**Supermarket sales**

Historical record of sales data in 3 different supermarkets

**Team #70**

|  |  |
| --- | --- |
| **20201700506** | **علي ماهر طلعت السيد** |
| **20201701115** | **كريمة صبحي ابو العلا عبد الخالق** |
| **20201701030** | **يوسف محمد توفيق** |
| **20201701025** | **يوسف عمرو عطيه عبدالله** |
| **20201701034** | **يوسف محمد غريب احمد** |

## **About Dataset**

The dataset is one of the historical sales of supermarket company which has recorded in **3** different branches for **3** months data.

## **Attribute information:**

Dataset of **17** columns and **1000** rows.

**Invoice id**: Computer generated sales slip invoice identification number  
  
**Branch:** Branch of supercenter (3 branches are available identified by A, B and C).  
  
**City:** Location of supercenters  
  
**Customer type:** Type of customers, recorded by Members for customers using member card and Normal for without member card.

**Gender:** Gender type of customer  
  
**Product line:** General item categorization groups - Electronic accessories, Fashion accessories, Food and beverages, Health and beauty, Home and lifestyle, Sports and travel  
  
**Unit price:** Price of each product in $  
  
**Quantity:** Number of products purchased by customer  
  
**Tax:** 5% tax fee for customer buying  
  
**Total:** Total price including tax  
  
**Date:** Date of purchase (Record available from January 2019 to March 2019)  
  
**Time:** Purchase time (10am to 9pm)  
  
**Payment:** Payment used by customer for purchase (3 methods are available – Cash, Credit card and Ewallet)  
  
**COGS:** Cost of goods sold  
  
**Gross margin percentage**: Gross margin percentage  
  
**Gross income**: Gross income  
  
**Rating:** Customer stratification rating on their overall shopping experience (On a scale of 1 to 10)

**[i]Using Suitable Statistics:**

**# ======== DATA PREPROCESSING and STATISTICS ==============**

**[1] Check if there are any duplicated rows.**

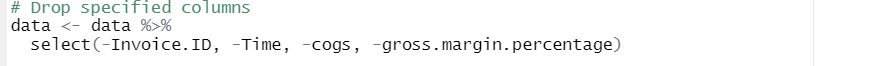
**🡺 there are no duplicated rows in our data**

**[2] extract column month from the Date column**

**🡺 to be used for visualization later**

**🡺 This line creates a new column 'Month' by extracting the month abbreviation from the 'Date' column.**

**[3] Drop specified columns.**

**🡺 drop columns that won’t be used in our project.**

**[4] View unique values in 'Branch' , 'Customer type' and 'Gender'**

**A computer code with text

Description automatically generated with medium confidence**

**🡺 This section prints the unique values in the specified columns.**

**There is 3 branches A , B and C , 2 customer types Member and Normal., And 2 genders male and female.**

**[5]Summary statistics for numeric columns**

**A close-up of several black text

Description automatically generated**

**🡺 provides summary statistics for the numeric columns in the dataset.**

**[6] check nulls**

****

**🡺 there is no nulls in our data**

**[7] get columns info.**

****

**A white background with black text

Description automatically generated**

**🡺 The columns Branch, City , Customer type, Gender, product line, Date, Payment and Month are character types.**

**Unit price , total ,tax , gross income and rating are number type**

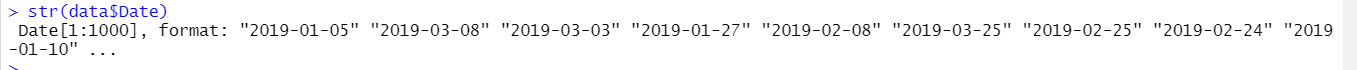
**Quantity is integer type.**

**[8] Convert Date Column to Date Type:**

**A close-up of a word

Description automatically generated**

**🡺 convert date type to Date and check it’s type again.**

****

**[9] Display the first few rows of the data**

**A close-up of a list

Description automatically generated**

**[ii] Dashboard instructions** **on how to use the dashboard effectively:**

Welcome to the Supermarket Sales Dashboard! This documentation provides step-by-step instructions on how to use the various features and functionalities of the dashboard.

**[1] Category and Comparison tab**

**1. Select a Category:**

- Use the dropdown menu to select a category (`Product.line`, `Gender`, `Branch`, `Payment`, `Customer.type`).

**2. Switch Between Units:**

- Use radio buttons to choose between "Absolute Values" and "Percentage Changes" for units.

**3. Select Categories for Comparison:**

- Use checkboxes to select categories for comparison (`Rating`, `Unit.price`, `Quantity`, `Tax.5.`, `Total`, `gross.income`).

**4. View Box Plots:**

- Box plots for selected categories will be displayed based on your choices.

**5. Distribution Pie Chart:**

- Observe the distribution pie chart based on the selected category.

**6. Customize Analysis:**

- Experiment with different category selections and units to customize your analysis.

**[2] Date and Mode tab**

**1. Choose a Category:**

- Select a category from the dropdown menu.

**2. Select Branches, Date Range, and Data Mode:**

- Choose branches, a date range, and between daily and cumulative data modes.

**3. View Scatter Plots and Histograms:**

- Observe scatter plots and histograms for the selected data.

**4. Customize Analysis:**

- Adjust visualizations based on your preferences.

**5. Toggle Between Daily and Cumulative Modes:**

- Switch between "Daily" and "Cumulative" modes for date histograms.

**[3] Visualization Options tab**

1. **Select Product Line and Customer Type:**

Use dropdown menus to choose a specific product line and customer type.

1. **Select Visualizations to Display:**

Check the checkboxes in the group labeled "Select Visualizations to Display" to dynamically show/hide different visualizations.

Use checkboxes to choose visualizations (`Line Chart`, `Histogram`, `Heatmap`, `Scatter Total Gender Plot`, `Combined Plot`).

1. **Explore Visualizations:**

View line charts, histograms, heatmaps, scatter plots, and combined plots based on your selections.

**4. Clear Selected Visualizations:**

- Uncheck checkboxes to clear selected visualizations.

**[4] City Dropdown Tab:**

**1. Select City:**

Choose a city from the dropdown menu labeled "Select City."

**2. View Bar Plot:**

Observe the bar plot for the selected city.

**[5] Dark/Light Mode**

**1. Toggle Between Dark and Light Mode:**

- Use the "Toggle Dark Mode" button to switch between dark and light modes.

**2. Adjust Styles:**

- Styles such as background color, text color, and title color will change based on the selected mode.

**3. Enhance Dashboard Appearance:**

- Toggle the mode to match your working environment and enhance dashboard appearance.

**[6] Tips and Notes**

**1. Hover Over Elements:**

- Hover over different elements in plots for additional information.

**2. Experiment with Combinations:**

- Experiment with different combinations of selections for a more in-depth analysis.

**3. 6. How to Use tab:**

- a tab Include clear and concise tooltips or help sections for each Visualization.

**4. Explore Dataset:**

- Explore the dataset from various angles to gain insights.

**5. Have Fun Exploring!**

- Enjoy exploring the Supermarket Sales data!

**[iii] insights about each visualization**

A screenshot of a computer

Description automatically generated**1 - Pie chart**

The percentage of electronic accessories , fashion accessories and food and beverage products line seem to be close to each other and have higher percentage rather than other products line while health and beauty product line has lowest percentage.

**2 - Box Plot**

A screenshot of a graph

Description automatically generatedThe average rating received of the supermarkets are 7 and minimum rate is 4 while maximum is 10.

**3 - Line Chart**

A graph of a graph

Description automatically generated

The total gross income is increasing within last 3 months.

9 mars 2019 is the highest day have gross income.

**4 – Histogram**

A graph of a bar graph

Description automatically generated with medium confidence

The highest percentage of people, which are 174 gave rating 6 out 10.

**5 - Histogram with multiple attributes and interactive**

A screenshot of a graph

Description automatically generated

Most of the time equal percentage of normal and member customers buy from health and beauty product line.

**6 – Scatter plot**

A graph with blue and red lines

Description automatically generated

The females and male customers of the supermarkets are equal .

**7- Combined Plot**

A graph of a number of bars

Description automatically generated with medium confidencein last three months people bought from electronic and accessories , health and beauty, Home and lifestyle product line most in mars.

, fashion and accessories, sports and travel product line in Jan and from food and beverage product line in February .

**8- Scatter plot**

A screenshot of a computer screen

Description automatically generated

We can see that branch a and b are most widespread and sales than branch c.

**9 - Histogram day mode**

A graph of a graph

Description automatically generated with medium confidence

We noticed that February is the month with the highest sales.

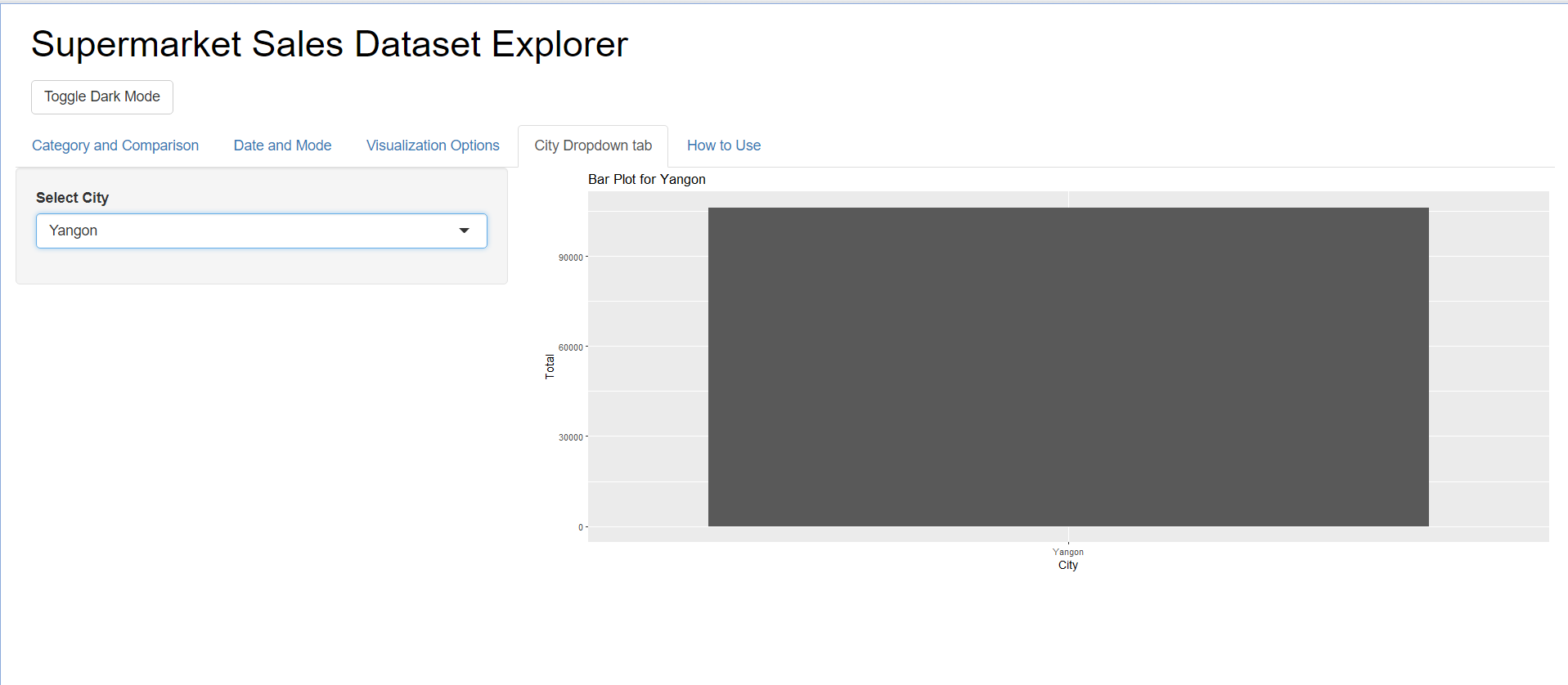
**10 - Histogram cumulative mode**

A graph of a graph

Description automatically generated

Here we can see that Profits are steadily and consistently increasing.

**11 – bar plot**



Here we can see that the city of Yargon is more widespread than other cities.