Assignment 1

Data: Sales_Data

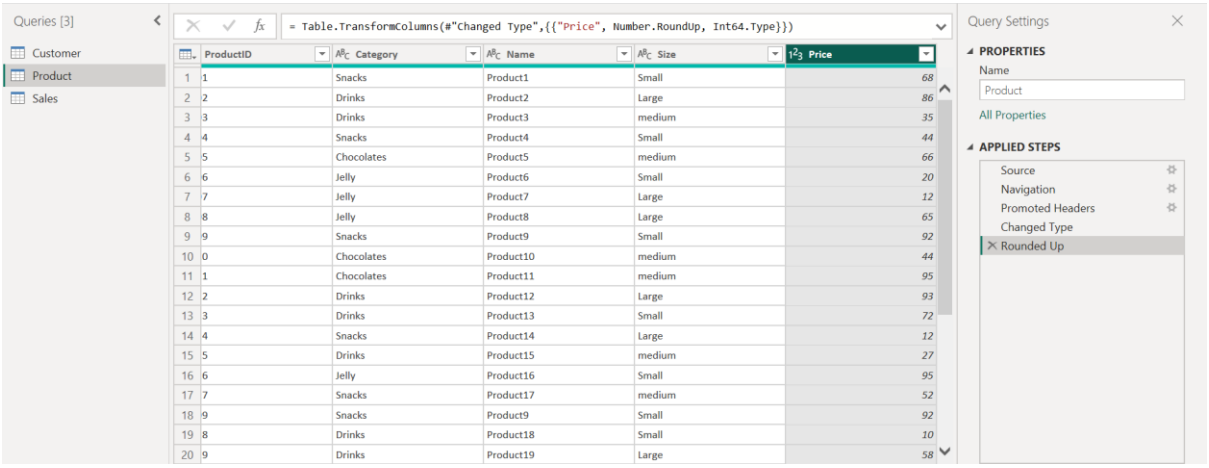
You have three tables: Customer, Product, and Sales.

Task 1:

- Round the 'Price' column in the Product dataset to the nearest integer for simplicity.
- Split the 'Customer' column in the Customer table into two columns: 'FirstName' and 'LastName'.

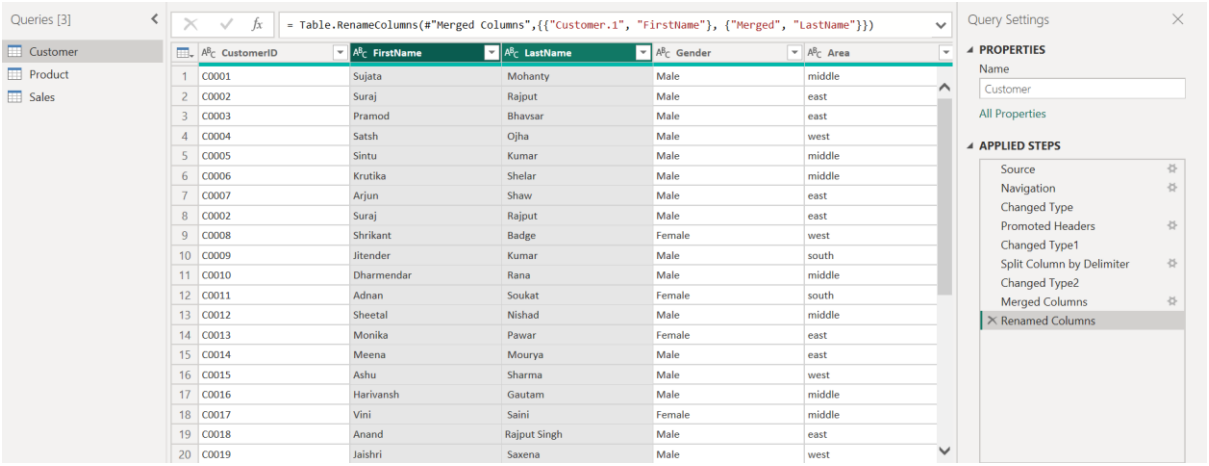
Key 1:

a) Select the 'Price' variable.
b) Click on Transform → Rounding → Choose Round Up.



Key 1.2:

a) Select 'Customer' variable.
b) Transform → Split Column → By Delimiter.
c) Choose Space from select or enter delimiter → Ok.
d) They will split into three variables. Select the second and third variables.
e) Select Merge columns → choose Space from the Separator.
f) Rename both variables.

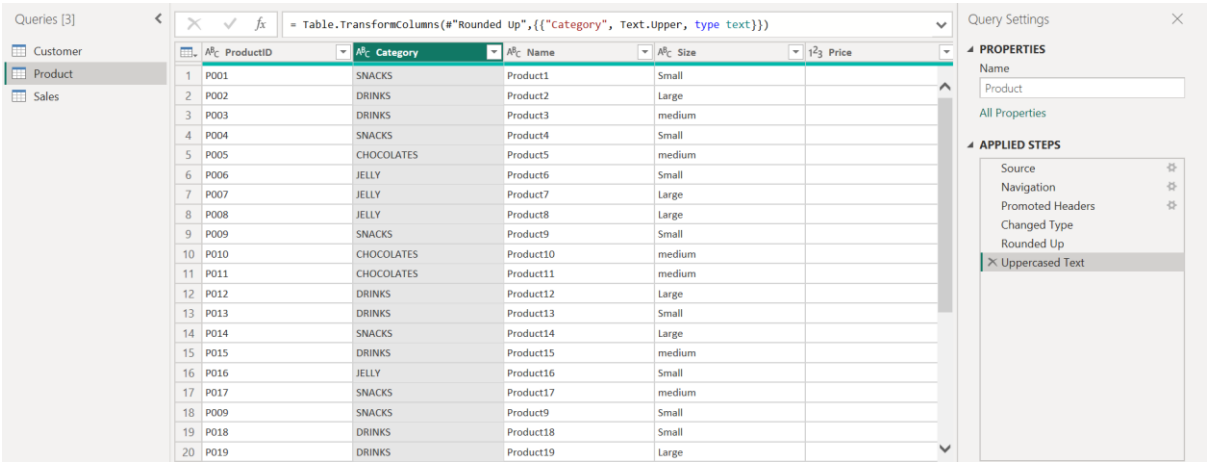


Task 2:

- Convert all entries in the 'Category' column in the Product table to uppercase.
- Replace all occurrences of 'unemployment' with 'Unemployed' in the 'Profession' column of the Customer table.

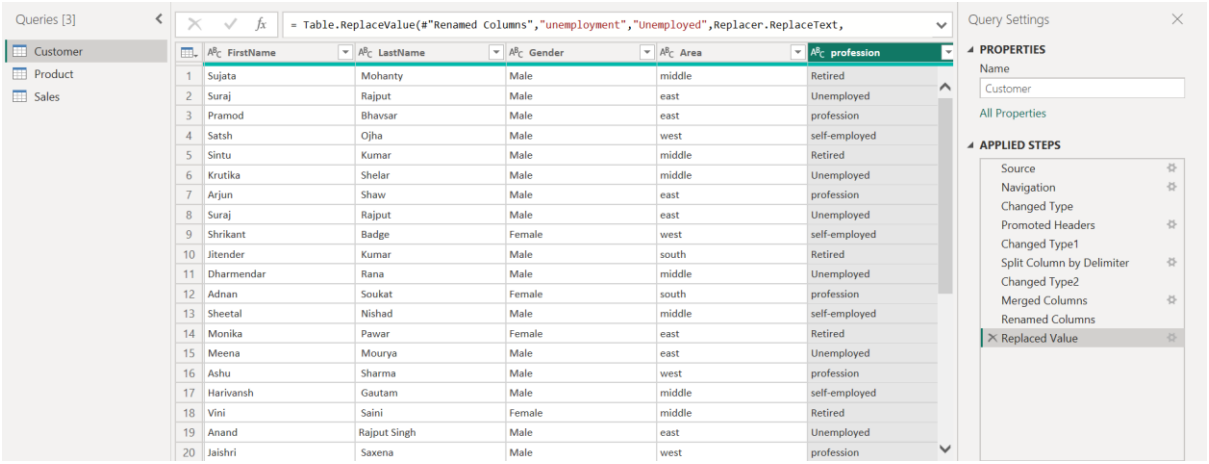
Key 2.1:

a) Select ‘Category’ variable.
b) Transform → Format → UPPERCASE.



Key 2.2:

a) Select ‘Profession’ column.
b) Home → Replace Values.
c) In Value to find enter ‘unemployment’ and replace with ‘Unemployed’.
d) Click Ok.

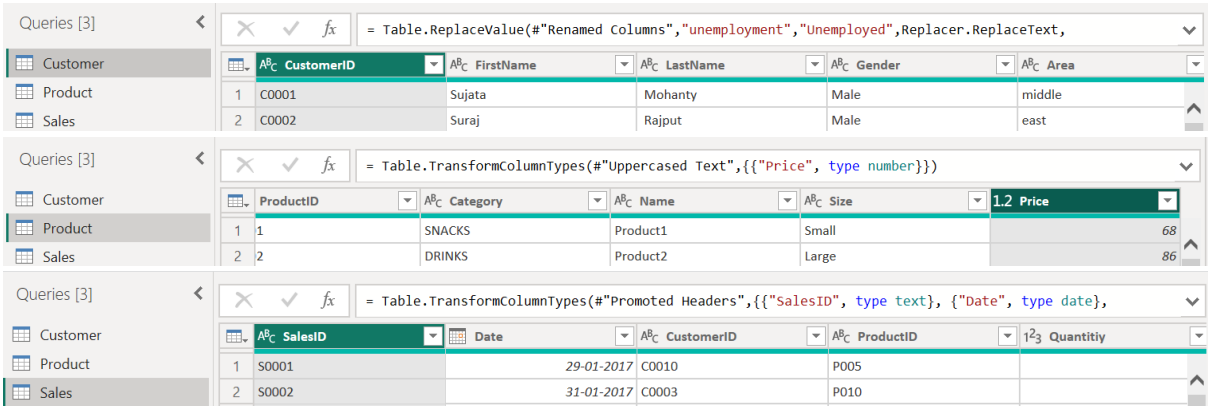


Task 3:

- Ensure all columns in the datasets have appropriate data types, e.g., 'Date' column as Date type, 'Price' as Decimal type.
- Identify and replace any inconsistent values in the 'Size' column of the Product dataset to ensure uniformity (e.g., replace "medium" with "Medium").

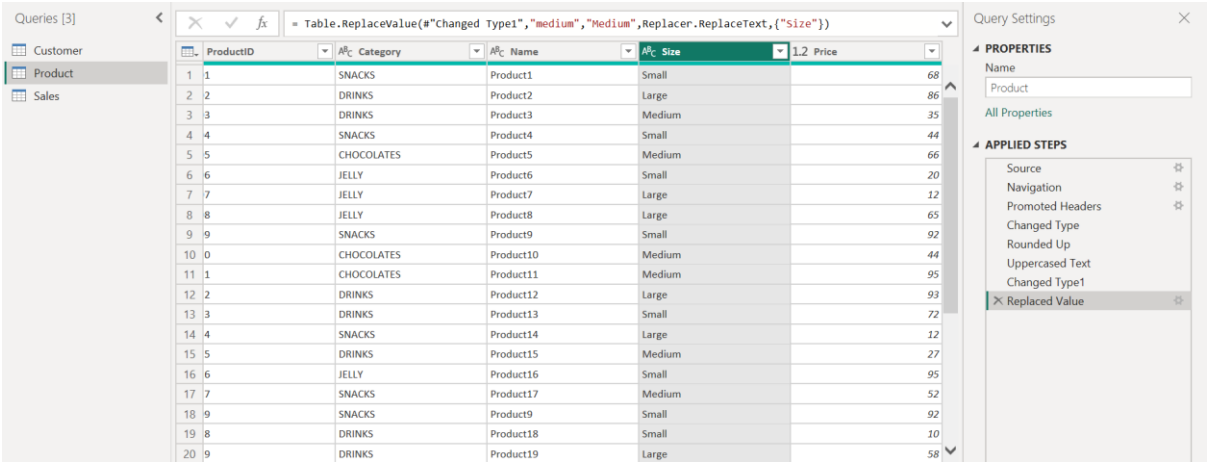
Key 3.1:

a) All columns in the datasets have appropriate data types now. I have attached the screenshot of three tables: Customer, Product, and Sales.



Key 3.2:

a) Select the ‘Size’ variable.
b) Home → Replace Values
c) In Value to find enter ‘medium’ and replace with ‘Medium’.
d) Click Ok.



Task 4:

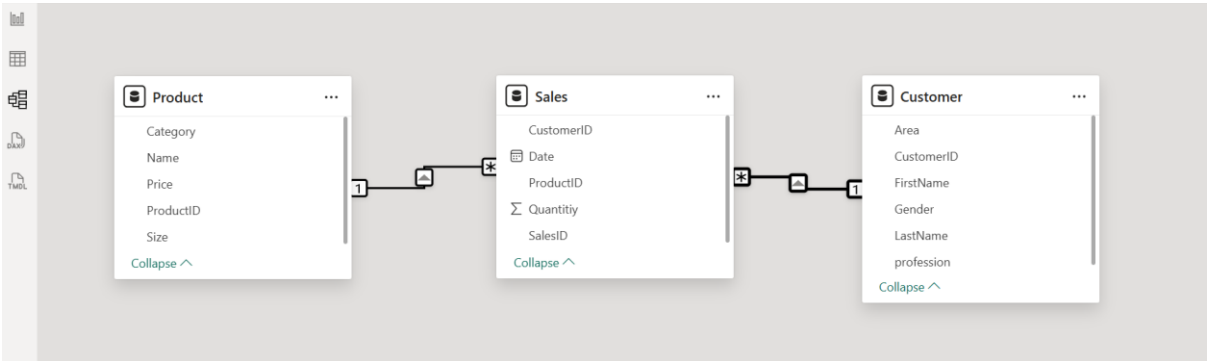
- Create relationships between the tables using 'CustomerID' and 'ProductID' as keys.
- Clean the data by removing any duplicate entries in the Customer and Product tables.

Key 4.1:

a) Click on model view.

b) From the Product table, drag ProductID to the ProductID field in the Sales table. Then, from the Sales table, drag CustomerID to the CustomerID field in the Customer table.

c) Set a one-to-many (1:M) relationship from Product to Sales by selecting the appropriate cardinality. Then, set a many-to-one (M:1) relationship from Sales to Customer by selecting the appropriate cardinality, and click Save.



Key 4.2:

a) Click Transform Data to open the Power Query Editor.

b) In the left panel, select the Customer table.

c) Select all columns.

d) Go to the Home tab and click Remove Duplicates.

e) Repeat the same process for the Product table.

Task 5:

Sales by Category:

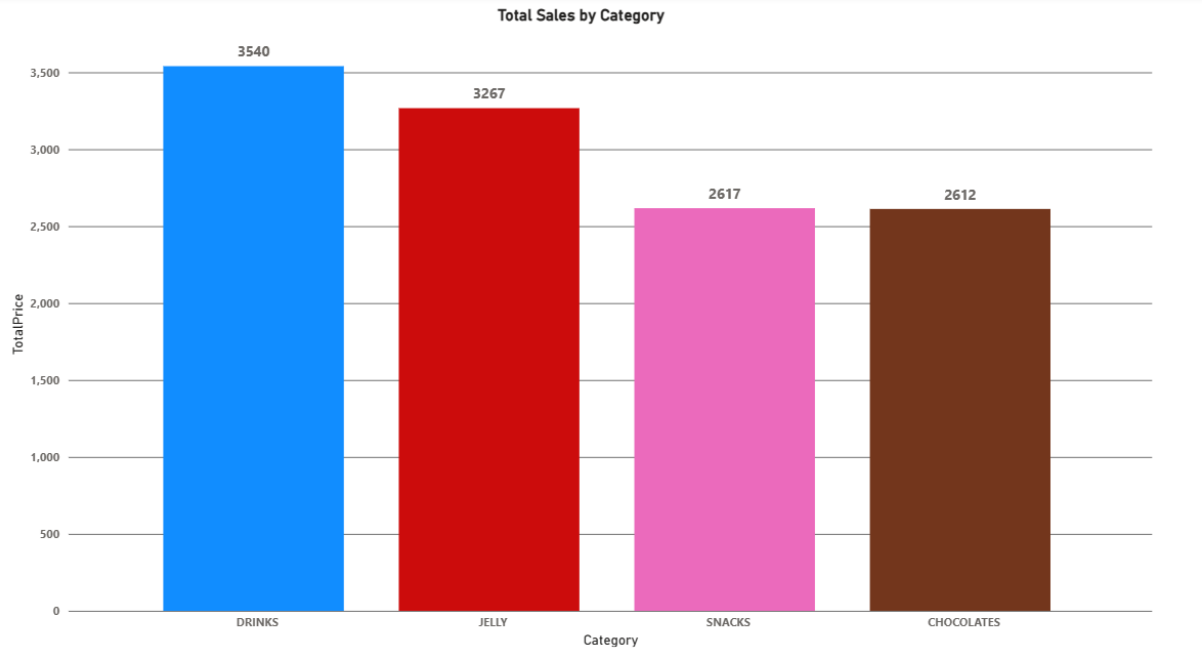
- Create a chart showing total sales (TotalPrice) by product category.
- Customize colors, refine the title, and add data labels for exact sales amounts.

Sales Percentage by Category:

- Create a chart showing the percentage share of sales (Price) for each product category.
- Keep the chart compact and easy to understand.

Key 5.1:

a) In Power BI Desktop, click on Report View.
b) In the Fields pane, click on the Sales table.
c) Right-click the Sales table and select New Measure.
d) Enter the following formula:
TotalSales = SUMX(Sales, Sales[Quantity] * RELATED(Product[Price]))
e) For visualizations, select the Clustered Column Chart.
f) Drag Category from the Product table into the X-Axis field.
g) Drag the newly created TotalSales measure from the Sales table into the Y-Axis field.
h) Go to Add to your visual → Size and style → Columns → customize the colors.
i) Click on Format → select Title → edit the text.
j) Click on Format → Data Labels → turn them On.



Key 5.2:

a) click on Report View.
b) From the Visualizations panel, select the Donut Chart.
c) Drag Product[Category] into the Legend field.
d) Drag TotalSales into the Values field.
e) Click on Format Visual.
f) Go to Detail labels →Label contents → select Percent of total.

g) The Donut Chart will now display the percentage share of sales for each product category.