date, CustomerName, Location, \$ Amount, Profit, and Quantity. A context menu is open over the 'Order Date' column, with 'Sort Descending' highlighted. The table data includes rows for various customers like Hitika, Bhishm, and Pinky across different locations and dates from August 31 to September 21, 2018. The 'APPLIED STEPS' pane on the right shows the step 'Sorted Rows'.

- In the ‘Orders Data’ table, Filter the orders to focus only on a specific state (e.g., Tamil Nadu) for regional analysis.

The screenshot shows the 'Orders Data' table in Power BI, filtered to show only orders from Tamil Nadu. The columns are Order Date, CustomerName, Location, \$ Amount, Profit, Quantity, and Category. A context menu is open over the 'Location' column, with 'Chennai,Tamil Nadu' selected. The table data shows various purchases from Chennai, such as clothing, electronics, and furniture items. The 'APPLIED STEPS' pane on the right shows the step 'Filtered Rows'.

Grouping and Aggregating Data:

- Duplicate the “Order Details” table and calculate the count of each Order ID, average profit by Category or total amount by Sub-Category.

Order Details table duplicated and named as “Order Details Aggregate”

The screenshot shows the Power BI Data View interface. On the left, the data source tree includes "List of Orders", "Order Details" (selected), "Sales target", "Orders Data", "Order Details Aggregate" (highlighted with a yellow box), and "Sales target Aggregate". In the main area, a table view shows columns "Amount" and "Order ID". A context menu is open over the table, with "Duplicate" highlighted in yellow. Other options in the menu include Copy, Paste, Delete, Rename, Enable load, Include in report refresh, Reference, Move To Group, Move Up, Move Down, Create Function..., Convert To Parameter, Advanced Editor, and Properties... .

Group By function to calculate the count of each Order ID

The screenshot shows the "Group By" dialog box. The left sidebar lists the same data sources as the previous screenshot. The main area has a title "Group By" and instructions "Specify the column to group by and the desired output." It includes a radio button for "Basic" mode, which is selected. Below this are fields for "New column name" (set to "Count of Order ID"), "Operation" (set to "Count Rows"), and "Column" (a dropdown menu). At the bottom right are "OK" and "Cancel" buttons. The bottom of the screen shows a preview of the data table with columns "Order ID", "Amount", and "Count of Order ID".

- Grouping by Order ID and calculated the Count of each Order ID.

	Order ID	Count of Order ID
1	B-25601	4
2	B-25602	5
3	B-25603	8
4	B-25604	2
5	B-25605	1
6	B-25606	1
7	B-25607	1
8	B-25608	4
9	B-25609	2
10	B-25610	6
11	B-25611	1
12	B-25612	1
13	B-25613	1
14	B-25614	2
15	B-25615	1
16	B-25616	4
17	B-25617	1
18	B-25618	2
19	B-25619	1
20	B-25620	1
21	B-25621	3
22	B-25622	1

PROPERTIES

Name: Order Details Aggregate
All Properties

APPLIED STEPS

- Source
- Promoted Headers
- Changed Type
- Added Custom
- Changed Type1
- Added Conditional Column
- Changed Type2
- Removed Duplicates
- Grouped Rows1**
- Grouped Rows
- Renamed Columns

- Grouping by Category and calculated the Average Profit.

	Category	1.2 Average Profit
1	Furniture	9.456790123
2	Clothing	11.76290832
3	Electronics	34.07142857

PROPERTIES

Name: Order Details Aggregate
All Properties

APPLIED STEPS

- Source
- Promoted Headers
- Changed Type
- Added Custom
- Changed Type1
- Added Conditional Column
- Changed Type2
- Removed Duplicates
- Grouped Rows1**
- Grouped Rows
- Renamed Columns

- Grouping by Subcategory and calculated the Total Amount.

	Sub-Category	1.2 Total Amount
1	Bookcases	56861
2	Stole	18546
3	Hankerchief	14608
4	Electronic Games	39168
5	Phones	46119
6	Saree	53511
7	Trousers	30039
8	Chairs	34222
9	Kurti	3361
10	T-shirt	7382
11	Shirt	7555
12	Leggings	2106
13	Tables	22614
14	Printers	58252
15	Accessories	21728
16	Furnishings	13484
17	Skirt	1946

PROPERTIES

Name: Order Details Aggregate
All Properties

APPLIED STEPS

- Source
- Promoted Headers
- Changed Type
- Added Custom
- Changed Type1
- Added Conditional Column
- Changed Type2
- Removed Duplicates
- Grouped Rows**
- Renamed Columns

- “Sales Target” table duplicated and named as “Sales target Aggregate” and aggregate the total target amount by Month of Order Date.

Sales target Aggregate

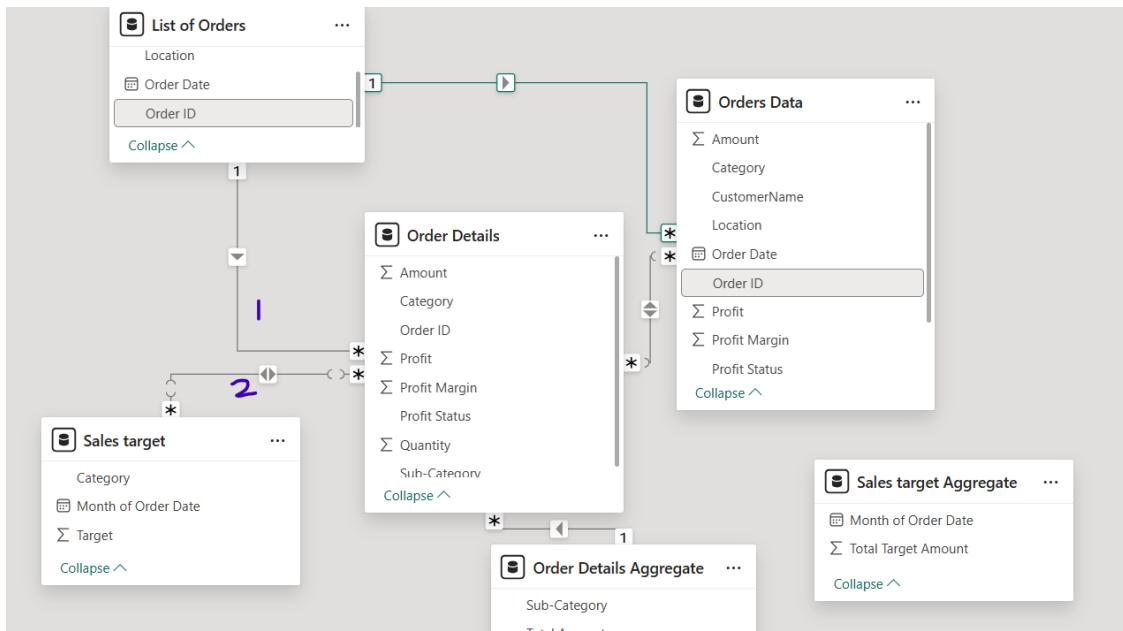
	Month of Order Date	Total Target Amount
1	01-04-2018	31400
2	01-05-2018	31500
3	01-06-2018	31600
4	01-07-2018	33800
5	01-08-2018	33900
6	01-09-2018	34000
7	01-10-2018	36100
8	01-11-2018	36300
9	01-12-2018	36400
10	01-01-2019	43500
11	01-02-2019	43600
12	01-03-2019	43800

APPLIED STEPS

- Source
- Promoted Headers
- Changed Type
- Grouped Rows**

Data Modelling

- Relationship established between the “List of Orders” and “Order Details” tables using the ‘Order ID’ column.
- Relationship created between the “Order Details” and “Sales Target” tables based on the ‘Category’ column.



- Click "Manage relationships" and ensure this relationship is active.

<input type="button"/> New relationship	Autodetect	Edit	Delete	Filter
From: table (column)	Relationship	To: table (column)	Status	
<input type="checkbox"/> Order Details (Order ID)		Orders Data (Order ID)	Active	
<input type="checkbox"/> Order Details (Order ID)		List of Orders (Order ID)	Active	
<input type="checkbox"/> Order Details (Sub-Category)		Order Details Aggregate (Sub-...)	Active	
<input type="checkbox"/> Orders Data (Order ID)		List of Orders (Order ID)	Active	
<input type="checkbox"/> Sales target (Category)		Order Details (Category)	Active	