

Assignment 1

Data: [Sales Data](#)

We 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'.

Explication: (a)

To round the 'Price' column to the nearest integer in Power BI, we use Power Query. I round prices directly in the Power Query.

- We selected the 'Price' column and
- Click on Transform → Rounding → Round to 0 decimal places.

ProductID	Category	Name	Size	Price
1	Snacks	Product1	Small	67
2	Drinks	Product2	Large	85
3	Drinks	Product3	medium	35
4	Snacks	Product4	Small	44
5	Chocolates	Product5	medium	65
6	Jelly	Product6	Small	20
7	Jelly	Product7	Large	11
8	Jelly	Product8	Large	65
9	Snacks	Product9	Small	91
10	Chocolates	Product10	medium	43
11	Chocolates	Product11	medium	95
12	Drinks	Product12	Large	92
13	Drinks	Product13	Small	72
14	Snacks	Product14	Large	12
15	Drinks	Product15	medium	26
16	Jelly	Product16	Small	94
17	Snacks	Product17	medium	52
18	Snacks	Product18	Small	91
19	Drinks	Product19	Small	9
20	Drinks	Product20	Large	57

Explication: (b)

- Split the column:**
 - Selected the 'Customer' column.
 - Clicked Transform → Split Column → By Delimiter.
 - Chose Space (" ") as the delimiter.
 - Selected "Each occurrence of the delimiter". Clicked OK.
 - This created three new columns: Customer.1 (First Name), Customer.2 (Middle Name or Last Name), Customer.3 (Last Name).

- **Handled Three-Word Names:**

1. Clicked at add column and chose custom column.
2. Created another custom column for last name.
3. If the third column was null, took customer. 2 as the last name.
4. If there was a third name, merged customer. 2 and customer. 3.
5. used this formula LastName = if [Customer.3] = null then [Customer.2] else [Customer.2] & " " & [Customer.3].
6. And after deleted the customer.2 and customer.3 columns.
7. Then reordered the column and renamed the columns.

Table.TransformColumnTypes(#"Removed Columns",{{"Last Name", type text}})

	CustomerID	First Name	Last Name	Gender	Area
1	C0001	Sujata	Mohanty	Male	middle
2	C0002	Suraj	Rajput	Male	east
3	C0003	Pramod	Bhavsar	Male	east
4	C0004	Satsh	Ojha	Male	west
5	C0005	Sintu	Kumar	Male	middle
6	C0006	Krutika	Shelar	Male	middle
7	C0007	Arjun	Shaw	Male	east
8	C0002	Suraj	Rajput	Male	east
9	C0008	Shrikant	Badge	Female	west
10	C0009	Jitender	Kumar	Male	south
11	C0010	Dharmendar	Rana	Male	middle
12	C0011	Adnan	Soukat	Female	south
13	C0012	Sheetal	Nishad	Male	middle
14	C0013	Monika	Pawar	Female	east
15	C0014	Meena	Mourya	Male	east
16	C0015	Ashu	Sharma	Male	west
17	C0016	Harivansh	Gautam	Male	middle
18	C0017	Vini	Saini	Female	middle
19	C0018	Anand	Singh Rajput	Male	east
20	C0019	Jaishri	Saxena	Male	west
21					

Query Settings

PROPERTIES

Name
Customer

All Properties

APPLIED STEPS

- Source
- Navigation
- Changed Type
- Promoted Headers
- Changed Type1
- Split Column by Delimiter
- Changed Type2
- Added Custom
- Reordered Columns
- Renamed Columns
- Removed Columns
- Changed Type3

Preview based on top 1000 rows

PREVIEW DOWNLOADED AT 09:15

Task 2:

(c). Convert all entries in the 'Category' column in the Product table to uppercase.

(d). Replace all occurrences of 'unemployment' with 'Unemployed' in the 'Profession' column of the Customer table.

Explication: (c)

1. Selected category column.
2. Clicked on Transform → Format → Uppercase.
3. Then, all category names be in **UPPERCASE**.

Query Settings

Query Settings

Table.TransformColumns(#"Changed Type1",{{"Category", Text.Upper, type text}})

	ProductID	Category	Name	Size	Price
1	P001	SNACKS	Product1	Small	
2	P002	DRINKS	Product2	Large	
3	P003	DRINKS	Product3	medium	
4	P004	SNACKS	Product4	Small	
5	P005	CHOCOLATES	Product5	medium	
6	P006	JELLY	Product6	Small	
7	P007	JELLY	Product7	Large	
8	P008	JELLY	Product8	Large	
9	P009	SNACKS	Product9	Small	
10	P010	CHOCOLATES	Product10	medium	
11	P011	CHOCOLATES	Product11	medium	
12	P012	DRINKS	Product12	Large	
13	P013	DRINKS	Product13	Small	
14	P014	SNACKS	Product14	Large	
15	P015	DRINKS	Product15	medium	
16	P016	JELLY	Product16	Small	
17	P017	SNACKS	Product17	medium	
18	P009	SNACKS	Product9	Small	
19	P018	DRINKS	Product18	Small	
20	P019	DRINKS	Product19	Large	

PROPERTIES

Name

Product

APPLIED STEPS

Source

Navigation

Promoted Headers

Changed Type

Rounded Off

Changed Type1

× Uppercased Text

Explication: (d)

1. Selected Customer Table in Power Query editor.
2. Then selected the 'Profession' column.
3. Clicked on Transform → Replace Values.
4. In the 'Value to Find' box, entered: unemployment.
5. In the 'Replace With' box, entered: Unemployed.
6. Then clicked ok.
7. Then, all occurrences of 'unemployment' have been replaced with 'Unemployed'.

Query Settings

Query Settings

Table.ReplaceValue(#"Changed Type3", "unemployment", "Unemployed", Replacer.ReplaceText, {"profession"})

Last Name	Gender	Area	profession
Mohanty	Male	middle	Retired
Rajput	Male	east	Unemployed
Bhavsar	Male	east	profession
Ojha	Male	west	self-employed
Kumar	Male	middle	Retired
Shelar	Male	middle	Unemployed
Shaw	Male	east	profession
Rajput	Male	east	Unemployed
Badge	Female	west	self-employed
Kumar	Male	south	Retired
Rana	Male	middle	Unemployed
Soukat	Female	south	profession
Nishad	Male	middle	self-employed
Pawar	Female	east	Retired
Mourya	Male	east	Unemployed
Sharma	Male	west	profession
Gautam	Male	middle	self-employed
Saini	Female	middle	Retired
Singh Rajput	Male	east	Unemployed
Saxena	Male	west	profession

PROPERTIES

Name

Customer

APPLIED STEPS

Source

Navigation

Changed Type

Promoted Headers

Changed Type1

Split Column by Delimiter

Changed Type2

Added Custom

Reordered Columns

Renamed Columns

Removed Columns

Changed Type3

× Replaced Value

Task 3:

(e). Ensure all columns in the datasets have appropriate data types, e.g., 'Date' column as Date type, 'Price' as Decimal type.

(f). Identify and replace any inconsistent values in the 'Size' column of the Product dataset to ensure uniformity (e.g., replace "medium" with "Medium").

Explication: (e)

1. All columns in the datasets have appropriate data types now. I have attached the screenshot. Three tables: Customer, Product, and Sales.

= Table.ReplaceValue("#Changed Type3","unemployment","Unemployed",Replacer.ReplaceText,{"profession"})						
AB_C CustomerID	AB_C First Name	AB_C Last Name	AB_C Gender	AB_C Area	AB_C profession	
1 C0001	Sujata	Mohanty	Male	middle	Retired	
2 C0002	Surai	Rainut	Male	east	Unemployed	

= Table.ReplaceValue("#Changed Type2","medium","Medium",Replacer.ReplaceText,{"Size"})				
AB_C ProductID	AB_C Category	AB_C Name	AB_C Size	1.2 Price
1 P001	SNACKS	Product1	Small	67
2 P002	DRINKS	Product2	Large	85

= Table.TransformColumnTypes("#Promoted Headers",{{"SalesID", type text}, {"Date", type date}, {"CustomerID", type text}, {"ProductID", type text}, {"Quantity", type decimal}}					
AB_C SalesID	Date	AB_C CustomerID	AB_C ProductID	1.2 Quantity	
1 S0001	29-01-2017	C0010	P005	7	
2 S0002	31-01-2017	C0003	P010	10	

Explication: (f)

1. Selected Product Table in Power Query editor.
2. Then selected the 'Size' column.
3. Clicked on Transform → Replace Values.
4. In the 'Value to Find' box, entered: medium.
5. In the 'Replace With' box, entered: Medium.
6. Then clicked ok.
7. Then, all occurrences of 'medium' have been replaced with 'Medium'.

= Table.ReplaceValue("#Changed Type2","medium","Medium",Replacer.ReplaceText,{"Size"})				
AB_C ProductID	AB_C Category	AB_C Name	AB_C Size	1.2 Price
1 P001	SNACKS	Product1	Small	67
2 P002	DRINKS	Product2	Large	85
3 P003	DRINKS	Product3	Medium	35
4 P004	SNACKS	Product4	Small	44
5 P005	CHOCOLATES	Product5	Medium	65
6 P006	JELLY	Product6	Small	20
7 P007	JELLY	Product7	Large	11
8 P008	JELLY	Product8	Large	65
9 P009	SNACKS	Product9	Small	91
10 P010	CHOCOLATES	Product10	Medium	43
11 P011	CHOCOLATES	Product11	Medium	95
12 P012	DRINKS	Product12	Large	92
13 P013	DRINKS	Product13	Small	72
14 P014	SNACKS	Product14	Large	12
15 P015	DRINKS	Product15	Medium	26

Query Settings

PROPERTIES

Name
Product

All Properties

APPLIED STEPS

Source

Navigation

Promoted Headers

Changed Type

Rounded Off

Changed Type1

Uppercased Text

Changed Type2

Replaced Value

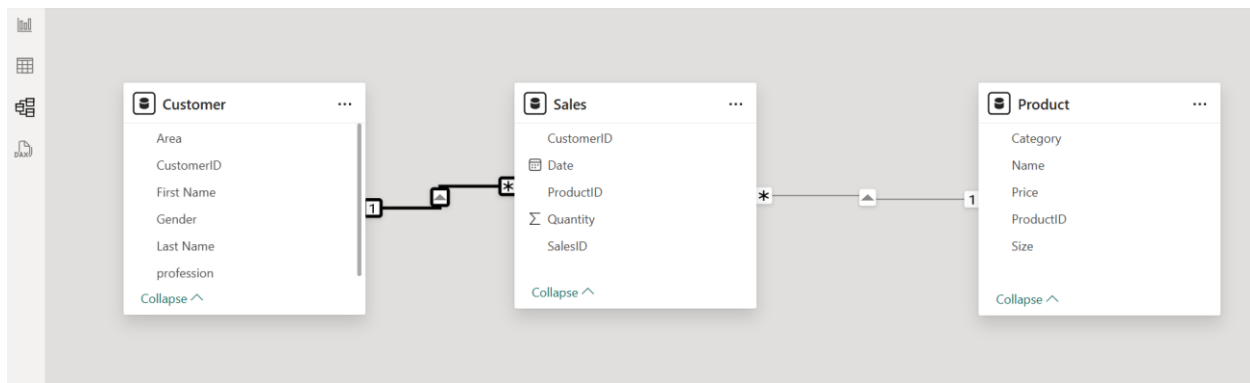
Task 4:

(g). Create relationships between the tables using 'CustomerID' and 'ProductID' as keys.

(h). Clean the data by removing any duplicate entries in the Customer and Product tables.

Explication: (g)

1. After completing the above Task 1, Task 2, and Task 3 we click File → Close & Apply.
2. Click on model view icon from the left pane.
3. From the customer table drag CustomerID in sales table to CustomerID.
4. A one-to-many (1:M) relationship (Customer to Sales) selected from cardinality then click on save.
5. From the Sales table we drag ProductID in Product Table to ProductID.
6. A many-to-one (M:1) relationship (Sales to Product) selected from cardinality then click on save.



Explication: (h)

1. Clicked on Transform Data to open Power Query Editor.
2. Selected the Customer table from the left panel.
3. Clicked on the CustomerID column.
4. clicked to the Home tab and click Remove Duplicates.
5. Same process for the Product table I selected ProductId columns.

Task 5:

Sales by Category:

(i). Create a chart showing total sales (TotalPrice) by product category.

(j). Customize colors, refine the title, and add data labels for exact sales amounts.

Sales Percentage by Category:

(k). Create a chart showing the percentage share of sales (Price) for each product

category.

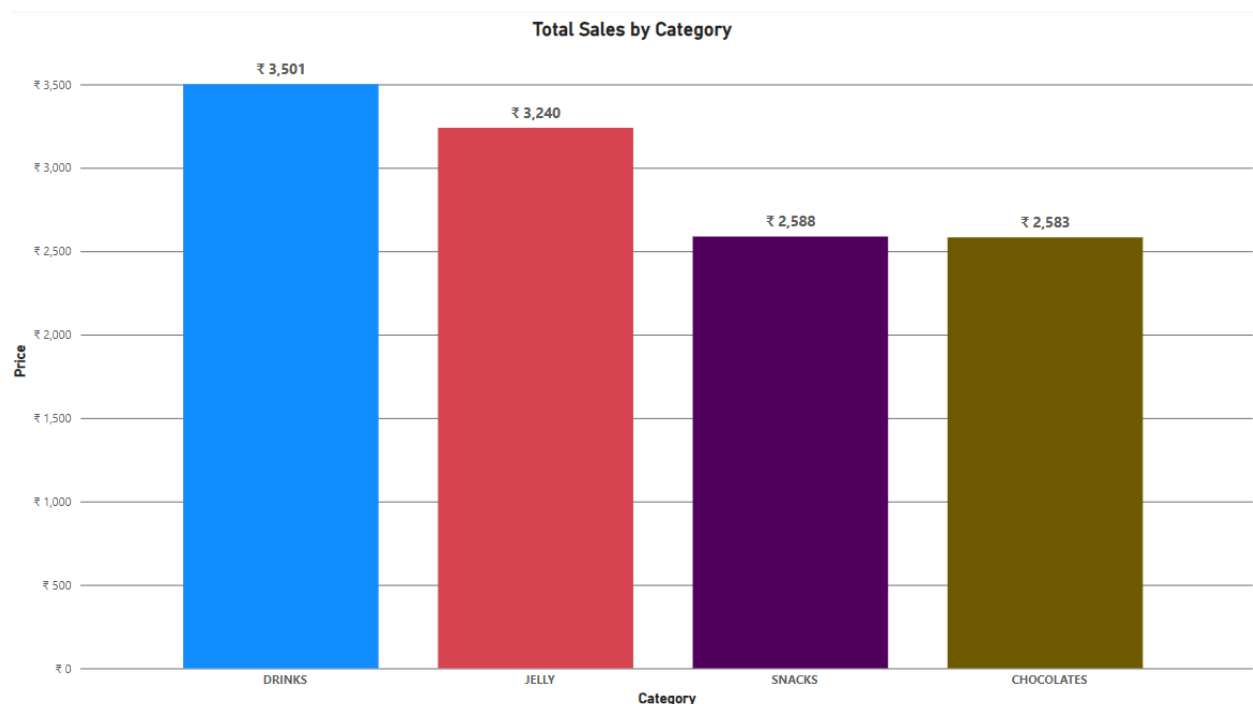
(I). Keep the chart compact and easy to understand.

Explication: (i)

1. Click on report view in Power BI Desktop.
2. Click on the sales table in the fields pane.
3. Right click on sales table → select new measure and used this formula $\text{TotalSales} = \text{SUMX}(\text{Sales}, \text{Sales}[\text{Quantity}] * \text{RELATED}(\text{Product}[\text{Price}]))$. Formula of Totalsales = Numberofitemssold X Salesprice.

Explication: (j)

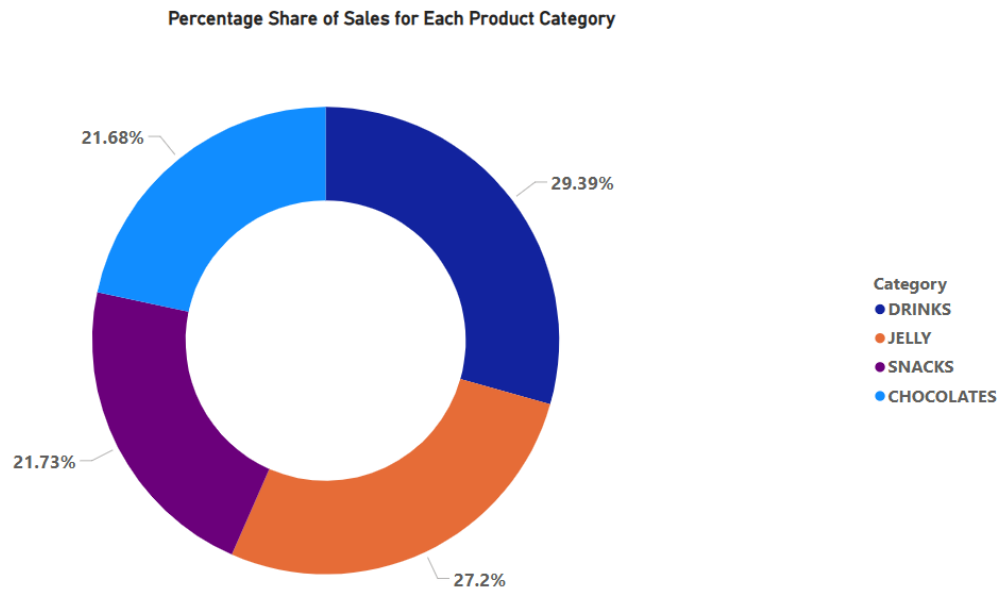
4. For Visualizations, selected Clustered Column Chart.
5. Drag Category from the Product table into the X-Axis field.
6. Drag the newly created TotalSales measure into the Y-Axis field from sales table.
7. Click on add to your visual → size and style → Columns → Customize the colors
8. click on Format → Select Title → Text
9. click on format → Data Labels → on



Explication: (k)

1. click to Report View in Power BI.
2. Select Donut Chart from the visualizations panel.
3. Assign Data Fields Drag Product[Category] into the Legend field.

4. Drag TotalSales into the Values field.
 5. Click on the chart, then go to the Format Pane (Paint Roller Icon).
 6. click on the detail labels →Label contents→Percent of total
 7. Now, our Donut Chart are displaying the percentage share of sales for each product category.
-



Assignment completed by Ravi Kant