

Power BI Assignment 1 – Data Transformation & Data Modeling

Data Transformation:

Restrict the "List of Orders" table to only the first 500 rows.

The screenshot shows the Power BI desktop interface with the 'List of Orders' query selected. In the 'Transform' tab, a context menu is open over the 'Order Date' column, specifically targeting the 'Keep Top Rows' option. The 'Keep Top Rows' dialog box is displayed, showing the value '500' entered. The 'APPLIED STEPS' pane on the right lists the step 'Kept First Rows'.

Ensure the "Order Date" column in the "List of Orders" table is set to data type 'Date'.

The screenshot shows the Power BI desktop interface with the 'List of Orders' query selected. In the 'Transform' tab, a context menu is open over the 'Order Date' column, with the 'Data Type' option selected. A dropdown menu shows various data types, with 'Date' highlighted. The 'APPLIED STEPS' pane on the right lists the step 'Changed order date type'.

Change the data type of "Amount" and "Target" columns to 'Fixed Decimal Number'

Amount :

The screenshot shows the Power BI desktop interface with the 'Order Details' query selected. In the 'Transform' tab, a context menu is open over the 'Amount' column, with the 'Data Type' option selected. A dropdown menu shows various data types, with 'Fixed decimal number' highlighted. The 'APPLIED STEPS' pane on the right lists the step 'Changed Amount Datatype'.

Target :

The screenshot shows the Power BI Data Editor interface. A table named "Sales target" is displayed with columns "Month of Order Date", "Category", and "Value". The "Category" column contains values like "Furniture" and "Target". The "Applied Steps" pane on the right shows the step "Changed Target Datatype" under the "APPLIED STEPS" section.

Format the "CustomerName" column into proper case, ensuring consistent capitalization for each word.

The screenshot shows the Power BI Data Editor interface. A table named "List of Orders" is displayed with columns "Order ID", "Order Date", and "CustomerName". The "CustomerName" column has been formatted using the "Capitalize Each Word" function. The "Applied Steps" pane on the right shows the step "Capitalized CustomerName" under the "APPLIED STEPS" section.

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

The screenshot shows the Power BI Data Editor interface. A table named "List of Orders" is displayed with columns "Order ID", "Order Date", "CustomerName", "State", and "City". A "Merge Columns" dialog box is open, prompting the user to choose how to merge the selected columns "State" and "City" into a new column "Location". The "Applied Steps" pane on the right shows the step "Capitalized CustomerName" under the "APPLIED STEPS" section.

Create a new custom column named "Profit Margin" as the percentage of "Profit" divided by "Amount".

The screenshot shows the Power Query Editor interface. In the ribbon, 'Add Column' is selected. A dialog box titled 'Custom Column' is open, prompting for a new column name ('Profit Margin') and a formula ('= [Profit]/[Amount]'). The formula bar shows '= [Profit]/[Amount]'. An 'Available columns' list includes 'Order ID', 'Amount', 'Profit', 'Quantity', 'Category', 'Sub-Category', and 'Profit Margin'. The right pane displays the 'Query Settings' and 'APPLIED STEPS' sections, which show the step 'Added Custom'.

Add a new conditional column 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".

The screenshot shows the Power Query Editor interface. In the ribbon, 'Conditional Column' is selected. A dialog box titled 'Add Conditional Column' is open, prompting for a new column name ('Profit Status'). It contains three clauses: 'If Profit is less than 0 Then "Loss"', 'Else If Profit equals 0 Then "Break-Even"', and 'Else "Profit"'. The right pane displays the 'Query Settings' and 'APPLIED STEPS' sections, which show the step 'Added Conditional Column'.

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.

The screenshot shows the Power Query Editor interface. In the ribbon, 'Transform' is selected. A dialog box titled 'Table.ExpandTableColumn' is open, showing the expansion of the 'Order Details' table into columns 'Amount', 'Profit', 'Quantity', 'Category', 'Sub-Category', 'Profit Margin', and 'Profit'. The right pane displays the 'Query Settings' and 'APPLIED STEPS' sections, which show the step 'Expanded Order Details'.

Handling Missing Data & Duplicate Data:

Identify missing values in the data and determine a strategy to address them. Check for duplicate rows and define a strategy to handle duplicates.

The screenshot shows the Power BI desktop interface with the 'Transform' tab selected. The 'Queries' list on the left shows four items: 'List of Orders', 'Order Details', 'Sales target', and 'Orders Data'. The 'Orders Data' query is currently selected, displaying a table with the following data:

Order ID	Order Date	CustomerName	Location	Amount	Profit	Quantity
B-25601	01-04-2018	Bharat	Ahmedabad, Gujarat	1,275.00	-1148	
B-25601	01-04-2018	Bharat	Ahmedabad, Gujarat	66.00	-12	
B-25601	01-04-2018	Bharat	Ahmedabad, Gujarat	8.00	-2	
B-25601	01-04-2018	Bharat	Ahmedabad, Gujarat	80.00	-56	
B-25602	01-04-2018	Pearl	Pune, Maharashtra	168.00	-111	
B-25602	01-04-2018	Pearl	Pune, Maharashtra	424.00	-272	
B-25602	01-04-2018	Pearl	Pune, Maharashtra	2,617.00	1151	
B-25602	01-04-2018	Pearl	Pune, Maharashtra	561.00	212	
B-25602	01-04-2018	Pearl	Pune, Maharashtra	119.00	-5	

The 'Applied Steps' pane on the right shows the following steps:

- Properties: Name = Orders Data
- Applied Steps:
 - Source: Expanded Order Details
 - Removed Duplicates

Sorting and Filtering Data:

In the 'Orders Data' table, utilize sorting and filtering techniques on columns like Order Date, State or Category to analyze data based on specific criteria:

◆ Sort the orders by Order Date in descending order to analyze recent trends.

The screenshot shows the Power BI desktop interface with the 'Transform' tab selected. The 'Queries' list on the left shows four items: 'List of Orders', 'Order Details', 'Sales target', and 'Orders Data'. The 'Orders Data' query is currently selected, displaying a table with the following data:

Order ID	Order Date	CustomerName	Location	Amount	Profit	Quantity
Hitika	Indore, Madhya Pradesh	72.00	16			
Hitika	Indore, Madhya Pradesh	828.00	230			
Hitika	Indore, Madhya Pradesh	34.00	10			
Bhishm	Mumbai, Maharashtra	207.00	37			
Bhishm	Mumbai, Maharashtra	2,366.00	552			
Bhishm	Mumbai, Maharashtra	9.00	3			
Bhishm	Mumbai, Maharashtra	835.00	267			
Pinky	Kashmir, Jammu and Kashmir	46.00	14			
Pinky	Kashmir, Jammu and Kashmir	497.00	179			

The 'Applied Steps' pane on the right shows the following steps:

- Properties: Name = Orders Data
- Applied Steps:
 - Source: Expanded Order Details
 - Removed Duplicates
 - Sorted Order Date in Descending Order

◆ Filter the orders to focus only on a specific state (e.g., Tamil Nadu) for regional analysis.

The screenshot shows the Power BI desktop interface with the 'Transform' tab selected. The 'Queries' list on the left shows four items: 'List of Orders', 'Order Details', 'Sales target', and 'Orders Data'. The 'Orders Data' query is currently selected, displaying a table with the following data:

Order ID	Order Date	CustomerName	Location	Amount	Profit	Quantity
B-26081	22-03-2019	Aarushi	Chennai, Tamil Nadu	79.00	33	
B-26081	22-03-2019	Aarushi	Chennai, Tamil Nadu	169.00	0	
B-26081	22-03-2019	Aarushi	Chennai, Tamil Nadu	359.00	-338	
B-26081	22-03-2019	Aarushi	Chennai, Tamil Nadu	93.00	-84	
B-26081	22-03-2019	Aarushi	Chennai, Tamil Nadu	24.00	11	
B-26081	22-03-2019	Aarushi	Chennai, Tamil Nadu	637.00	50	
B-26018	14-02-2019	Aarushi	Chennai, Tamil Nadu	61.00	8	
B-26018	14-02-2019	Aarushi	Chennai, Tamil Nadu	326.00	107	
B-26008	09-02-2019	Kalyani	Chennai, Tamil Nadu	22.00	4	
B-26008	09-02-2019	Kalyani	Chennai, Tamil Nadu	706.00	51	

The 'Applied Steps' pane on the right shows the following steps:

- Properties: Name = Orders Data
- Applied Steps:
 - Source: Expanded Order Details
 - Removed Duplicates
 - Sorted Order Date in Descending Order
 - 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.

Count of each Order ID

The screenshot shows the Microsoft Power BI interface in the 'Transform' tab. The ribbon menu includes Home, Transform (selected), Add Column, View, Tools, and Help. The 'Transform' tab has several icons: Group By, Use First Row as Headers, Transpose, Reverse Rows, Count Rows, Detect Data Type, Rename, Replace Values, Unpivot Columns, Fill, Move, Pivot Column, Convert to List, and Any Column. The 'Queries [5]' pane on the left lists 'List of Orders', 'Order Details', 'Sales target', 'Orders Data', and 'Order Details (2)', with 'Order Details (2)' selected. The main area displays a table with two columns: 'Order ID' and 'Order_Count'. The table contains 21 rows of data, with the last row (B-25621) highlighted in grey. The formula bar at the top shows the formula: = Table.RenameColumns(#"Grouped Rows",{{"Count", "Order_Count"}}).

Order ID	Order_Count
B-25601	4
B-25602	5
B-25603	8
B-25604	2
B-25605	1
B-25606	1
B-25607	1
B-25608	4
B-25609	2
B-25610	6
B-25611	1
B-25612	1
B-25613	1
B-25614	2
B-25615	1
B-25616	4
B-25617	1
B-25618	2
B-25619	1
B-25620	1
B-25621	3

Average profit by each category

Home **Transform** Add Column View Tools Help

Transpose Reverse Rows Detect Data Type Fill Move Unpivot Columns
Group Use First Row By Count Rows Rename Pivot Column Convert to List
Table Any Column

Merge Columns ABC Extract abc Parse
Split Column Format Text Column

Queries [5]

= Table.SelectRows(#"Grouped Rows", each true)

	Category	1.2 Average_Profit
1	Furniture	9.456790123
2	Clothing	11.76290832
3	Electronics	34.07142857

List of Orders Order Details Sales target Orders Data Order Details (2)

Total amount by sub-category

Home **Transform** Add Column View Tools Help

Transpose Reverse Rows Detect Data Type Fill Move Unpivot Columns
Group Use First Row By Count Rows Rename Pivot Column Convert to List
Table Any Column

Merge Columns ABC Extract abc Parse
Split Column Format Text Column

Statistics Standard Nur

Queries [5]

= Table.Group(#"Filtered Rows", {"Sub-Category"}, {"Total_Amount", each List.Sum([Amount]), type nullable number})

	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

List of Orders Order Details Sales target Orders Data Order Details (2)

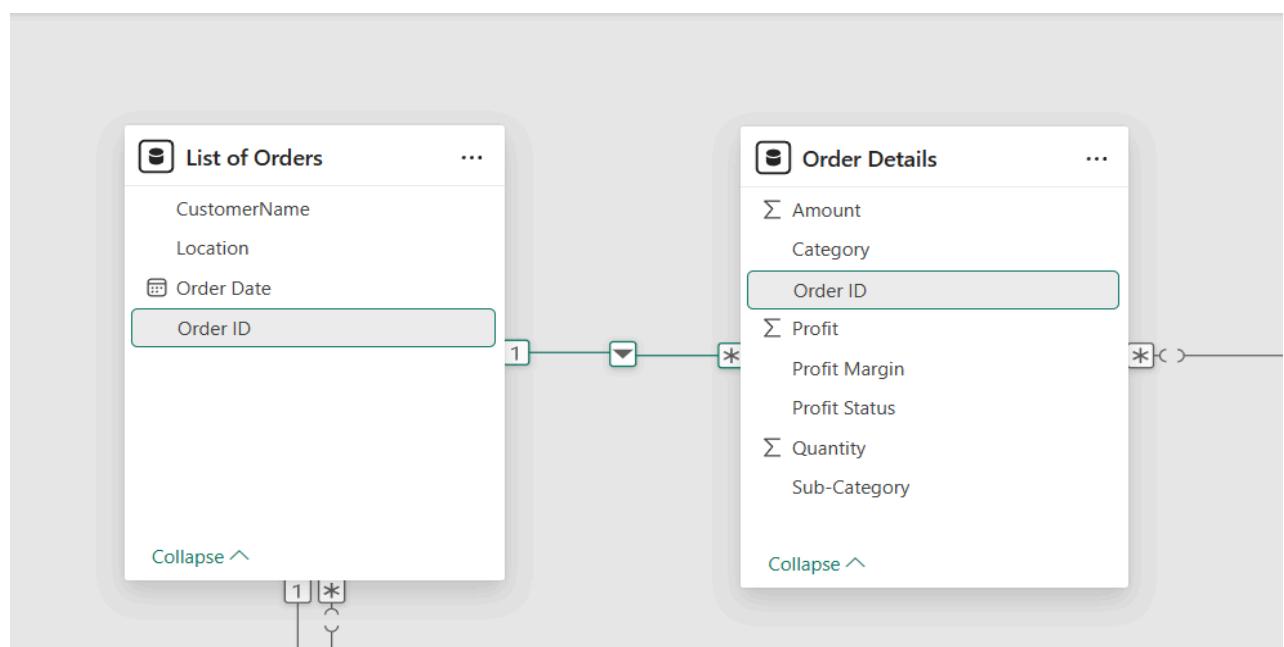
Duplicate the “Sales Target” table and aggregate the total target amount by Month of Order Date.

The screenshot shows the Microsoft Power BI Data Editor interface. The top navigation bar includes Home, Transform (which is selected), Add Column, View, Tools, and Help. The Transform ribbon has various tools like Transpose, Reverse Rows, Detect Data Type, and Pivot Column. Below the ribbon, the Queries list shows six items: List of Orders, Order Details, Sales target, Orders Data, Order Details (2), and Sales target (2). The 'Sales target (2)' query is currently selected. Its preview pane displays a table with two columns: 'Month of Order Date' and '1.2 Total_Target'. The data shows monthly target amounts from January 2018 to March 2019. The formula bar at the top of the preview pane shows the command: `= Table.RenameColumns(#"Grouped Rows",{{"Count", "Total_Target"}})`.

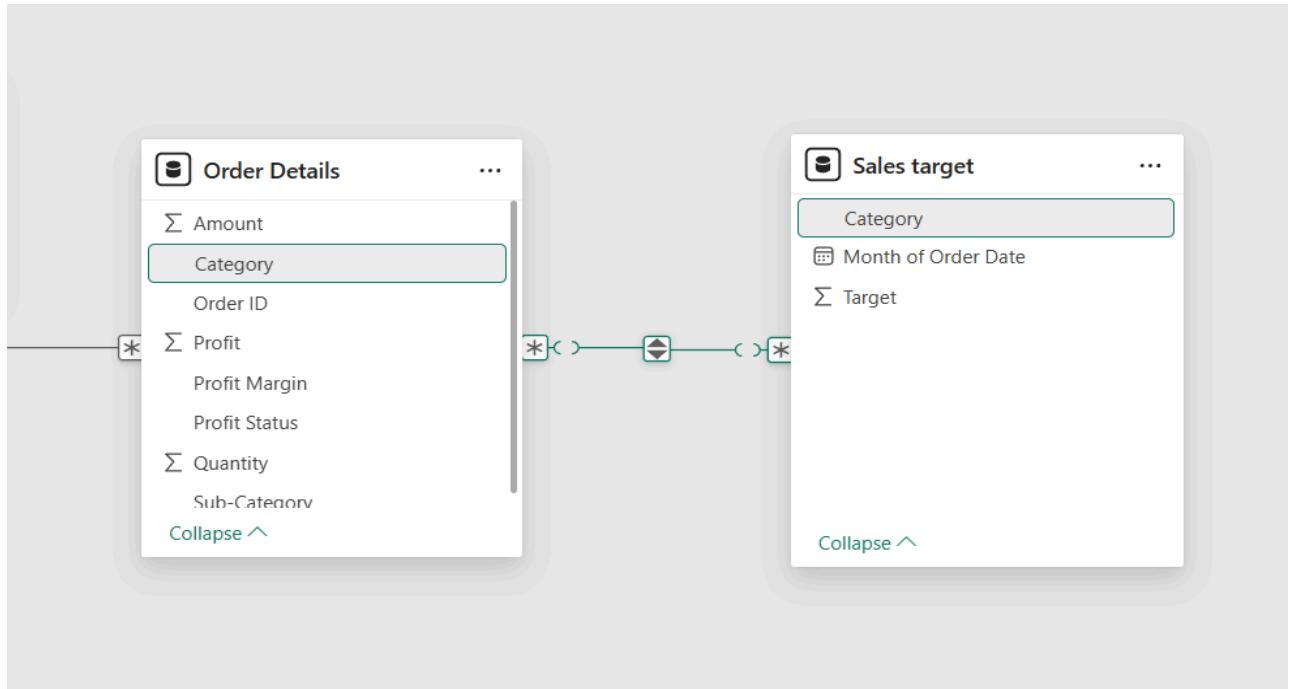
Month of Order Date	1.2 Total_Target
01-04-2018	31400
01-05-2018	31500
01-06-2018	31600
01-07-2018	33800
01-08-2018	33900
01-09-2018	34000
01-10-2018	36100
01-11-2018	36300
01-12-2018	36400
01-01-2019	43500
01-02-2019	43600
01-03-2019	43800

Data Modeling:

Establish a relationship between the “List of Orders” and “Order Details” tables using the ‘Order ID’ column.



Build a relationship between the “Order Details” and “Sales Target” tables based on the ‘Category’ column.



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

Manage relationships

+ New relationship		Autodetect	Edit	Delete	Filter
From: table (column)	To: table (column)	Relationship	Status		
Order Details (Category)	Sales target (Category)		Active	...	
Order Details (Order ID)	List of Orders (Order ID)		Active	...	
Orders Data (CustomerName)	List of Orders (CustomerName)		Active	...	
Orders Data (Order ID)	List of Orders (Order ID)		Active	...	

