**Assignment 1**

**Data:** [**Sales\_Data**](https://drive.google.com/file/d/10kmZayKtKYv7BByv9djEAZYeDCrf0AjY/view)

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

**Task 1:**

(a). Round the 'Price' column in the Product dataset to the nearest integer for simplicity.

(b). 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.

1. We selected the ‘Price’ column and
2. Click on Transform → Rounding → Round to 0 decimal places.

A screenshot of a computer

AI-generated content may be incorrect.

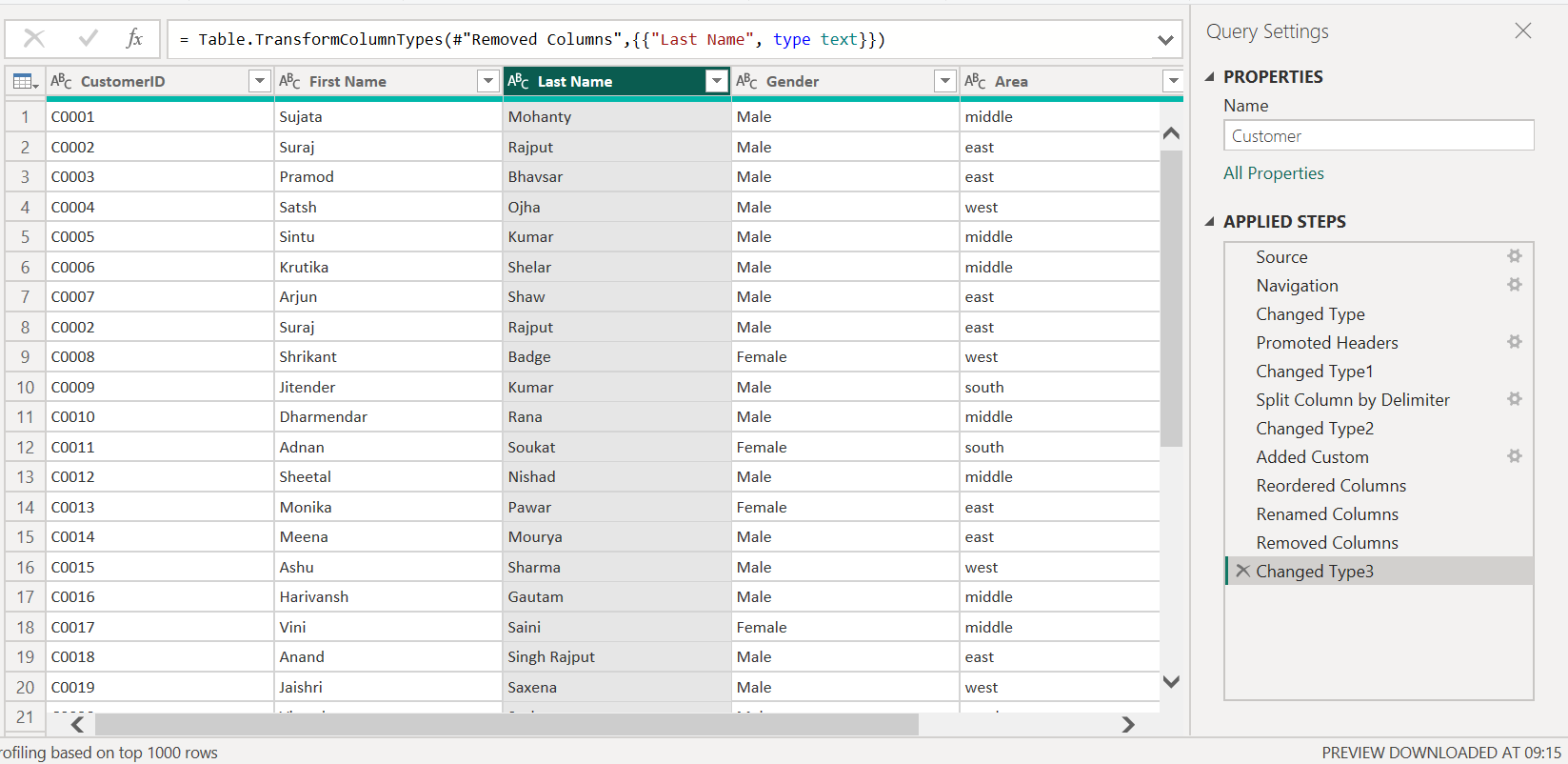
**Explication: (b)**

* **Split the column:**

1. Selected the 'Customer' column.
2. Clicked Transform → Split Column → By Delimiter.
3. Chose Space (" ") as the delimiter.
4. Selected "Each occurrence of the delimiter". Clicked OK.
5. 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.



**Task 2:**

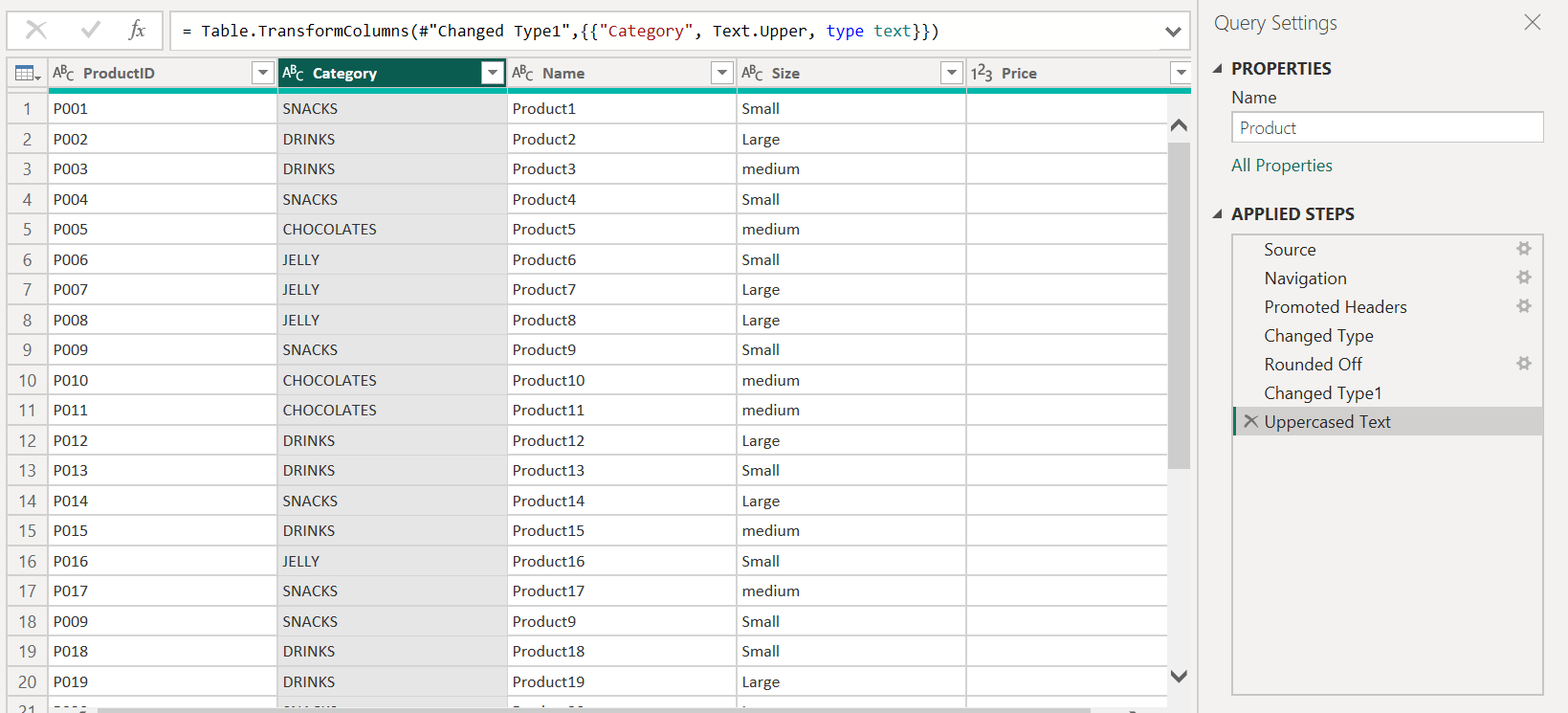
(c). Convert all entries in the 'Category' column in the Product table to uppercase.

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

of the Customer table.

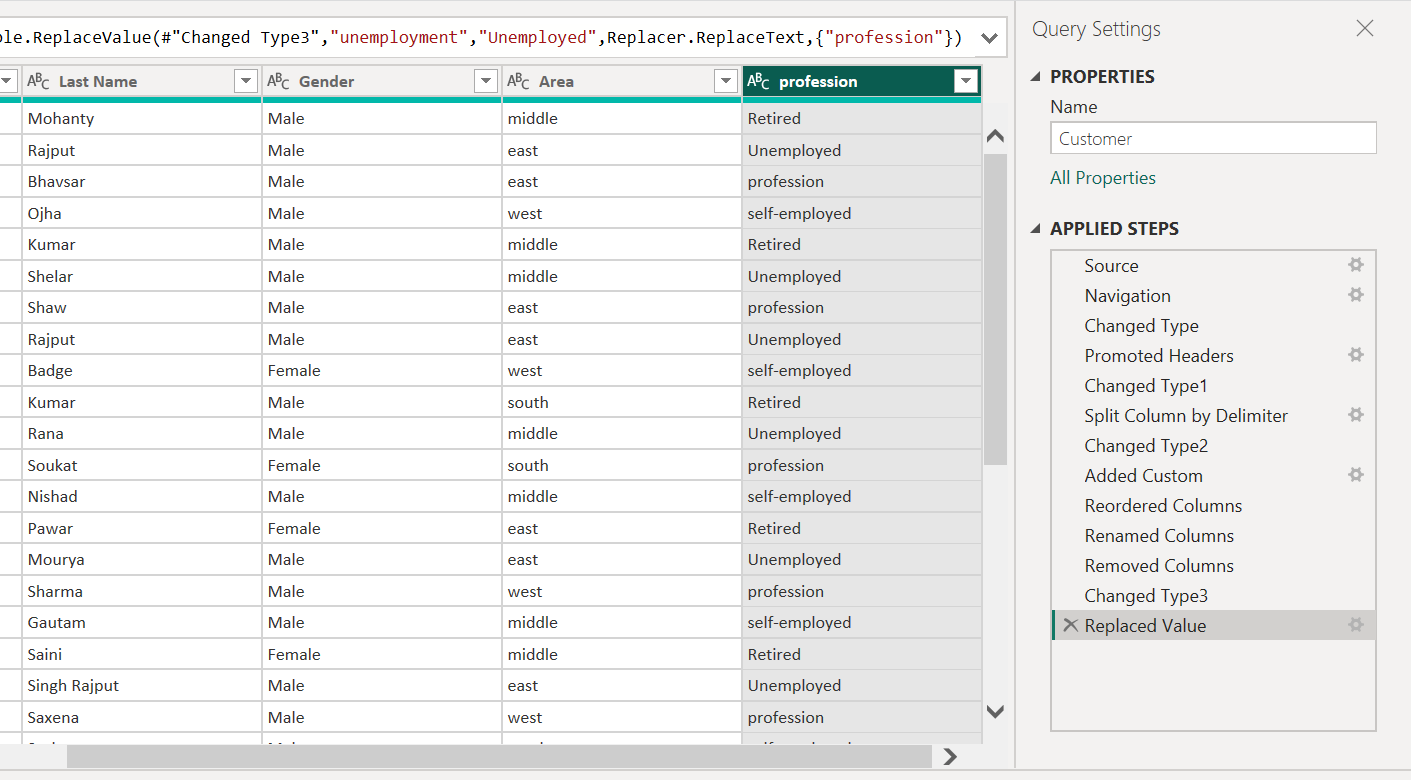
**Explication: (c)**

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



**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'.



**Task 3:**

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

Date type, 'Price' as Decimal type.

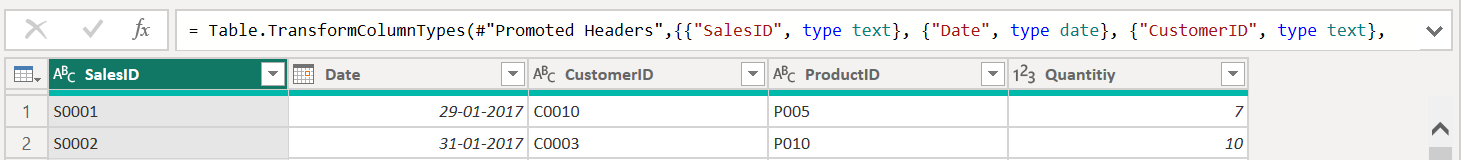
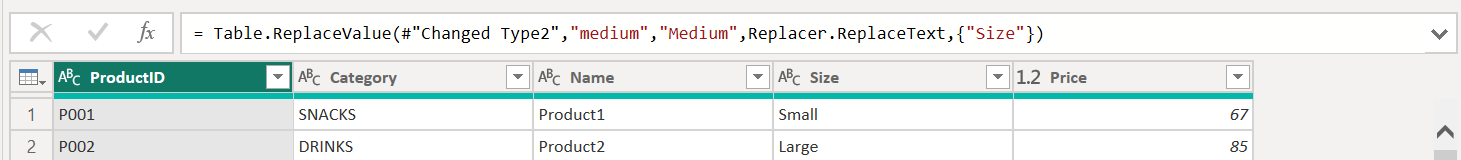
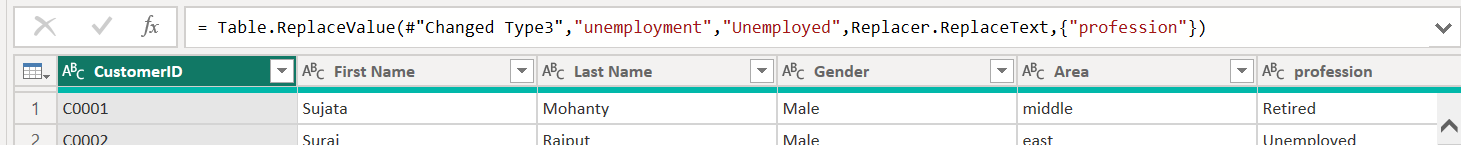
(f). Identify and replace any inconsistent values in the 'Size' column of the Product dataset

to ensure uniformity (e.g., replace "medium" with "Medium").

**Explication: (e)**

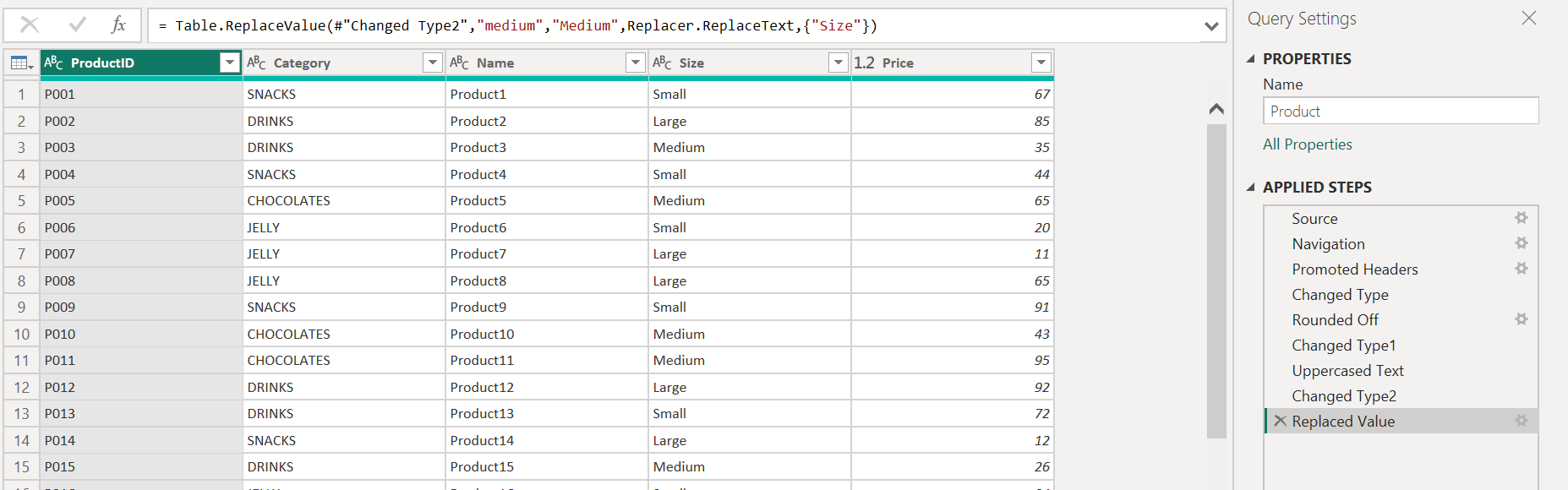
1. All columns in the datasets have appropriate data types now. I have attached the screenshot.

Three tables: Customer, Product, and Sales.



**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'.



**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.

A screenshot of a computer

AI-generated content may be incorrect.

**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.

(l). 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 TotalSales = SUMX(Sales, Sales[Quantity] \* RELATED(Product[Price])). Formula of Totalsales = Numberofitemssold X Salesprice.

**Explication: (j)**

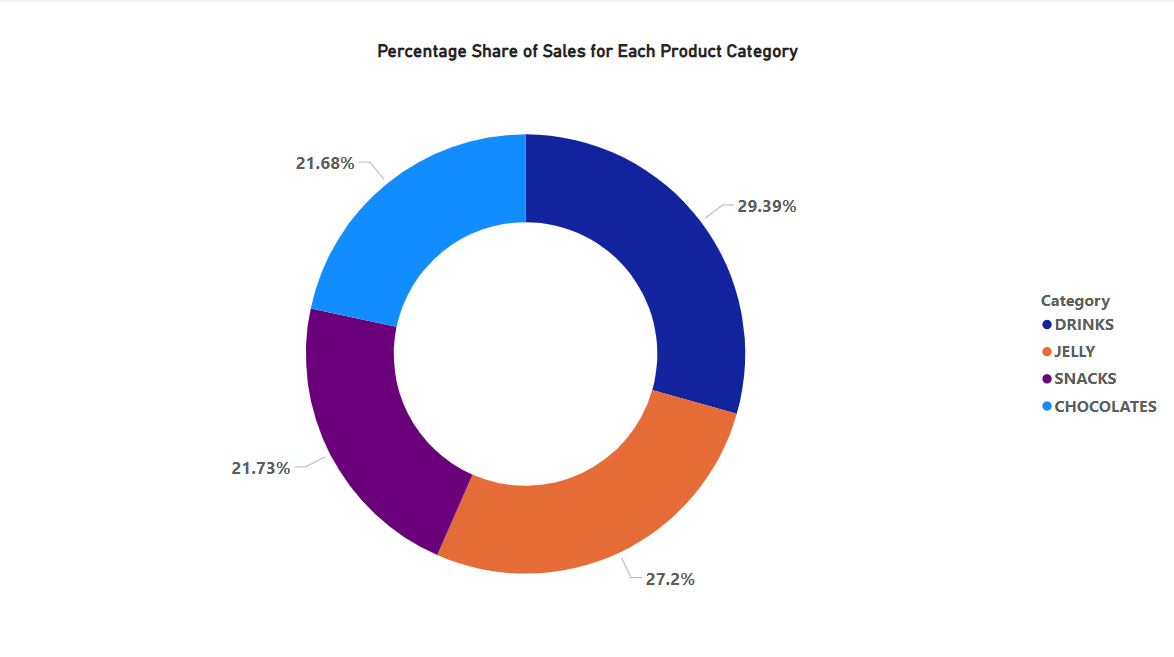
1. For Visualizations, selected Clustered Column Chart.
2. Drag Category from the Product table into the X-Axis field.
3. Drag the newly created TotalSales measure into the Y-Axis field from sales table.
4. Click on add to your visual🡪size and style🡪Columns🡪Customize the colors
5. click on Format 🡪Select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

A graph of sales by category

AI-generated content may be incorrect.

**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**