Power BI Assignment 1 – Data Transformation & Data Modeling

Student Name: Banupriya N

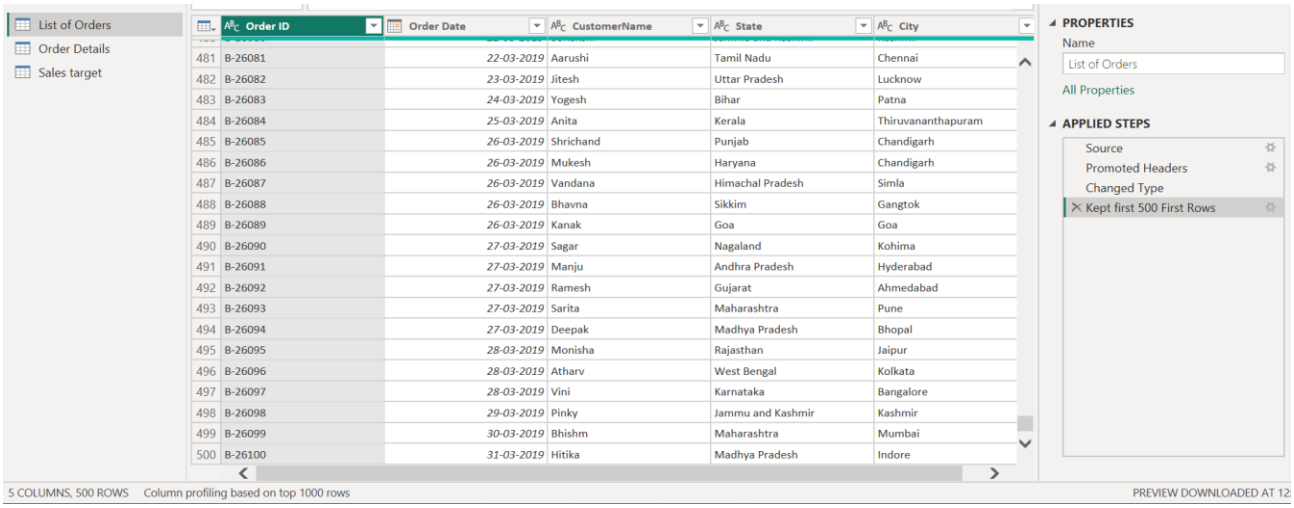
Topic Name: E-Commerce Sales Analysis

Import Data:

List of Orders Table imported to Power Bi through Ribbon→Home Tab→Data Group→Get Data→Excel Work Book and transformed to power query editor using Transform Data in Power BI and other two tables are imported from power query editor through Ribbon →Home tab→New Query Group→ New Source→Excel Workbook.

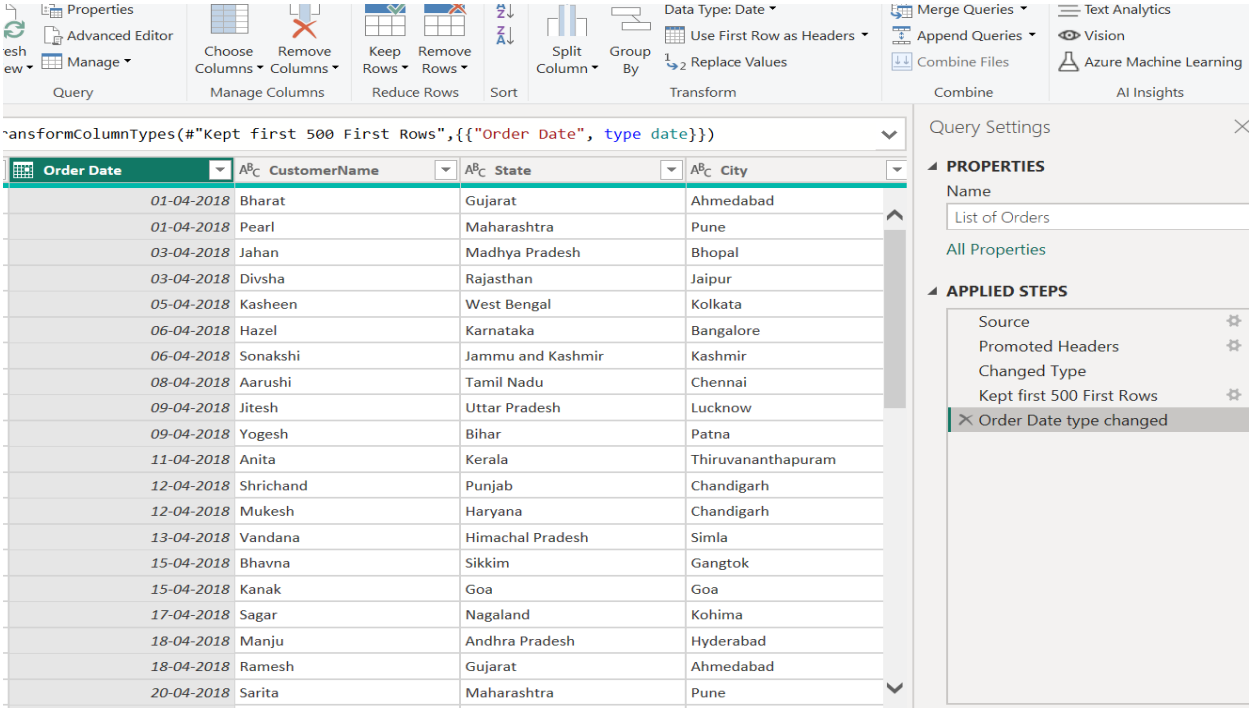
Data Transformation:

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



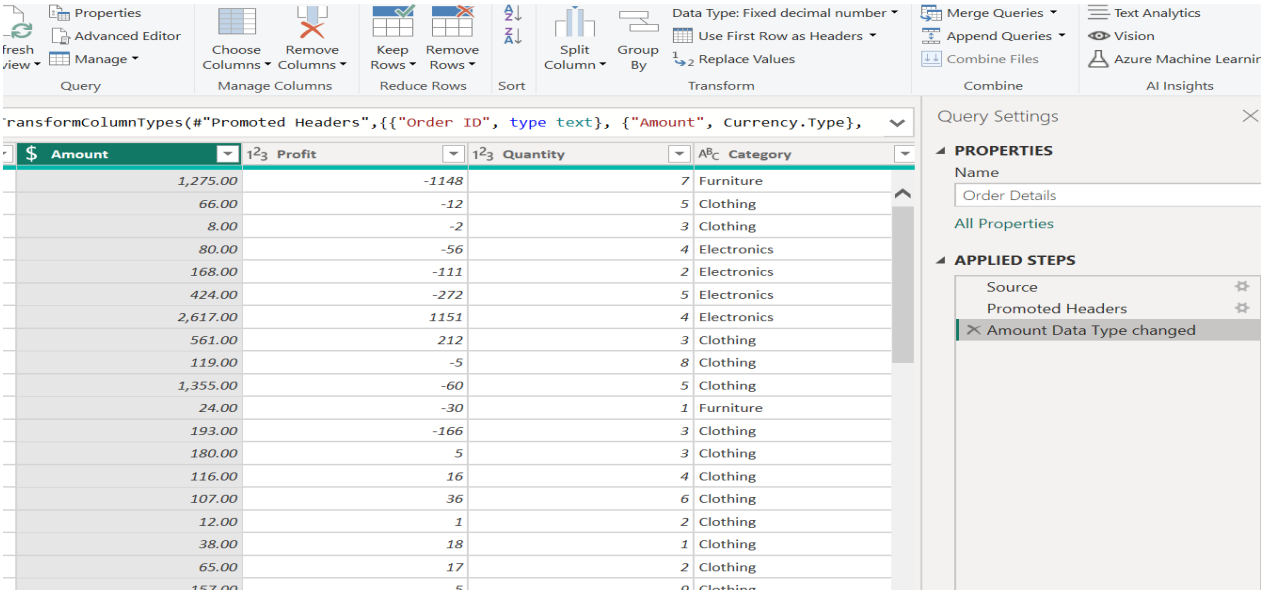
Restricted the first 500 rows in “List of Orders” table using keep rows menu. Ribbon→Home Tab→Reduce rows→Keep rows→Keep Top Rows→Number of Rows “500”.

Ensure the “Order Date” column in the “List of Orders” table is set to data type 'Date':

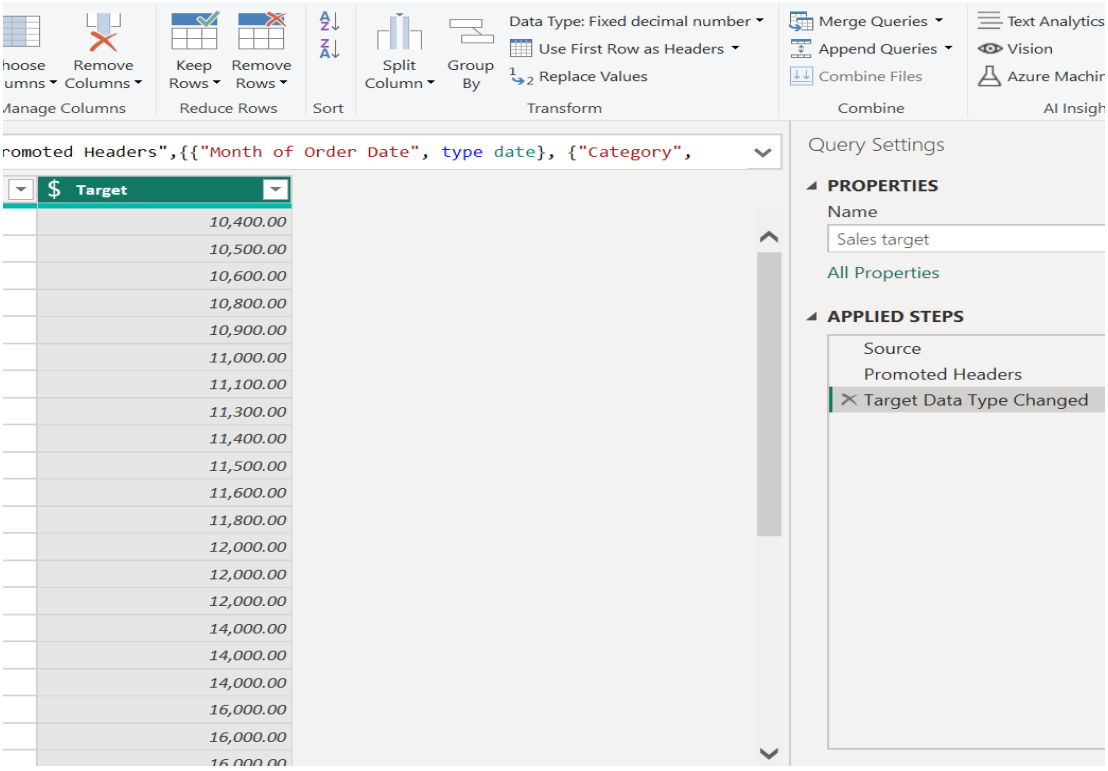


Order date data type changed as “Date”.

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



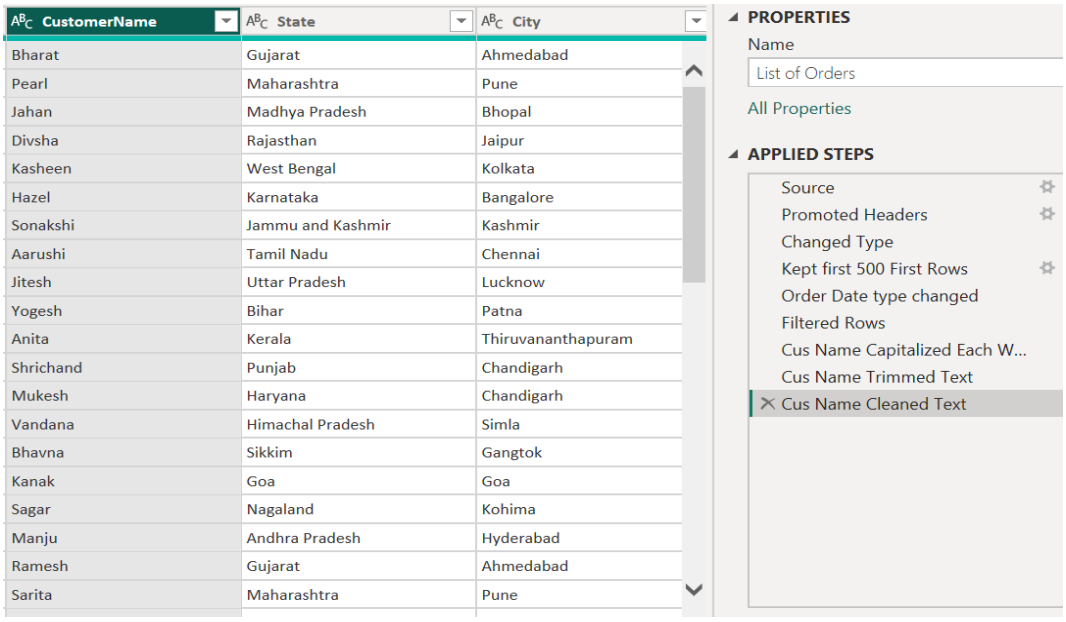
Column “Amount” Data type changed as “Fixed Decimal Number” in Order Details table



Column “Target” Data type changed as “Fixed Decimal Number” in Sales Target table

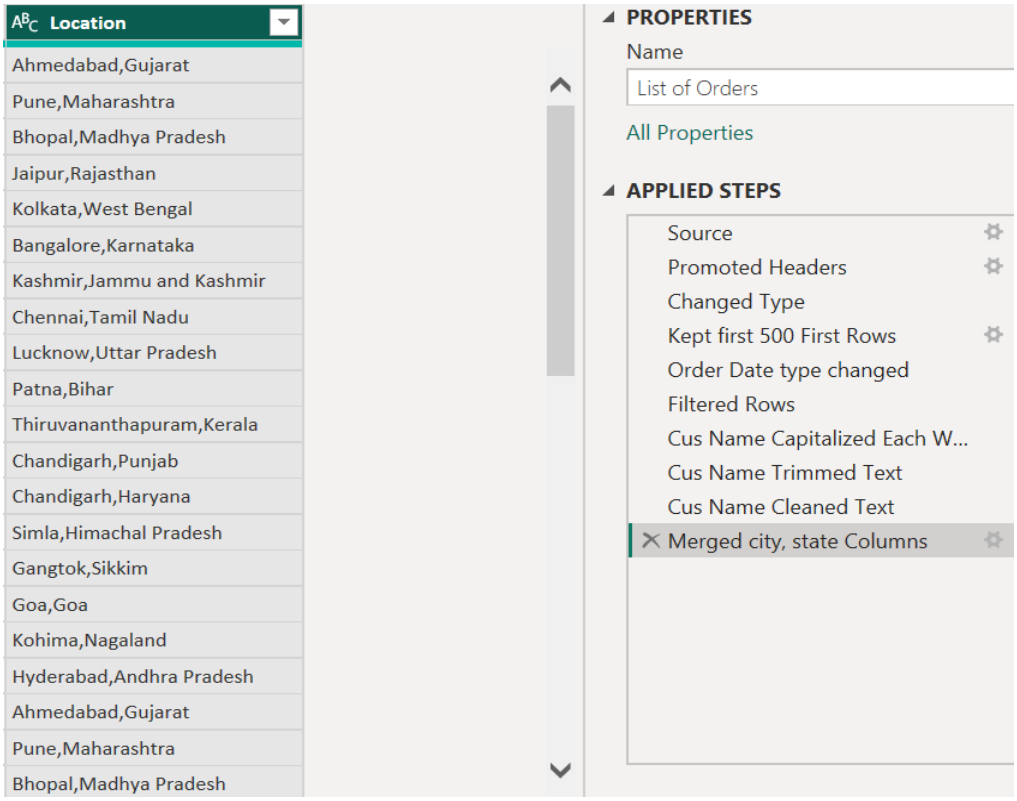
Format the "Customer Name" column into proper case, ensuring consistent capitalization for each word:

Customer Name column capitalized each word through Ribbon→Transform tab→Text Column group→Format menu→ Capitalize each word option.



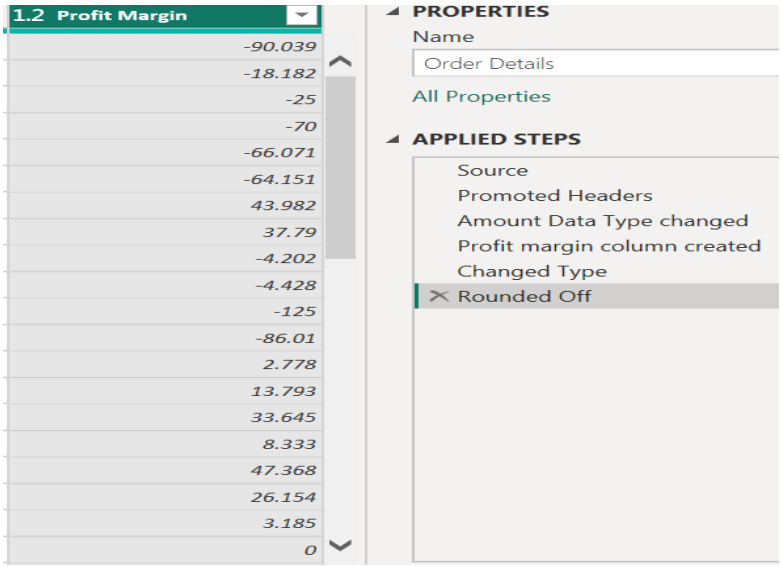
CustomerName	State	City
Bharat	Gujarat	Ahmedabad
Pearl	Maharashtra	Pune
Jahan	Madhya Pradesh	Bhopal
Divsha	Rajasthan	Jaipur
Kasheen	West Bengal	Kolkata
Hazel	Karnataka	Bangalore
Sonakshi	Jammu and Kashmir	Kashmir
Aarushi	Tamil Nadu	Chennai
Jitesh	Uttar Pradesh	Lucknow
Yogesh	Bihar	Patna
Anita	Kerala	Thiruvananthapuram
Shrichand	Punjab	Chandigarh
Mukesh	Haryana	Chandigarh
Vandana	Himachal Pradesh	Simla
Bhavna	Sikkim	Gangtok
Kanak	Goa	Goa
Sagar	Nagaland	Kohima
Manju	Andhra Pradesh	Hyderabad
Ramesh	Gujarat	Ahmedabad
Sarita	Maharashtra	Pune

Merge the "State" and "City" columns to create a new column named "Location" in the format ‘City, State’:



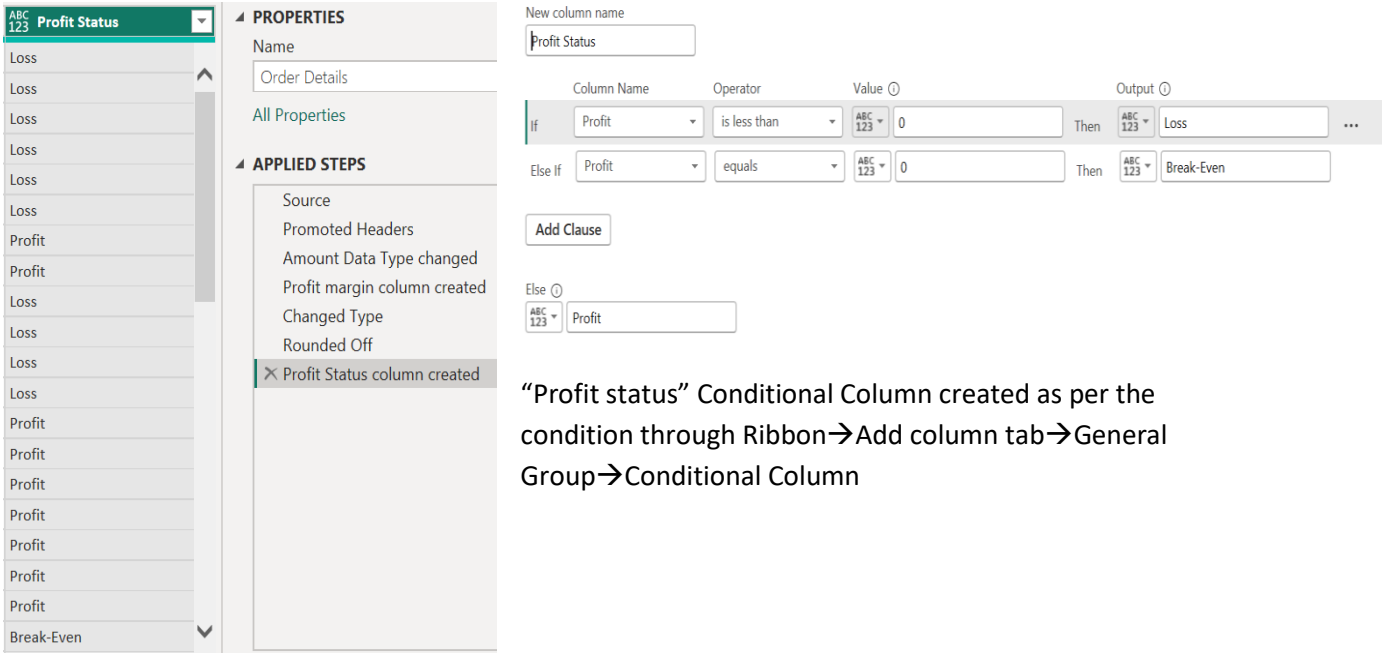
State and city column merged in one column named Location in the format City, State using Merge Columns in Transform Tab.

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



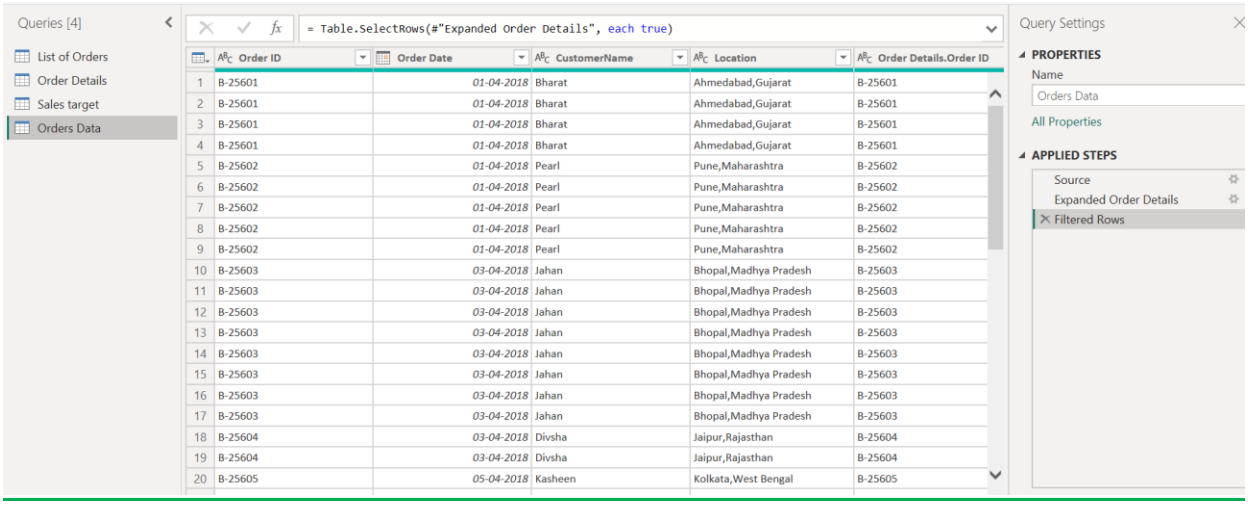
Profit Margin Customer column created as the percentage of profit divided by amount through Ribbon→Add column tab→General Group→Custom Column menu using formula **=([Profit] / [Amount]) * 100**

Add a new conditional column named "Profit Status" based on the values in the "Profit" column:



"Profit status" Conditional Column created as per the condition through Ribbon→Add column tab→General Group→Conditional Column

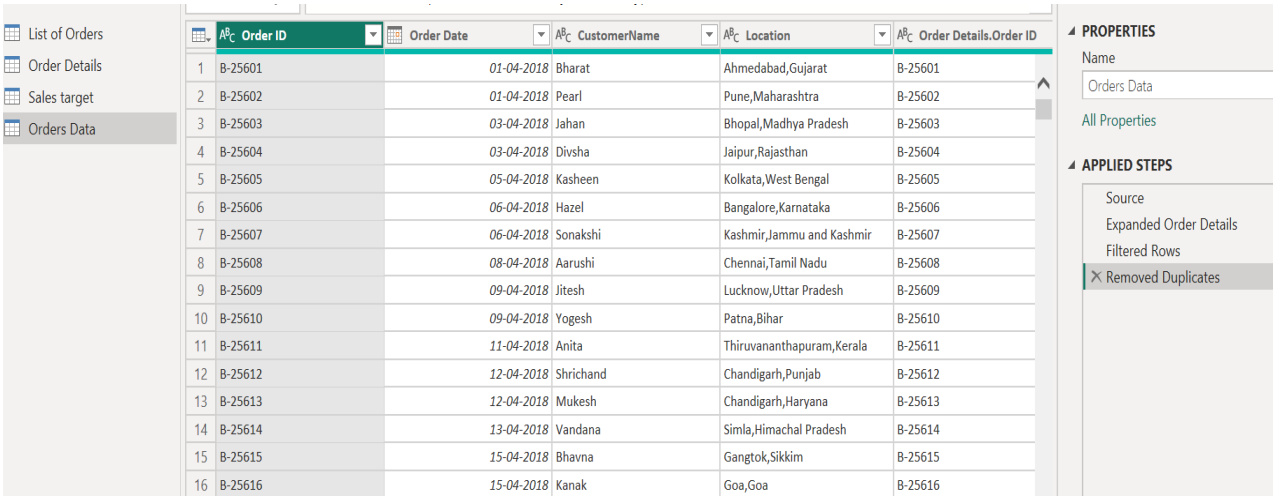
Merging Data (Joins):



New Single Table created named "Orders Data" by merging "List of Orders" and "Order Details" Table Through Ribbon→Home Tab→Combine Group→Merge Queries→ Merge Queries as New→Join Kind→Full Outer (All rows from both) based on the relationship "ORDER ID".

Handling Missing Data & Duplicate Data:

Duplicate rows are removed from order data table based on Order Id as unique by selecting order id column→ right clicking the order id column→Select Remove duplicates



List of Orders
Order Details
Sales target
Orders Data

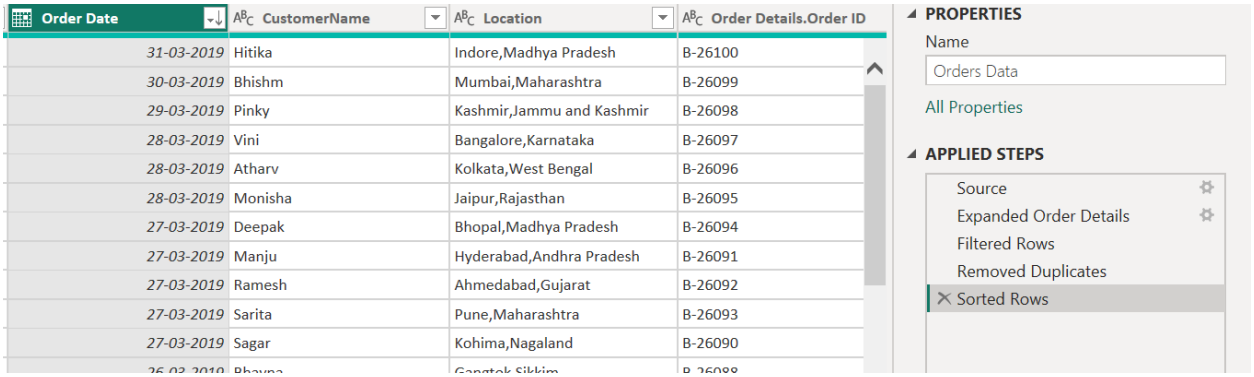
	Order ID	Order Date	CustomerName	Location	Order Details.Order ID
1	B-25601	01-04-2018	Bharat	Ahmedabad,Gujarat	B-25601
2	B-25602	01-04-2018	Pearl	Pune,Maharashtra	B-25602
3	B-25603	03-04-2018	Jahan	Bhopal,Madhya Pradesh	B-25603
4	B-25604	03-04-2018	Divsha	Jaipur,Rajasthan	B-25604
5	B-25605	05-04-2018	Kasheen	Kolkata,West Bengal	B-25605
6	B-25606	06-04-2018	Hazel	Bangalore,Karnataka	B-25606
7	B-25607	06-04-2018	Sonakshi	Kashmir,Jammu and Kashmir	B-25607
8	B-25608	08-04-2018	Aarushi	Chennai,Tamil Nadu	B-25608
9	B-25609	09-04-2018	Jitesh	Lucknow,Uttar Pradesh	B-25609
10	B-25610	09-04-2018	Yogesh	Patna,Bihar	B-25610
11	B-25611	11-04-2018	Anita	Thiruvananthapuram,Kerala	B-25611
12	B-25612	12-04-2018	Shrichand	Chandigarh,Punjab	B-25612
13	B-25613	12-04-2018	Mukesh	Chandigarh,Haryana	B-25613
14	B-25614	13-04-2018	Vandana	Simla,Himachal Pradesh	B-25614
15	B-25615	15-04-2018	Bhavna	Gangtok,Sikkim	B-25615
16	B-25616	15-04-2018	Kanak	Goa,Goa	B-25616

PROPERTIES
Name
Orders Data
All Properties
APPLIED STEPS
Source
Expanded Order Details
Filtered Rows
Removed Duplicates

Sorting and Filtering Data:

Sorting order date column:

Order date column sorted in descending order.

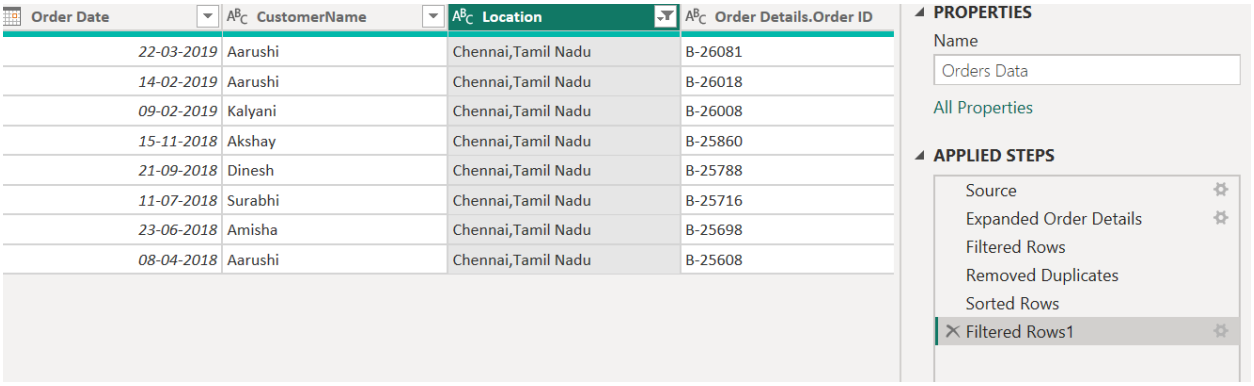


Order Date	CustomerName	Location	Order Details.Order ID
31-03-2019	Hitika	Indore,Madhya Pradesh	B-26100
30-03-2019	Bhishm	Mumbai,Maharashtra	B-26099
29-03-2019	Pinky	Kashmir,Jammu and Kashmir	B-26098
28-03-2019	Vini	Bangalore,Karnataka	B-26097
28-03-2019	Atharv	Kolkata,West Bengal	B-26096
28-03-2019	Monisha	Jaipur,Rajasthan	B-26095
27-03-2019	Deepak	Bhopal,Madhya Pradesh	B-26094
27-03-2019	Manju	Hyderabad,Andhra Pradesh	B-26091
27-03-2019	Ramesh	Ahmedabad,Gujarat	B-26092
27-03-2019	Sarita	Pune,Maharashtra	B-26093
27-03-2019	Sagar	Kohima,Nagaland	B-26090
26-03-2019	Bhavna	Gangtok,Sikkim	B-26088

PROPERTIES
Name
Orders Data
All Properties
APPLIED STEPS
Source
Expanded Order Details
Filtered Rows
Removed Duplicates
Sorted Rows

Filtering location column:

Order date sorted in “descending” and location filtered by a specific state “Tamil Nadu”.

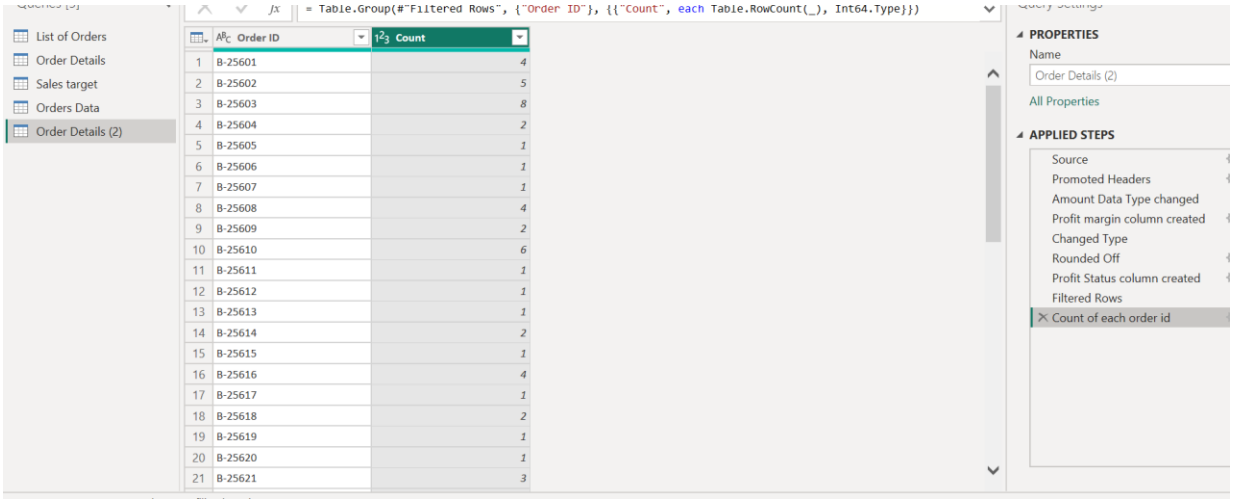


Order Date	CustomerName	Location	Order Details.Order ID
22-03-2019	Aarushi	Chennai,Tamil Nadu	B-26081
14-02-2019	Aarushi	Chennai,Tamil Nadu	B-26018
09-02-2019	Kalyani	Chennai,Tamil Nadu	B-26008
15-11-2018	Akshay	Chennai,Tamil Nadu	B-25860
21-09-2018	Dinesh	Chennai,Tamil Nadu	B-25788
11-07-2018	Surabhi	Chennai,Tamil Nadu	B-25716
23-06-2018	Amisha	Chennai,Tamil Nadu	B-25698
08-04-2018	Aarushi	Chennai,Tamil Nadu	B-25608

Grouping and Aggregating Data:

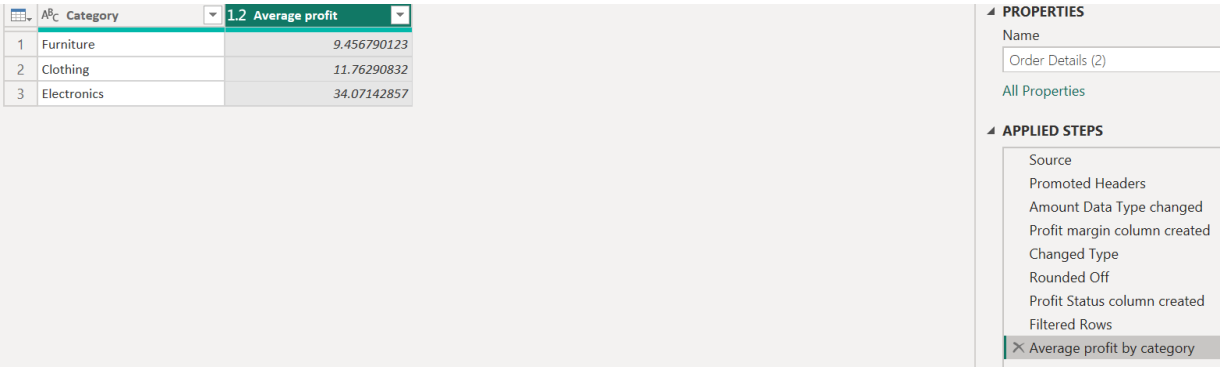
Count of Each order ID:

Order Details table duplicated by right clicking on Order details table and Count of each order id is calculated using Group by function in Home tab→Transform group.



	Order ID	Count
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

Average Profit By category:



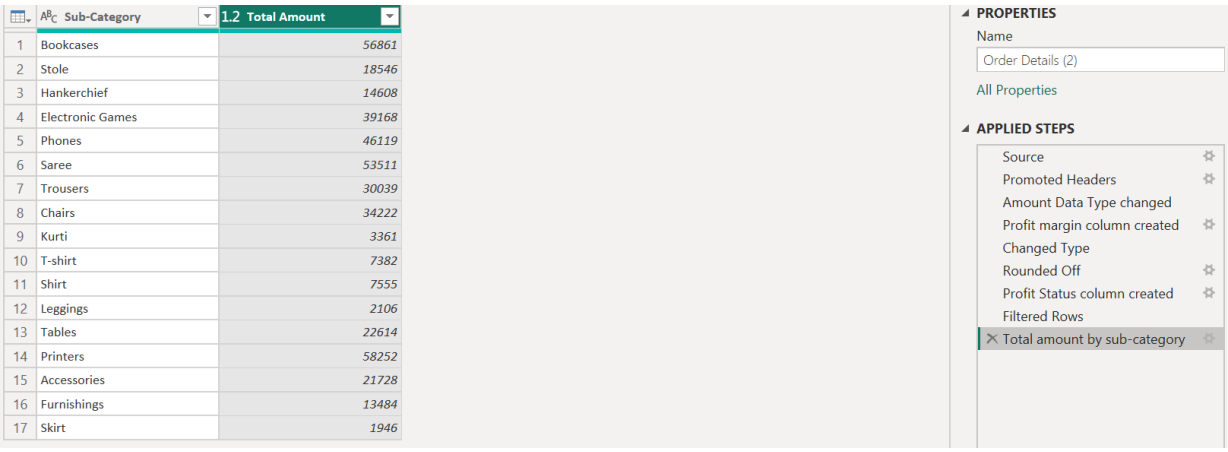
	Category	Average profit
1	Furniture	9.456790123
2	Clothing	11.76290832
3	Electronics	34.07142857

PROPERTIES
Name
Order Details (2)
All Properties

APPLIED STEPS
Source
Promoted Headers
Amount Data Type changed
Profit margin column created
Changed Type
Rounded Off
Profit Status column created
Filtered Rows
Average profit by category

Average profit by category calculated using Group By function.

Total Amount by Sub-Category:



	Sub-Category	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

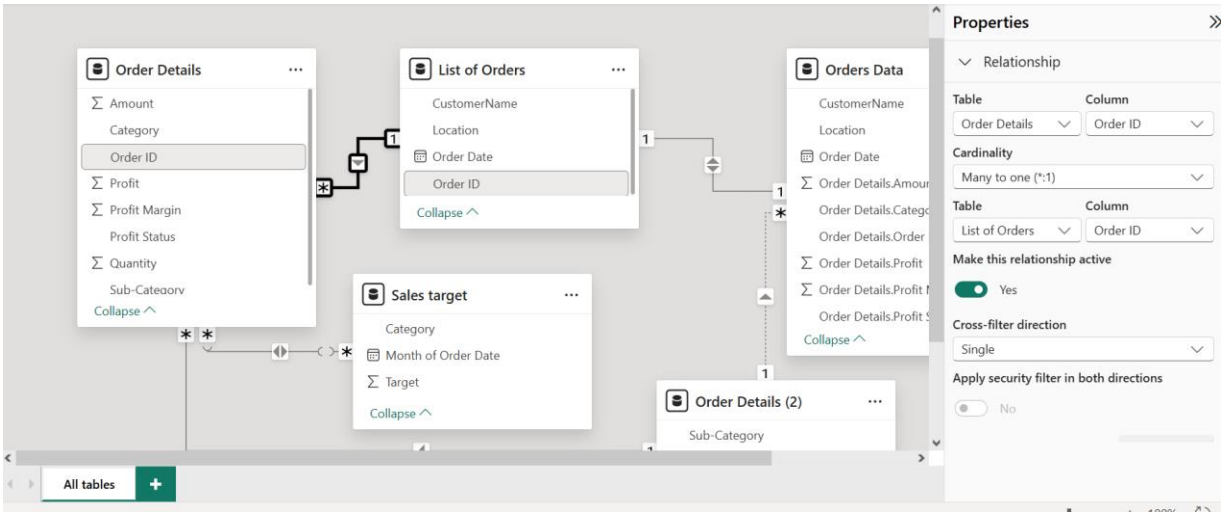
PROPERTIES
Name
Order Details (2)
All Properties

APPLIED STEPS
Source
Promoted Headers
Amount Data Type changed
Profit margin column created
Changed Type
Rounded Off
Profit Status column created
Filtered Rows
Total amount by sub-category

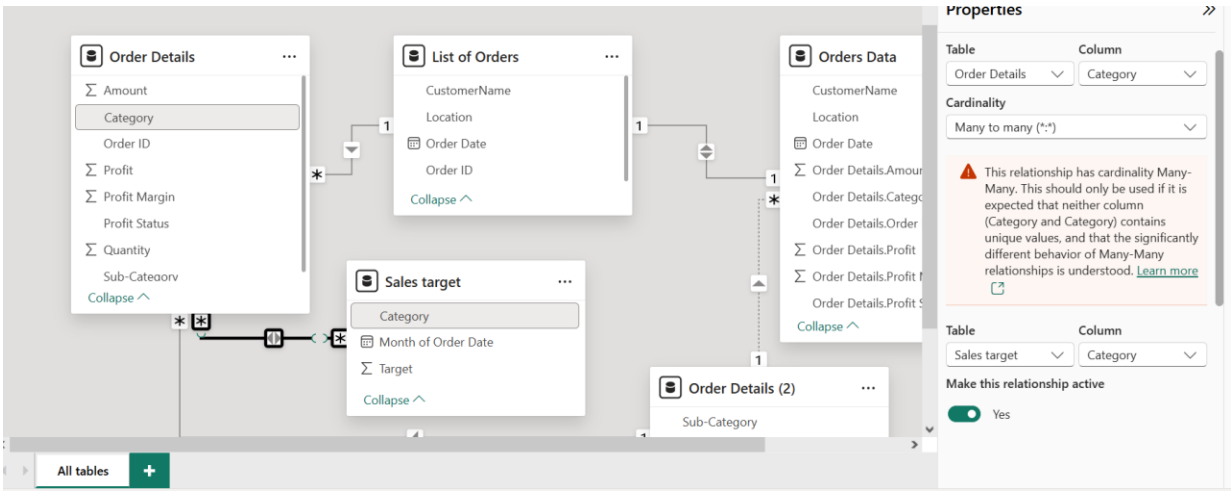
Total amount by sub-category is calculated using group y function.

Data Modeling:

Relationship between the “List of Orders” and “Order Details” tables using the ‘Order ID’ column:



Relationship between the “Order Details” and “Sales Target” tables based on the ‘Category’:



After clicking “Manage Relationship” relationships are active and working.