

About the Dataset :

This dataset contains historical sales data from a supermarket company. The data includes records from three different branches over a three-month period.

Data Description :

1. Invoice ID: Automatically generated identification number for sales slips.
2. Branch: Supercenter branch (identified as A, B, or C).
3. City: Location of the supercenters.
4. Customer Type: Type of customers, categorized as "Members" for those using a member card and "Normal" for those without.
5. Gender: Gender of the customer.
6. Product Line: Categorization of general items, including Electronic Accessories, Fashion Accessories, Food and Beverages, Health and Beauty, Home and Lifestyle, and Sports and Travel.
7. Unit Price: Price of each product in dollars (\$).
8. Quantity: Number of products purchased by the customer.
9. Tax: 5% tax fee for customers.
10. Total: Total price including tax.
11. Date: Purchase date (Recorded from January 2019 to March 2019).
12. Time: Purchase time (from 10 AM to 9 PM).
13. Payment: Payment method used by the customer (Cash, Credit Card, or Ewallet).
14. COGS: Cost of goods sold.
15. Gross Margin Percentage: Gross margin percentage.
16. Gross Income: Gross income.
17. Rating: Customer satisfaction rating based on their overall shopping experience (rated on a scale of 1 to 10)

In Mind Questions

- General Questions related to the existence of
 - missing values?
 - wrong datatypes for columns?
 - complete duplicates in the data?
 - outliers in each column?
- Univariate Analysis
 - Which branch has largest Sales numbers?
 - Which Gender is greatly come to branches?
 - What is the best-selling product line in the branches?
 - Which Payment Customers preferred?
 - Which Customer type comes to branches greater?
- Bivariate Questions
 - Which branch has the highest gross income?
 - There is relationship between gross income and customer rating?
 - What is gender who come to branches more?
 - What is male and female gross income in each branch?
 - Which product line in branches have greater gross income?
 - What is relationship between customer type, branches and gross income?
 - Which product line preferred for every gender?
 - Which product line have greater gross income?
 - What is Month has greater gross income?
 - What is greater sales product line in each month?

Conclusion

- Reached Results from Univariate Analysis
 - Branche A has largest Sales numbers.
 - Most Gender Come to branches is female.
 - Most sales in branches is Fashion accessories.
 - Customers preferred to Ewallet, cash payments rather than credit card.
 - Most Customers come to branches has member card.
- Reached Results from Bivariate Questions
 - Branch C stands out slightly with higher income compared to Branch A and B. Despite Branch A having slightly higher sales, it is Branch C that emerges as the most profitable branch in terms of gross income.
 - No relationship between rating and gross income.
 - Branch A, B has males greater than females, but branch C has females greater than males.
 - Female gross income greater than Male in each branch.

- Electronic accessories, Home and lifestyle is the most sales in branch A, Health and beauty, Sports and travel is the most sales in branch B and - Food and beverages, Fashion accessories is the most sales in branch C.
- Branch A and B, Normal customer greater than Member but Branch C, Member customer greater than Normal, When customer number of members in Branch increased, Total of gross income increased.
- When number of females increase, Fashion accessories Product line increase but When number of male increase, Health and beauty Product line increase.
- Gross income is highest in food and beverages.
- January month has greater gross income.
- January month Sports and travel, Fashion accessories is the most product line sales, March month Electronic accessories, Home and lifestyle is the most product line sales and February month Food and beverages, Fashion accessories is the most product line sales.

Summary

- Customer average rating of approximately 7. there was no relationship between customer ratings and gross income
- Among the three branches, Branch C as the most profitable in terms of gross income, despite Branch A having slightly higher sales
- Fashion Accessories and Food and Beverages were the most products. Therefore, must focus on these categories, with electronic accessories
- The preferred payment method for customers was found to be Ewallet and Cash
- Food and Beverage proved to be the category generating the highest gross income overall
- Females spend the most on fashion accessories, while males preference for health and beauty products Females also spending on sports and travel, which contributed to the highest overall income
- In the end, there was no relationship between customer ratings were not significantly related to any variable