ÇUKUROVA UNIVERSITY
DEPARTMENT OF COMPUTER ENGINEERING

# VISULATION OF SUPERMARKET SALES DATA

ALİ CAN SARIBOĞA - 2019556055

## Supermarket Sales

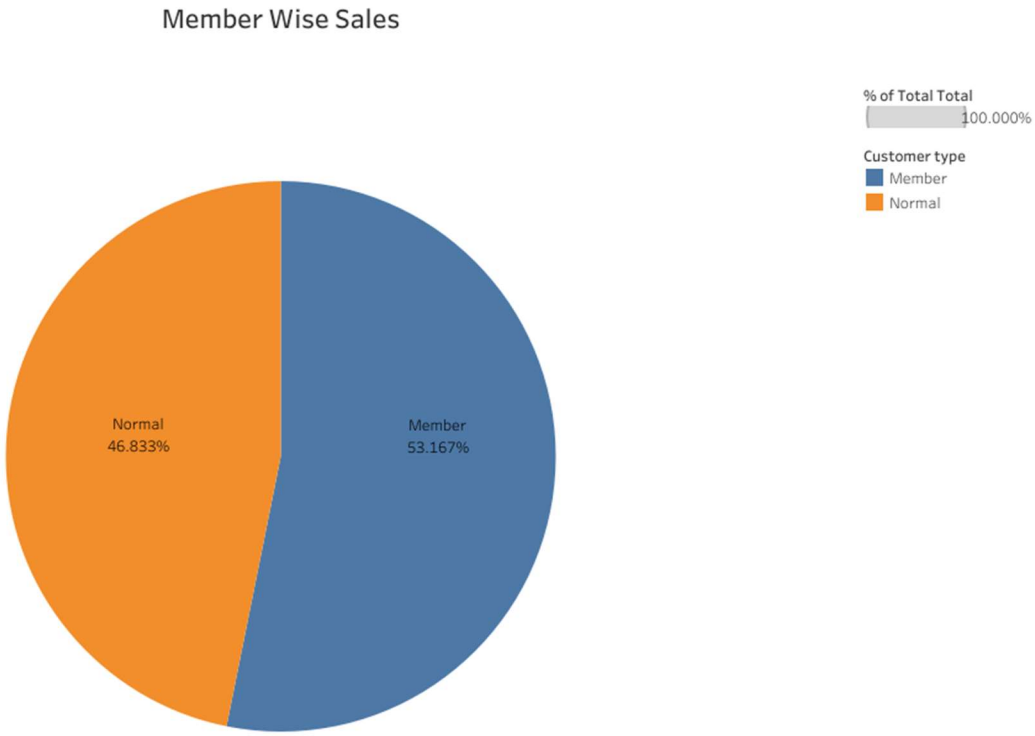
*Context*

The growth of supermarkets in most populated cities are increasing and market competitions are also high. The dataset is one of the historical sales of supermarket company which has recorded in 3 different branches for 3 months data. Predictive data analytics methods are easy to apply with this dataset.

*Attribute information*

| |
|---|
| 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). |

**Question 1 - What is the ratio of total sales to member and non-member customers?**



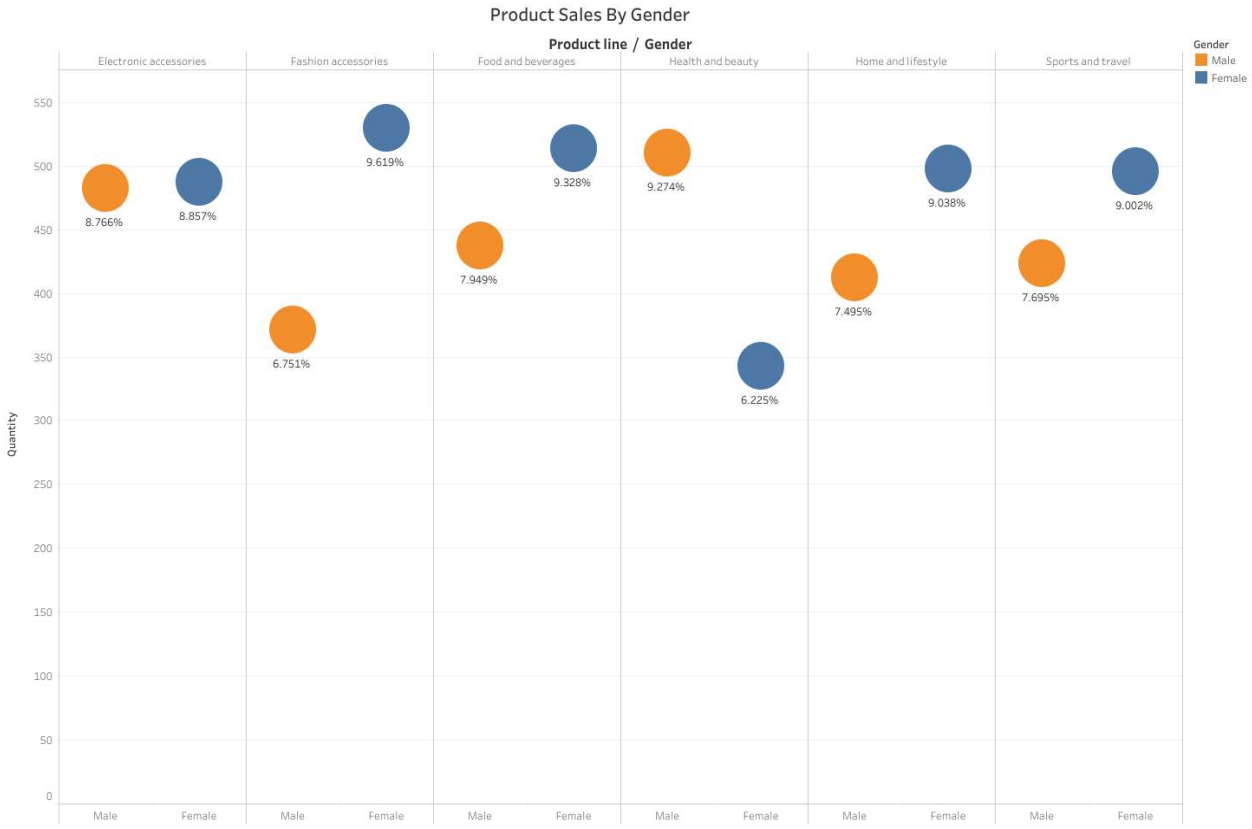

I have represented the shopping rate of customers in the supermarket using a pie chart. The chart is divided into two segments, each corresponding to a different membership status – members and non-members.

The larger segment of the pie chart represents the shopping rate of members, accounting for 53.17% of the total. This indicates that a significant portion of the customers in the supermarket are members.

The smaller segment of the pie chart represents the shopping rate of non-members, making up 46.83% of the total. While slightly less than the members, non-members still constitute a substantial portion of the customer base.

The pie chart illustrates the distribution of shopping rates between members and non-members. The larger share occupied by members suggests that a considerable percentage of customers are taking advantage of membership benefits, possibly indicating a positive response to the supermarket's membership program.

**Question 2 - What is the gender distribution of the number of products sold by product category?**



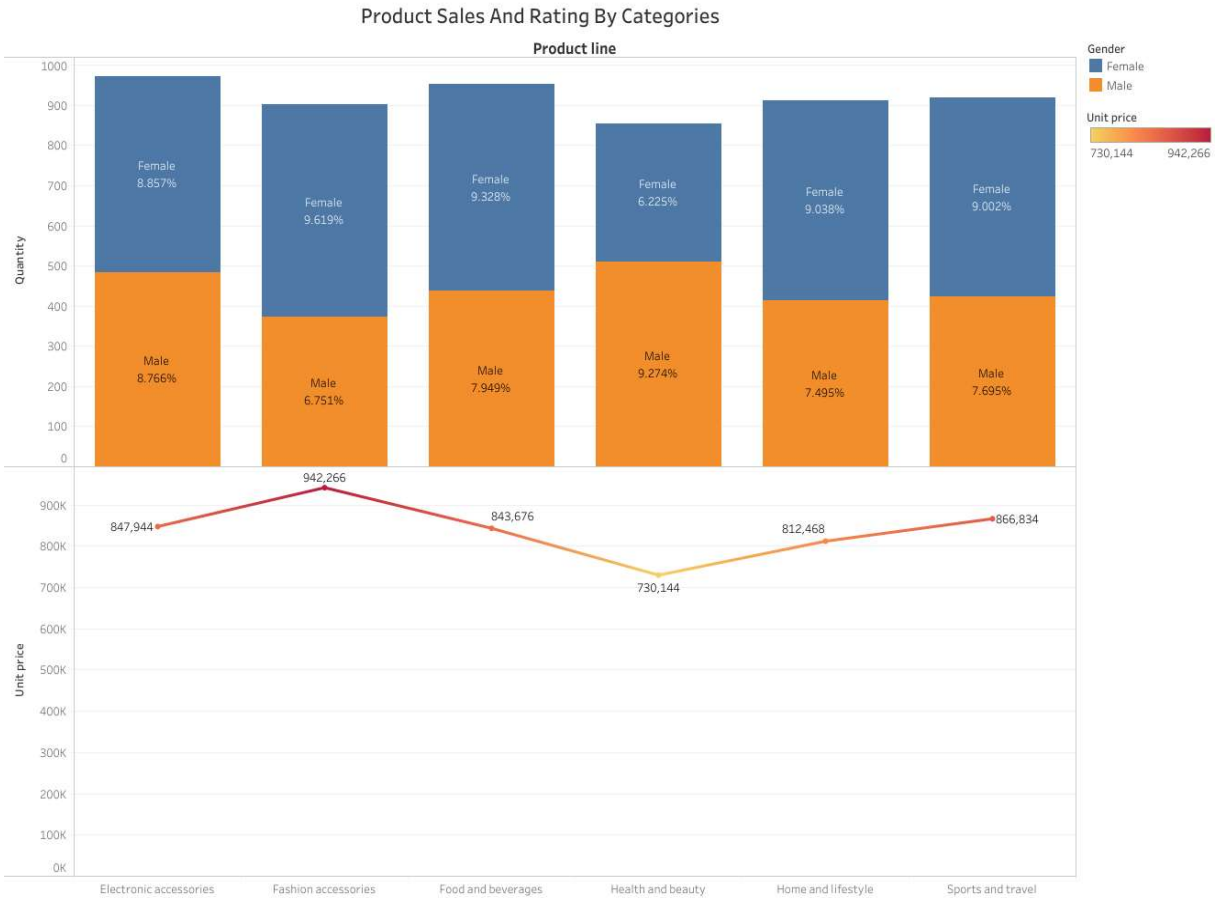

I used a scatter plot chart to visualize the ratio of the total number of products sold based on product categories and the genders of the shoppers. The chart represents product categories on the x-axis and the ratio of the total number of products sold on the y-axis.

The chart generally indicates that men have a higher shopping ratio in product categories compared to women. The overall trend of the points on the scatter plot suggests that men tend to shop more in many categories.

However, there is a notable exception. In the health and beauty category, women's shopping ratio is higher than that of men. This category exhibits a different trend from the others, and it is represented by a specific point on the scatter plot.

This graph not only illustrates general trends but also highlights gender-specific shopping patterns in specific categories. While men tend to dominate in overall shopping, the health and beauty category stands out with a higher interest from women. This information could be valuable when making strategic decisions, such as in marketing or inventory planning.

**Question 2.1 - What is the gender distribution of the number of products sold by product category? Do product prices have an impact on this distribution?**




Bar chart represents the quantity of each product category sold, displayed as a bar chart. Here are the comments related to this chart:

The category with the highest sales among women has a higher value compared to other categories. This indicates that a specific product category is popular among women.

The category with the highest sales among men has a lower quantity compared to the category for women. This suggests a marketing strategy with a focus on women's products.

Line chart displays the unit price of each product category as a line graph. Here are the comments related to this chart:

The category with the highest sales among women has a higher unit price compared to other categories. This implies that the popular category among women is associated with a higher price.

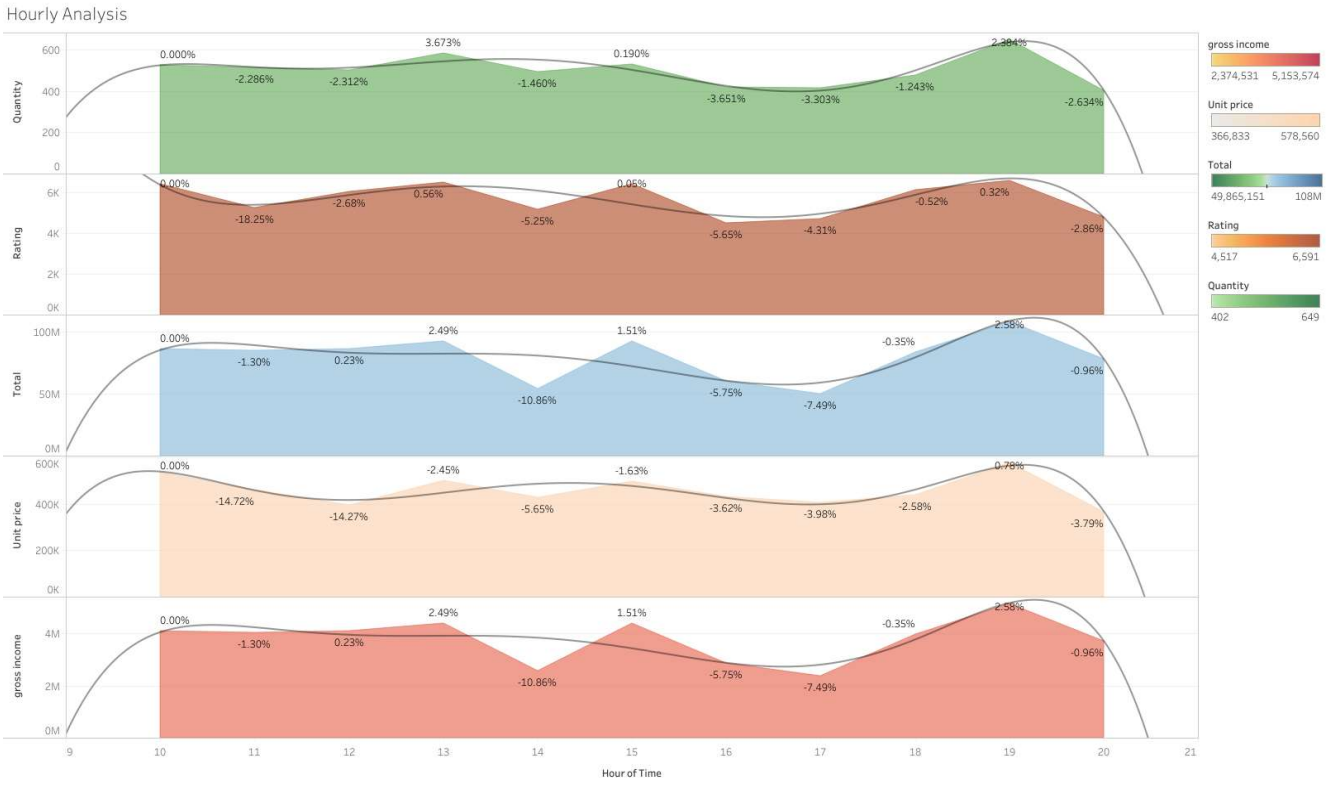
The category with the highest sales among men has a lower unit price compared to other categories. This indicates that the popular category among men is priced more affordably.

The popular category among women may involve premium products, given its higher unit price. This suggests a marketing strategy emphasizing premium products among women.

The popular category among men may involve more budget-friendly products, given its lower unit price. This implies a strategy focused on price-sensitive male customers.

These comments highlight that the marketing strategy is tailored to specific gender groups, with different pricing approaches. With this information, the store management can further optimize their sales strategy, adjust inventory based on demand, and enhance customer satisfaction.

**Question 3 - How do quantity, rating, total, unit price and gross revenue values change depending on the time of day?**




To analyze shopping patterns throughout the day, separate charts were created for quantity, rating, total, unit price, and gross income. These charts explored hourly data from 10 AM to 8 PM. Applying a compound growth rate to the hourly data, the values at 10 AM were set as a reference point, and subsequent values were observed relative to this reference. This method was employed to thoroughly understand shopping trends during different hours of the day. The notable peak in shopping activity at 7 PM suggests that customers tend to engage in more significant shopping activities in the evening. Conversely, the observed lowest shopping

volume, predominantly at 5 PM, may indicate a period of potentially lower demand or intensity.

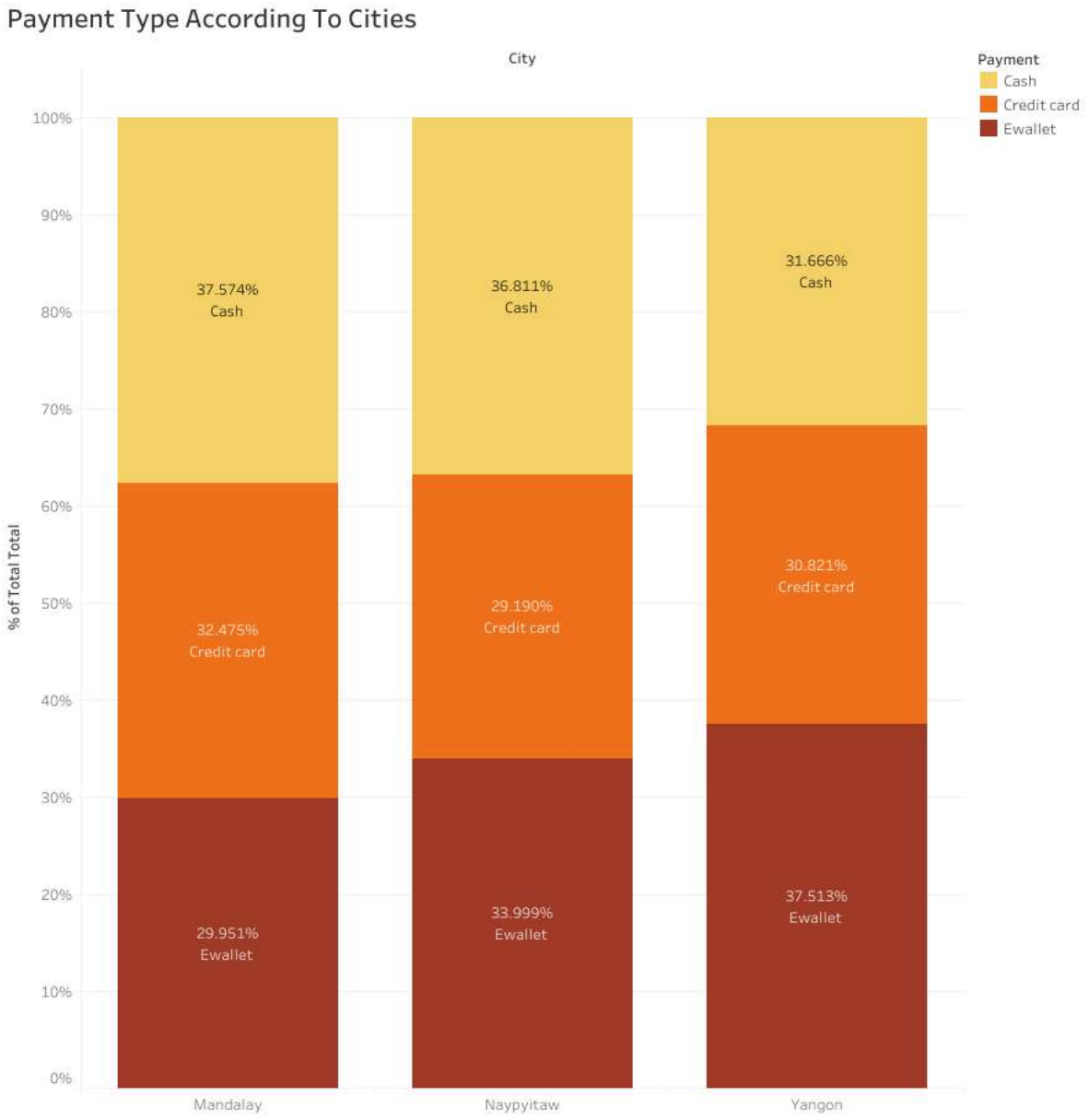
The hourly charts vividly illustrate the dynamics of shopping behaviors throughout the day.

Using compound growth rate with a reference point at 10 AM provides a more meaningful representation of changes during other hours.

The substantial peak in shopping volume at 7 PM indicates heightened shopping activities towards the end of the day.

The observed lowest shopping volume, largely at 5 PM, may point to a period of potentially lower demand or intensity.

**Question 4 - How do payment methods for sold products vary by city?**



The stacked bar chart is an effective choice in this case to illustrate the breakdown of total sales by payment methods across different cities. This chart type clearly shows the contribution of different payment methods to total sales and how these contributions are distributed among the cities of Mandalay, Naypyitaw, and Yangon.

This stacked bar chart provides valuable insights by showcasing the distribution of different payment methods across cities. Here are some key observations:

Mandalay:

The highest sales in Mandalay are made through cash payments. This suggests a tendency among customers in Mandalay to use cash. Ewallet payments remain at the lowest level, indicating a potentially lower demand for digital payment methods.

Naypyitaw:

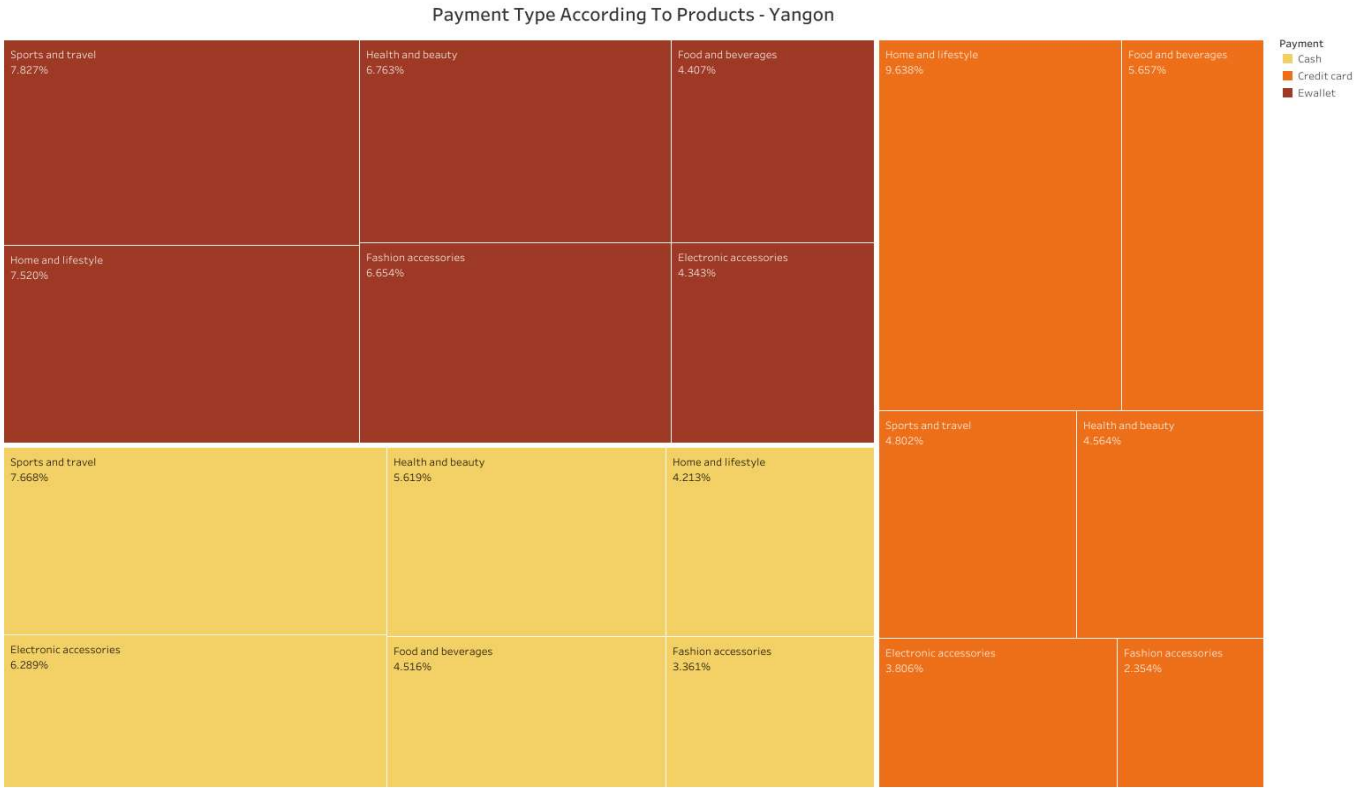
In Naypyitaw, the highest sales are made through cash payments, and credit card payments are at their lowest level. This indicates the shopping trends and payment preferences in Naypyitaw.

Yangon:

In Yangon, the highest sales are through ewallet payments, while credit card payments are at their lowest level. This suggests that customers in Yangon are more inclined to adopt and prefer digital payment methods.

This chart type, by highlighting the differences in payment methods across cities, provides a valuable perspective for making strategic decisions for the business.

**Question 4.1 - What is the variation and percentage of payment methods according to products in Yangon?**



In this treemap graph, the total products sold in Yangon are visualized based on payment methods and product categories. The treemap graph efficiently represents the payment distribution within each product category. Payment methods are differentiated by three distinct colors, and the percentage representation of product varieties is indicated on the graph. Ewallet payments are most prominent in the sports and travel category, while they are least in the electronic accessories category. For cash payments, the sports and travel category leads, while the fashion accessories category has the least. Credit card payments are most prevalent in the home and lifestyle category and least in the fashion accessories category.

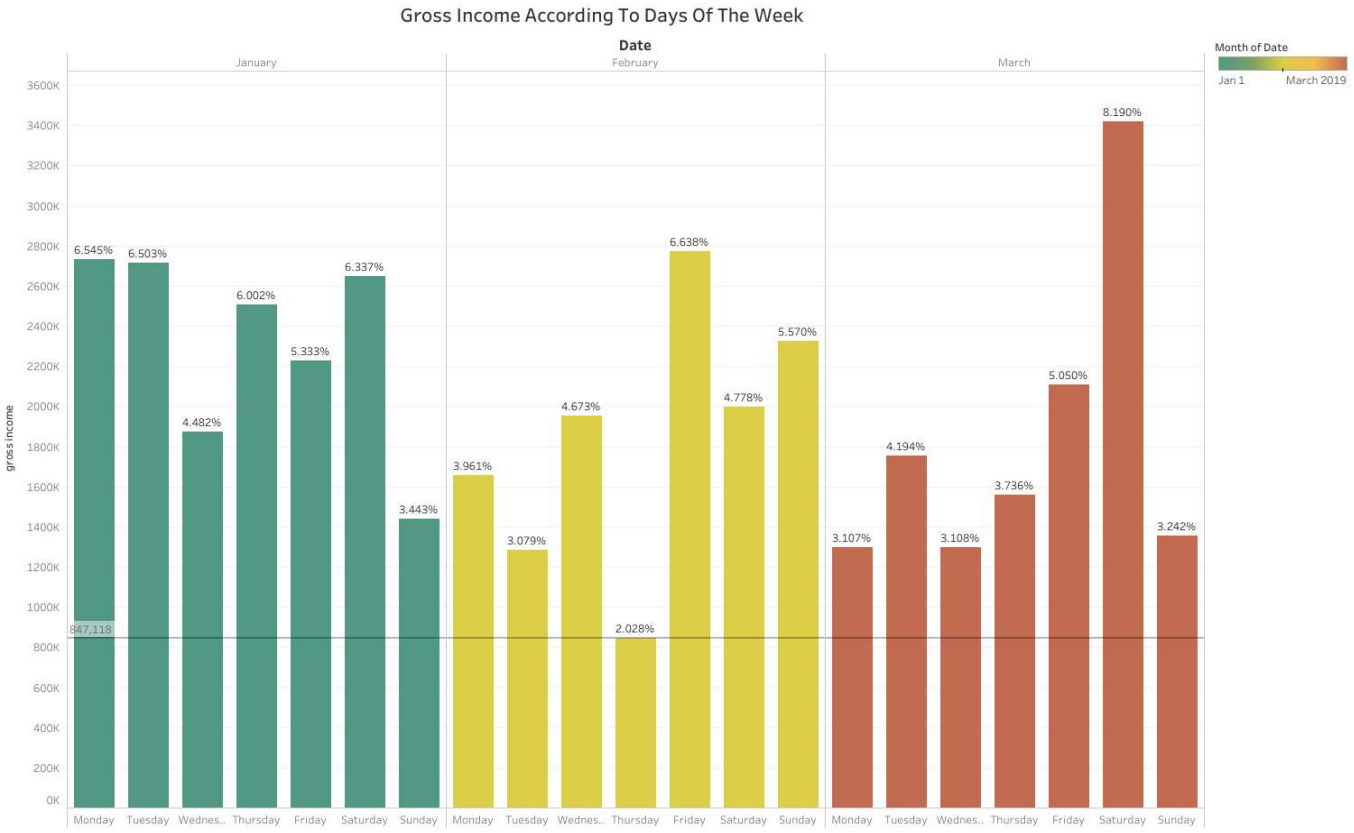
The treemap graph is chosen to present a hierarchical view of data, showing the proportion of each product category within the total products sold in Yangon. The color-coded payment methods provide a clear visual distinction, allowing for easy interpretation of the payment distribution within each category.

The treemap graph effectively communicates the distribution of payment methods within each product category in Yangon.

It highlights the variation in payment preferences, with different categories showing distinct patterns.

Understanding these patterns can inform business decisions, such as marketing strategies tailored to each category or adjustments in payment processing methods based on customer preferences.

**Question 5 - How does 90-day data change compared to the same days of that month? How does gross income vary by day?**



This graph utilizes a bar chart to visually represent the ratio of 90-day gross income data across weekdays for the months of January, February, and March. The bar chart serves as an effective tool to compare gross income ratios for each day of the week, offering a quick and insightful understanding of performance variations.

Factors Contributing to Differences:

Seasonal effects, holidays, promotions, and special events can impact the gross income ratios for each month.

Consumer shopping habits, preferences during weekdays and weekends, and other external factors may contribute to these variations.

Strategic Evaluation:

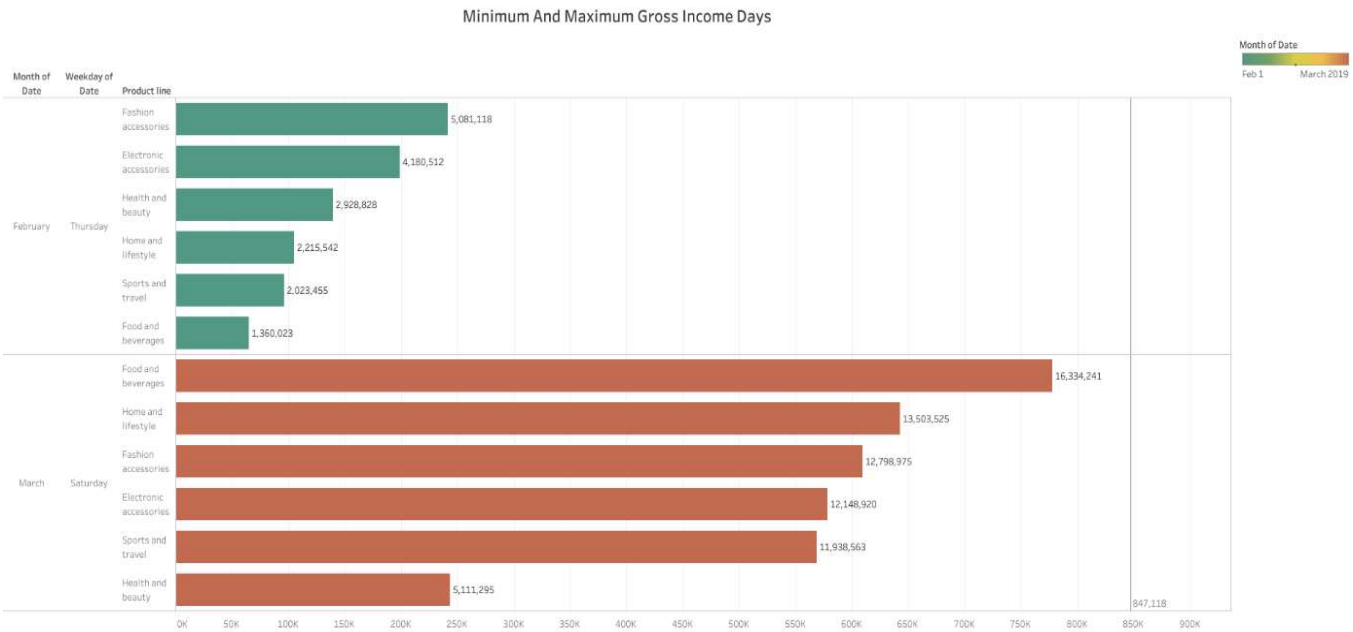
Marketing Strategies: Understanding peak sales days can inform marketing strategies, enabling the planning of targeted promotions or discounts on busy days.

Inventory and Staff Management: Optimizing inventory and staff management on high-traffic days is crucial for meeting demand and enhancing customer satisfaction.

Enhancements to Customer Experience: Improving customer experience on busy days through strategic improvements can lead to faster service and increased customer satisfaction.

Such analysis helps businesses gain insights into daily operations, enabling them to develop more effective strategies for improved performance.

**Question 5.1 - How does the gross income value on the lowest and highest days of that month change according to categories within the 90-day data?**



The graph employs a horizontal bar chart to assess 90-day gross income data across weekdays, categorizing sales to reveal insights into daily performance. This visual representation facilitates a quick understanding of gross income variations for each day.

The horizontal bar chart was selected for its effectiveness in comparing gross income values across days and categories. It visually presents the gross income percentages for each category, offering a comprehensive view.

Sales Strategy and Product Management:

February and March Sales Strategies:

February Thursdays: Focus on fashion accessories with targeted promotions, consider diversification in the food and beverages category.

March Saturdays: Highlight food and beverages, introduce promotions in health and beauty.

Product-Specific Strategies:

For Sales Growth: Encourage sales in the fashion accessories category.

For Maintaining Stability: Preserve diversity in the food and beverages category, consider promotions in health and beauty.

These strategies aim to optimize sales based on days and categories, responding effectively to customer demands.

*Ali Can SARIBOĞA*

*2019556055*