

Power BI Assignment 1 – Data Transformation & Data Modeling

Import Data:

Import data “List of Orders, Order Details & Sales Target” in Power Query Editor.

Queries [3]

	Month of Order Date	Category	Target
1	01-04-2018	Furniture	10400
2	01-05-2018	Furniture	10500
3	01-06-2018	Furniture	10600
4	01-07-2018	Furniture	10800
5	01-08-2018	Furniture	10900
6	01-09-2018	Furniture	11000
7	01-10-2018	Furniture	11100
8	01-11-2018	Furniture	11300
9	01-12-2018	Furniture	11400
10	01-01-2019	Furniture	11500
11	01-02-2019	Furniture	11600
12	01-03-2019	Furniture	11800
13	01-04-2018	Clothing	12000
14	01-05-2018	Clothing	12000
15	01-06-2018	Clothing	12000
16	01-07-2018	Clothing	14000
17	01-08-2018	Clothing	14000
18	01-09-2018	Clothing	14000
19	01-10-2018	Clothing	16000
20	01-11-2018	Clothing	16000
21	01-12-2018	Clothing	16000

3 COLUMNS, 36 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 11:16

Data Transformation:

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

Queries [3]

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

5 COLUMNS, 500 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED AT 11:39

- Ensure the “Order Date” column in the “List of Orders” table is set to data type ‘Date’.

Screenshot of the Power Query Editor showing the “List of Orders” query. The “Order Date” column has been converted to a Date type, as indicated by the green status bar at the bottom.

Order ID	Order Date
5601	01-04-2018
5602	01-04-2018
5603	03-04-2018
5604	03-04-2018
5605	05-04-2018
5606	06-04-2018
5607	06-04-2018
5608	08-04-2018
5609	09-04-2018
5610	09-04-2018
5611	11-04-2018
5612	12-04-2018
5613	12-04-2018
5614	13-04-2018
5615	15-04-2018
5616	15-04-2018
5617	17-04-2018
5618	18-04-2018
5619	18-04-2018
5620	20-04-2018
5621	-

- Change the data type of “Amount” and “Target” columns to ‘Fixed Decimal Number’.

Editor

Screenshot of the Power Query Editor showing two tables being transformed. The left table shows the “Amount” column being converted to a Fixed decimal number type. The right table shows the “Target” column being converted to a Fixed decimal number type.

Order ID	Amount
B-25601	1,275.00
B-25601	66.00
B-25601	8.00
B-25601	80.00
B-25602	168.00

Month of Order Date	Category	Target
01-04-2018	Furniture	10,400.00
01-05-2018	Furniture	10,500.00
01-06-2018	Furniture	10,600.00
01-07-2018	Furniture	10,800.00
01-08-2018	Furniture	10,900.00
01-09-2018	Furniture	11,000.00

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

	CustomerName
3	Bharat
3	Pearl
3	Jahan
3	Divsha
3	Kasheen
3	Hazel
3	Sonakshi
3	Aarushi
3	Jitesh
3	Yogesh
3	Anita
3	Shrichand
3	Mukesh
3	Vandana
3	Bhavna
3	Kanak
3	Sagar
3	Manju
3	Ramesh
3	Sarita

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

```
,{"City", "State"},Combiner.Combine
```

ne	AB_C Location
	Ahmedabad,Gujarat
	Pune,Maharashtra
	Bhopal,Madhya Pradesh
	Jaipur,Rajasthan
	Kolkata,West Bengal
	Bangalore,Karnataka
	Kashmir,Jammu And Kashmir
	Chennai,Tamil Nadu
	Lucknow,Uttar Pradesh
	Patna,Bihar
	Thiruvananthapuram,Kerala
	Chandigarh,Punjab
	Chandigarh,Haryana

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

Properties

Name: Order Details

Applied Steps

- Source
- Promoted Headers
- Changed Type
- Added Custom- Profit Margin
- Changed Type - PM as Whole...
- Filtered Rows

- 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".

Properties

Name: Order Details

Applied Steps

- Source
- Promoted Headers
- Changed Type
- Added Custom- Profit Margin
- Changed Type - PM as Whole...
- Filtered Rows
- Added Profit Status Condition...
- Changed Type1

PREVIEW DOWNLOADED AT 12:29

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 BI Data Editor interface. On the left, under 'Queries [4]', there are four tables: 'List of Orders', 'Order Details', 'Sales target', and 'Orders Data'. The 'Orders Data' table is selected. In the center, the table structure is displayed with columns: Location, Amount, Profit, Quantity, and Category. The formula bar at the top shows the query: = Table.ExpandTableColumn(Source, "Order Details", {"Amount", "Profit", "Quantity", "Category"}, true). On the right, the 'Query Settings' pane is open, showing 'Name' set to 'Orders Data' and the 'APPLIED STEPS' section showing the step 'Expanded Order Details Colu...'. The data preview shows 7 rows of merged data.

Handling Missing Data & Duplicate Data:

- Identify missing values in the data and determine a strategy to address them.

Answer: No missing data in the first 500rows.

- Check for duplicate rows and define a strategy to handle duplicates.

Answer: Duplicate rows check with “Order ID” in List of Orders and Order Details tables.

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 Data Editor interface. The 'Orders Data' table is selected. The formula bar at the top shows the query: = Table.Sort(#"Filtered Rows",{{"Order Date", Order.Descending}}). The table structure includes columns: Order ID, Order Date, CustomerName, and Location. The data preview shows 7 rows of sorted data, with the most recent order date (31-03-2019) at the top.

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

= Table.SelectRows(#"Sorted Rows", each ([Location] = "Chennai,Tamil Nadu"))

	A ^B _C Order ID	Order Date	A ^B _C CustomerName	A ^B _C Location
1	B-26081	22-03-2019	Aarushi	Chennai,Tamil Nadu
2	B-26018	14-02-2019	Aarushi	Chennai,Tamil Nadu
3	B-26008	09-02-2019	Kalyani	Chennai,Tamil Nadu
4	B-25860	15-11-2018	Akshay	Chennai,Tamil Nadu
5	B-25788	21-09-2018	Dinesh	Chennai,Tamil Nadu
6	B-25716	11-07-2018	Surabhi	Chennai,Tamil Nadu
7	B-25698	23-06-2018	Amisha	Chennai,Tamil Nadu
8	B-25608	08-04-2018	Aarushi	Chennai,Tamil Nadu

Grouping and Aggregating Data:

- Duplicate the “Order Details” table and calculate the count of each Order ID

= Table.Group(#"Removed Duplicates", {

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

Average profit by Category Wise

1.2 Average Profit by Category

	A ^B _C Category	1.2 Average Profit by Category
1	Furniture	20.44318182
2	Electronics	-19.37894737
3	Clothing	15.29022082

Total amount by Sub-Category Wise.

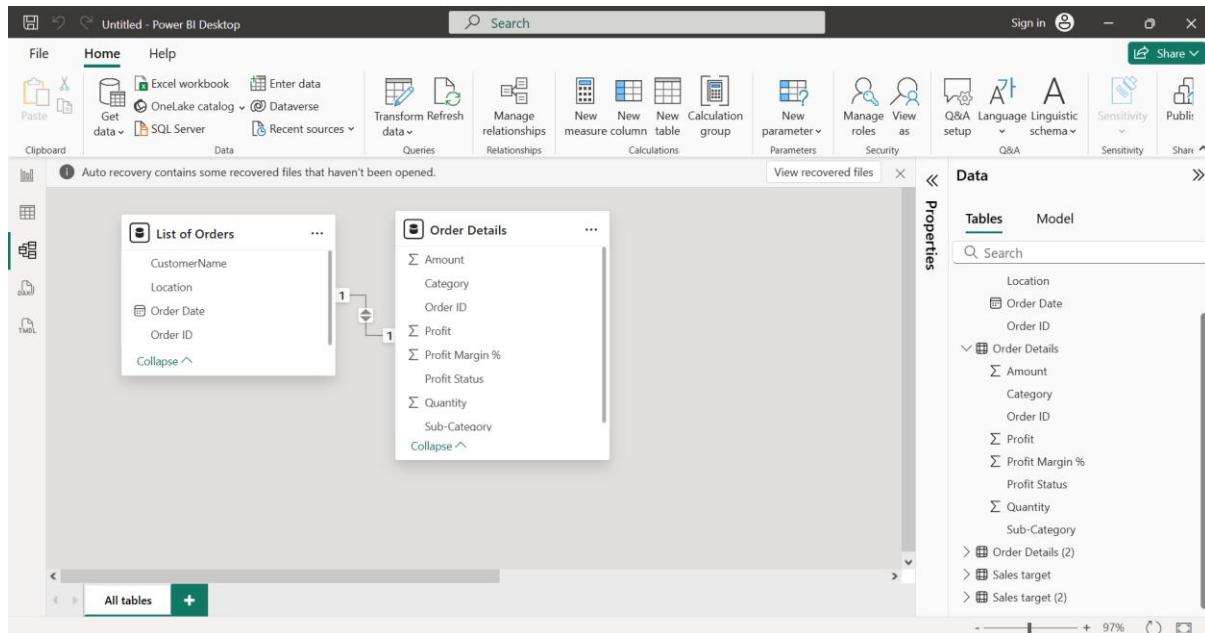
	A ^B C Sub-Category	1.2 Total amount by Sub-category
1	Bookcases	25800
2	Phones	11691
3	Trousers	7109
4	T-shirt	3205
5	Saree	17831
6	Shirt	2535
7	Leggings	851
8	Tables	3967
9	Electronic Games	11942
10	Printers	16822
11	Chairs	5724
12	Hankerchief	4362
13	Stole	7583
14	Skirt	822
15	Furnishings	4295
16	Accessories	9472
17	Kurti	1154

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

	Month of Order Date	1.2 Total Target by Month
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

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.

