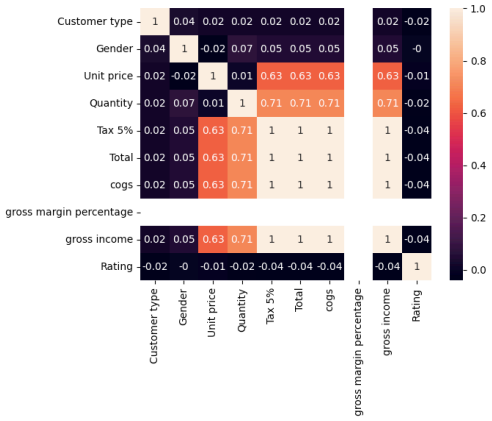
Supermarket Sales

Report

Introduction

The data given to us shows the invoices of 3 supermarket stores in Myanmar, Yangon, Naypyitaw / Naypyidaw, Mandalay referred to as A,B and C respectively for the rest of the report. This data been collected between 1st January 2019 and 3rd March 2019.

The first step in our analysis will be to clean this data, removing the NA values and performing the necessary transformations. I have covered analysis based on gender and time in this report.



The heatmap above represents the correlation of the dataset. We can see that the tax, total, cogs and gross income are completely correlated with each other, this is because they are derived quantities. Gross margin percentage is just a constant value throughout.

Data cleaning

The data has no NA or missing values that need to be accounted for and appears to be clean. Few unnecessary columns like Invoice ID and city name were dropped. We can also drop total, cogs, and gross margin percentage for the rest of the analysis and just keep gross income.

Transformations

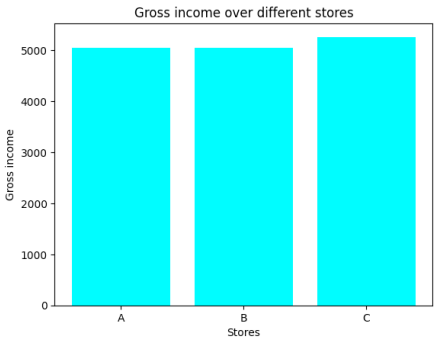
The date and time had been read as ‘object’ data type and have been converted to their proper format. The columns gender and customer type have been converted to binary variables. For gender 0 represents male and 1 represents female. For customer type 0 represents non-member while 1 represents members.

Exploratory Data Analysis

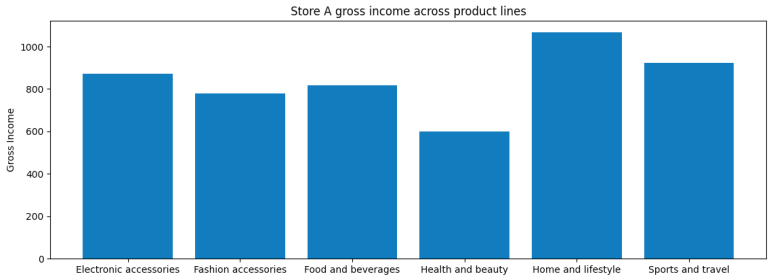
Store Based Analysis

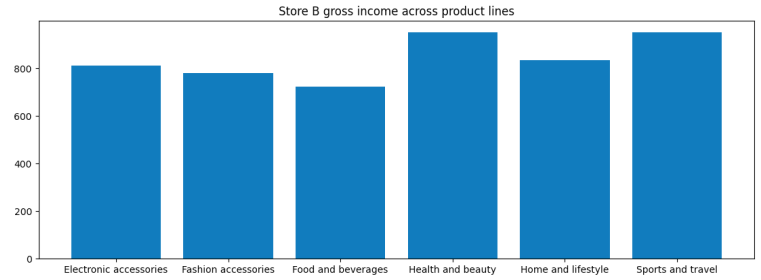
There are a 1000 invoices in the data frame. Store A has the most with 340 invoices followed by store B with 332 and store C has the least invoices 328. The difference however is not too significant.

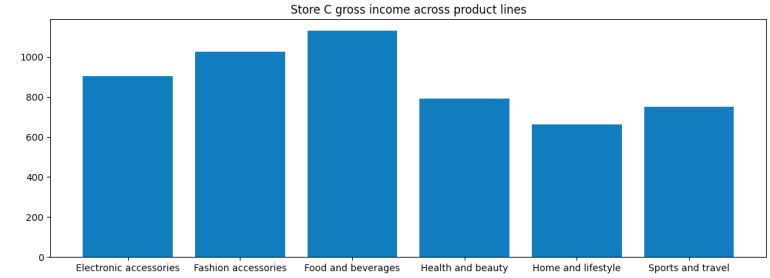
The average rating of customer experience across all stores is 6.97/10. This can be used as a metric to identify which stores are performing best in terms of customer satisfaction. Store A has an average rating of 7.02, store B is 6.81 and store C is 7.07. So we can see that store A is the best followed by store C and then store B.

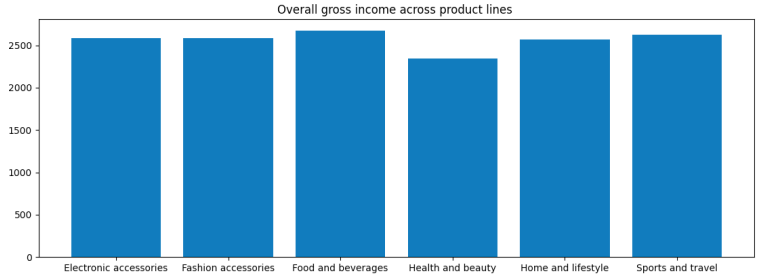


We can see that Store A sells the least health and beauty products and Store B sells the most. Health and beauty products are the least sold category overall. Home and lifestyle is the least sold product category at store C and the most sold product category at store A. Overall food and beverages sells the most. The number of invoices show that store C is the least popular but the gross income graph shows store C is making the most money.



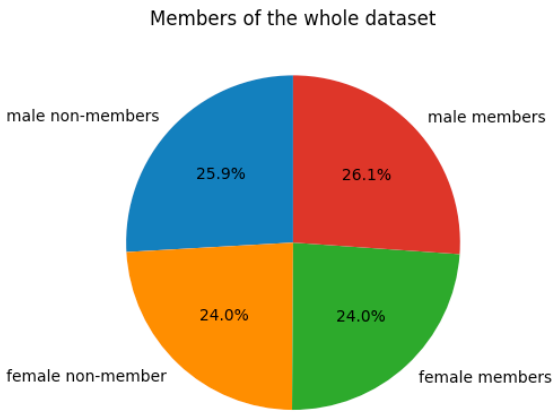






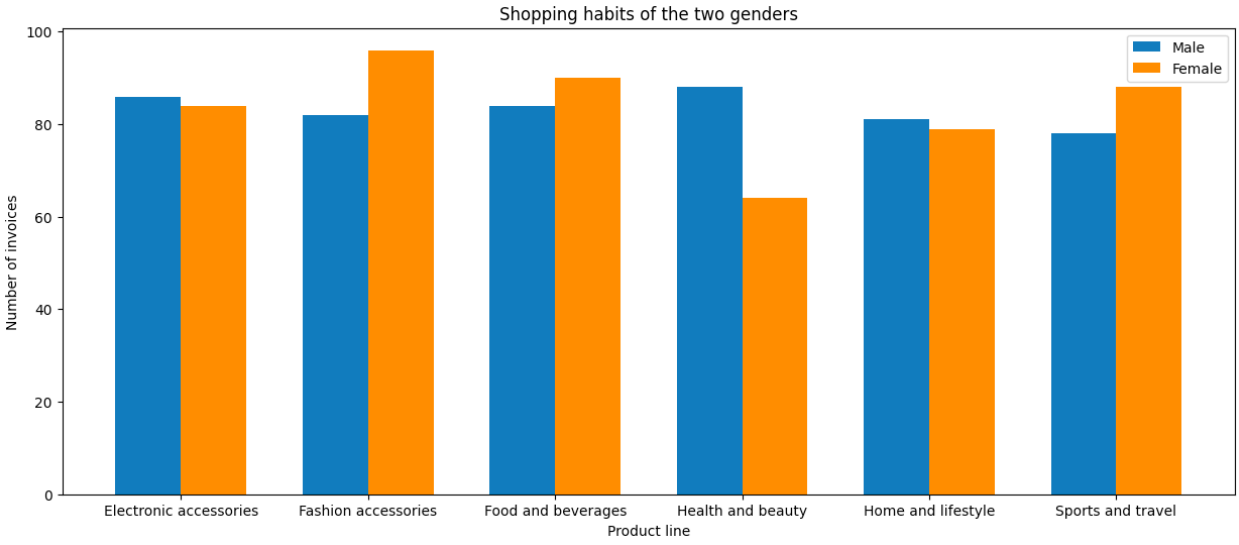
Gender Based Analysis

There 501 females shopping across all three stores and 499 males so the dataset is spread even among shoppers of both genders. The pie chart shows the distribution of male and female shoppers who are members/non-members of the supermarket chain.



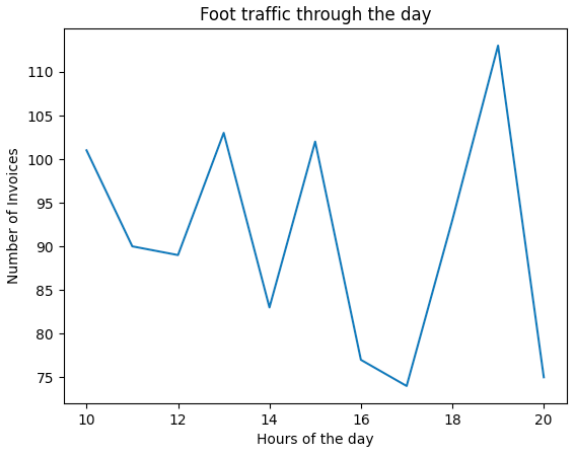
The grouped bar graph shows the total number of purchases made by the two genders on each product line. Key insights here are

1. Men spend more on health and beauty
2. Men spend more on health and lifestyle
3. Women spend more on sports and travel



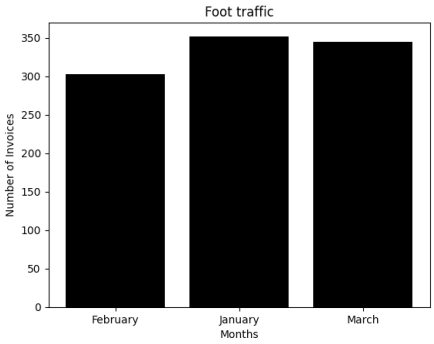
Analysis based on time

The dataset spreads 3 months in duration. From this we can do several kinds of analysis. We can see that based on days of week, all stores see peak traffic on Saturdays after which traffic falls to it's lowest till Tuesday when it gets back up.



From the above plot we can see that the stores run from 10am to 8pm. The foot traffic in the stores dies down right before it reaches its peak. The peak operation hours for the stores is between 5 and 8 PM.

The bar graph on the right shows that Feburary saw this least amount of traffic between the three months and Janurary saw the most.



Insights

1. Store Performance:

- Store A has the highest customer traffic, indicating its popularity among shoppers. However, store C generates the highest gross income, suggesting that it focuses on higher-priced items or has higher transaction values. Store B falls in between in terms of both customer traffic and gross income. This indicates potential opportunities for stores A and B to improve their profitability by analyzing factors contributing to store C's success.

2. Customer Experience:

- Store A receives the highest average rating for customer experience, indicating that customers are more satisfied with their shopping experience at this store. This suggests that store A excels in areas such as customer service, store layout, or product availability. Store C also receives a relatively high average rating, while store B has the lowest average rating.

Store B should focus on identifying the reasons behind lower customer satisfaction and take measures to improve its performance.

3. Product Category Analysis:

- Store B sells the most health and beauty products, indicating a demand for such items among customers. Stores A and C can capitalize on this opportunity by expanding their offerings in the health and beauty category to attract more customers and increase sales.

- Store A sells the most home and lifestyle products, indicating a potential focus area where it can further enhance its offerings and attract more customers.

- Food and beverages are the most sold product category across all stores, emphasizing its significance in driving sales. Stores should continue to prioritize this category and optimize their inventory and marketing strategies to maximize profitability.

4. Gender-Based Analysis:

- Men tend to spend more on health, beauty, and lifestyle products, suggesting an untapped market opportunity. Stores can introduce a wider range of products in these categories targeted towards men to capture their purchasing power and drive additional sales.

- Women demonstrate a higher spending inclination towards sports and travel. Stores can tailor marketing campaigns, promotions, or product displays to cater to women's preferences in these categories and increase their sales.

5. Analysis Based on Time:

- Tuesdays and Saturdays between 5 and 8 PM experience the highest foot traffic. This information can guide stores in managing their staffing resources effectively during peak hours to ensure a satisfactory shopping experience for customers.

- January emerges as the most profitable month for all stores, while February is the least profitable. Stores can leverage this insight by implementing targeted promotions or special offers during January and strategizing to boost sales during February through initiatives like seasonal promotions or collaborations.