

# 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.

The screenshot shows the Power Query Editor interface. On the left, the 'Queries [3]' pane lists 'Customer', 'Product', and 'Sales', with 'Product' selected. The main area displays a table with 20 rows and 5 columns: ProductID, Category, Name, Size, and Price. The Price column values are rounded up to the nearest integer. The formula bar at the top shows the M code: `= Table.TransformColumns(#"Changed Type",{{"Price", Number.RoundUp, Int64.Type}})`. On the right, the 'Query Settings' pane shows the 'Name' as 'Product' and the 'APPLIED STEPS' list containing 'Source', 'Navigation', 'Promoted Headers', 'Changed Type', and 'Rounded Up'.

	ProductID	Category	Name	Size	Price
1	1	Snacks	Product1	Small	68
2	2	Drinks	Product2	Large	86
3	3	Drinks	Product3	medium	35
4	4	Snacks	Product4	Small	44
5	5	Chocolates	Product5	medium	66
6	6	Jelly	Product6	Small	20
7	7	Jelly	Product7	Large	12
8	8	Jelly	Product8	Large	65
9	9	Snacks	Product9	Small	92
10	0	Chocolates	Product10	medium	44
11	1	Chocolates	Product11	medium	95
12	2	Drinks	Product12	Large	93
13	3	Drinks	Product13	Small	72
14	4	Snacks	Product14	Large	12
15	5	Drinks	Product15	medium	27
16	6	Jelly	Product16	Small	95
17	7	Snacks	Product17	medium	52
18	9	Snacks	Product9	Small	92
19	8	Drinks	Product18	Small	10
20	9	Drinks	Product19	Large	58

### 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.

Queries [3] ✕ ✓ fx = Table.RenameColumns(#"Merged Columns",{{"Customer.1", "FirstName"}, {"Merged", "LastName"}})

	APC CustomerID	APC FirstName	APC LastName	APC Gender	APC 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	Rajput Singh	Male	east
20	C0019	Jaishri	Saxena	Male	west

Query Settings ✕

**PROPERTIES**

Name  
Customer

All Properties

**APPLIED STEPS**

- Source
- Navigation
- Changed Type
- Promoted Headers
- Changed Type1
- Split Column by Delimiter
- Changed Type2
- Merged Columns
- Renamed Columns**

## 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:

- Select 'Category' variable.
- Transform → Format → UPPERCASE.

Queries [3] ✕ ✓ fx = Table.TransformColumns(#"Rounded Up",{{"Category", Text.Upper, type text}})

	APC ProductID	APC Category	APC Name	APC Size	APC 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	

Query Settings ✕

**PROPERTIES**

Name  
Product

All Properties

**APPLIED STEPS**

- Source
- Navigation
- Promoted Headers
- Changed Type
- Rounded Up
- Uppercased Text**

## Key 2.2:

- Select 'Profession' column.
- Home → Replace Values.
- In Value to find enter 'unemployment' and replace with 'Unemployed'.
- Click Ok.

Queries [3] ✕ ✓ fx = Table.ReplaceValue(#"Renamed Columns", "unemployment", "Unemployed", Replacer.ReplaceText, ✕

	FirstName	LastName	Gender	Area	profession
1	Sujata	Mohanty	Male	middle	Retired
2	Suraj	Rajput	Male	east	Unemployed
3	Pramod	Bhavsar	Male	east	profession
4	Satsh	Ojha	Male	west	self-employed
5	Sintu	Kumar	Male	middle	Retired
6	Krutika	Shelar	Male	middle	Unemployed
7	Arjun	Shaw	Male	east	profession
8	Suraj	Rajput	Male	east	Unemployed
9	Shrikant	Badge	Female	west	self-employed
10	Jitender	Kumar	Male	south	Retired
11	Dharmendar	Rana	Male	middle	Unemployed
12	Adnan	Soukat	Female	south	profession
13	Sheetal	Nishad	Male	middle	self-employed
14	Monika	Pawar	Female	east	Retired
15	Meena	Mourya	Male	east	Unemployed
16	Ashu	Sharma	Male	west	profession
17	Harivansh	Gautam	Male	middle	self-employed
18	Vini	Salni	Female	middle	Retired
19	Anand	Rajput Singh	Male	east	Unemployed
20	Jaishri	Saxena	Male	west	profession

Query Settings ✕

**PROPERTIES**

Name  
Customer

All Properties

**APPLIED STEPS**

- Source
- Navigation
- Changed Type
- Promoted Headers
- Changed Type1
- Split Column by Delimiter
- Changed Type2
- Merged Columns
- Renamed Columns
- ✕ Replaced Value

### 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.

Queries [3] ✕ ✓ fx = Table.ReplaceValue(#"Renamed Columns", "unemployment", "Unemployed", Replacer.ReplaceText, ✕

	CustomerID	FirstName	LastName	Gender	Area
1	C0001	Sujata	Mohanty	Male	middle
2	C0002	Suraj	Rajput	Male	east

Queries [3] ✕ ✓ fx = Table.TransformColumnTypes(#"Uppercased Text", {"Price", type number}) ✕

	ProductID	Category	Name	Size	Price
1	1	SNACKS	Product1	Small	68
2	2	DRINKS	Product2	Large	86

Queries [3] ✕ ✓ fx = Table.TransformColumnTypes(#"Promoted Headers", {"SalesID", type text}, {"Date", type date}, ✕

	SalesID	Date	CustomerID	ProductID	Quantity
1	S0001	29-01-2017	C0010	P005	
2	S0002	31-01-2017	C0003	P010	

### Key 3.2:

- Select the 'Size' variable.
- Home → Replace Values
- In Value to find enter 'medium' and replace with 'Medium'.
- Click Ok.

Queries [3]    `= Table.ReplaceValue(#"changed Type1","medium","Medium",Replacer.ReplaceText,{"Size"})`

	ProductID	Category	Name	Size	Price
1	1	SNACKS	Product1	Small	68
2	2	DRINKS	Product2	Large	86
3	3	DRINKS	Product3	Medium	35
4	4	SNACKS	Product4	Small	44
5	5	CHOCOLATES	Product5	Medium	66
6	6	JELLY	Product6	Small	20
7	7	JELLY	Product7	Large	12
8	8	JELLY	Product8	Large	65
9	9	SNACKS	Product9	Small	92
10	0	CHOCOLATES	Product10	Medium	44
11	1	CHOCOLATES	Product11	Medium	95
12	2	DRINKS	Product12	Large	93
13	3	DRINKS	Product13	Small	72
14	4	SNACKS	Product14	Large	12
15	5	DRINKS	Product15	Medium	27
16	6	JELLY	Product16	Small	95
17	7	SNACKS	Product17	Medium	52
18	9	SNACKS	Product9	Small	92
19	8	DRINKS	Product18	Small	10
20	9	DRINKS	Product19	Large	58

Query Settings

**PROPERTIES**

Name: Product

**APPLIED STEPS**

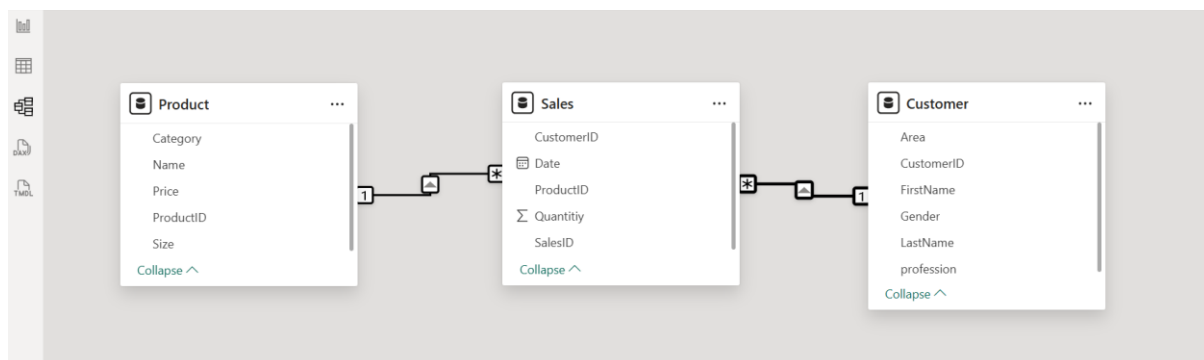
- Source
- Navigation
- Promoted Headers
- Changed Type
- Rounded Up
- Uppercased Text
- Changed Type1
- Replaced Value

#### 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:

- Click on model view.
- 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.
- 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:

- Click Transform Data to open the Power Query Editor.
- In the left panel, select the Customer table.
- Select all columns.
- Go to the Home tab and click Remove Duplicates.
- 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.

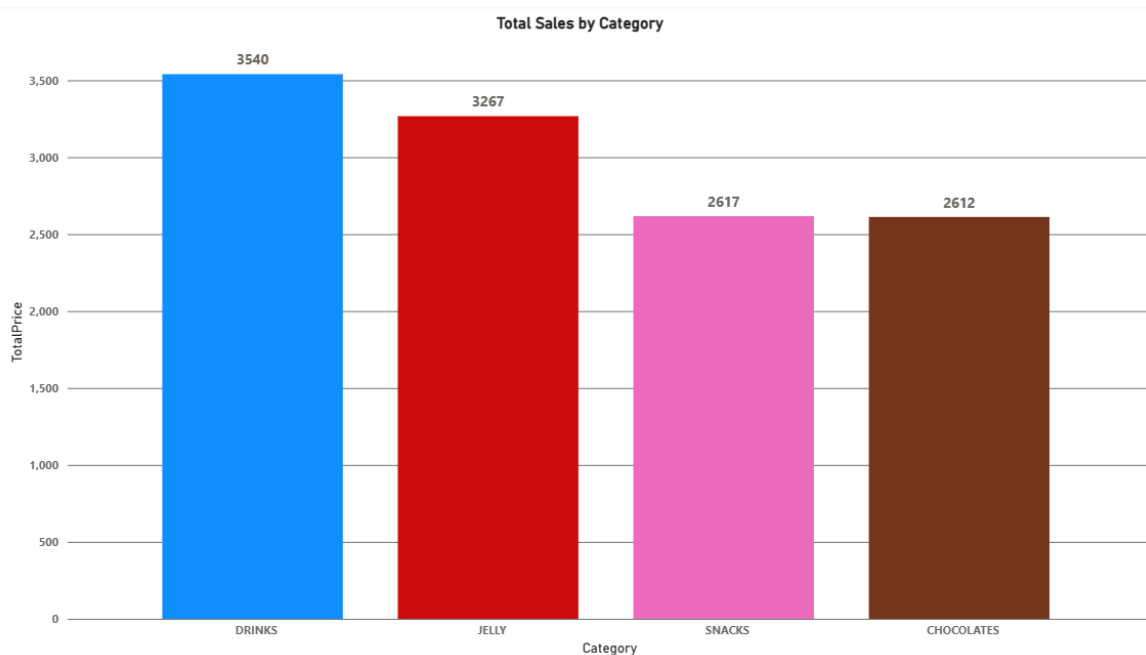
### Sales Percentage by Category:

- 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:

- In Power BI Desktop, click on Report View.
- In the Fields pane, click on the Sales table.
- Right-click the Sales table and select New Measure.
- Enter the following formula:  

$$\text{TotalSales} = \text{SUMX}(\text{Sales}, \text{Sales}[\text{Quantity}] * \text{RELATED}(\text{Product}[\text{Price}]))$$
- For visualizations, select the Clustered Column Chart.
- Drag Category from the Product table into the X-Axis field.
- Drag the newly created TotalSales measure from the Sales table into the Y-Axis field.
- Go to Add to your visual → Size and style → Columns → customize the colors.
- Click on Format → select Title → edit the text.
- Click on Format → Data Labels → turn them On.



### Key 5.2:

- click on Report View.
- From the Visualizations panel, select the Donut Chart.
- Drag Product[Category] into the Legend field.
- Drag TotalSales into the Values field.
- Click on Format Visual.
- 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.

