Supermarket Sales Data Report

Introduction:

Dataset Overview:

Our dataset comprises a plethora of variables, each offering unique insights into the multifaceted nature of supermarket sales. From fundamental transactional details such as Invoice ID, Date, Time, and Payment Method to more nuanced factors like Branch Location, Customer Type, Gender Demographics, Product Line, and Product Ratings, every facet has been meticulously documented.

Key Attributes:

1. Invoice ID: A unique identifier for each sales transaction, facilitating traceability and analysis.
2. Branch (A, B, C): The geographical location of the supermarket branch, allowing for regional comparisons and trend identification.
3. Customer Type (Normal, Member): Distinguishing between regular customers and members, offering insight into loyalty and engagement levels.
4. Gender (Male, Female): Demographic segmentation aiding in understanding purchasing preferences and patterns.
5. Product Line (Fashion Accessories, Electronic Accessories, Food and Beverages, Health and Beauty, Home and Lifestyle, Sports and Travel): Categorization of products facilitating analysis of sales trends across different product categories.
6. Unit Price, Quantity, Tax (5%): Fundamental transactional details crucial for revenue assessment and pricing strategies.
7. Payment Method (Credit Card, Cash, E-wallet): Reflecting evolving payment preferences and trends in consumer behavior.
8. Gross Margin Percentage, Gross Income, COGS: Performance metrics illuminating profitability and operational efficiency.
9. Rating (1 to 10): Customer feedback providing a qualitative assessment of product satisfaction and service quality.
10. City (Yangon, Mandalay, Naypyitaw): Regional segmentation enabling geographical analysis and market segmentation.

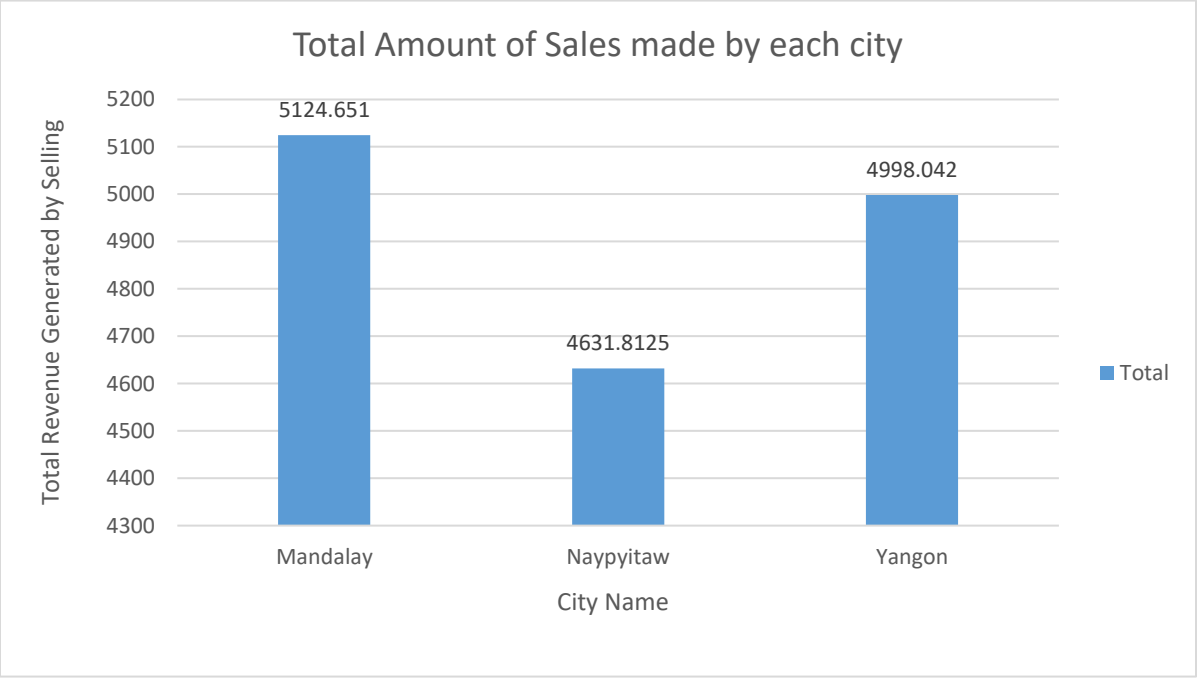
Questionnaire:

Q1. Which of the given cities having tax 5% slab performed better than all the others?
Q2. Which customer gender ordered most items from all the three branches?
Q3. Compare highest and lowest rating products on the basis of units sold.
Q4. Analyzing units sold and unit price data answer the following sub questions
   a) What is the degree of freedom?
   b) Co-relation of Unit price and revenue generated
   c) What result you can draw from regression of the two data

Q5. What product will you suggest as per the city data analysis to each type of customer

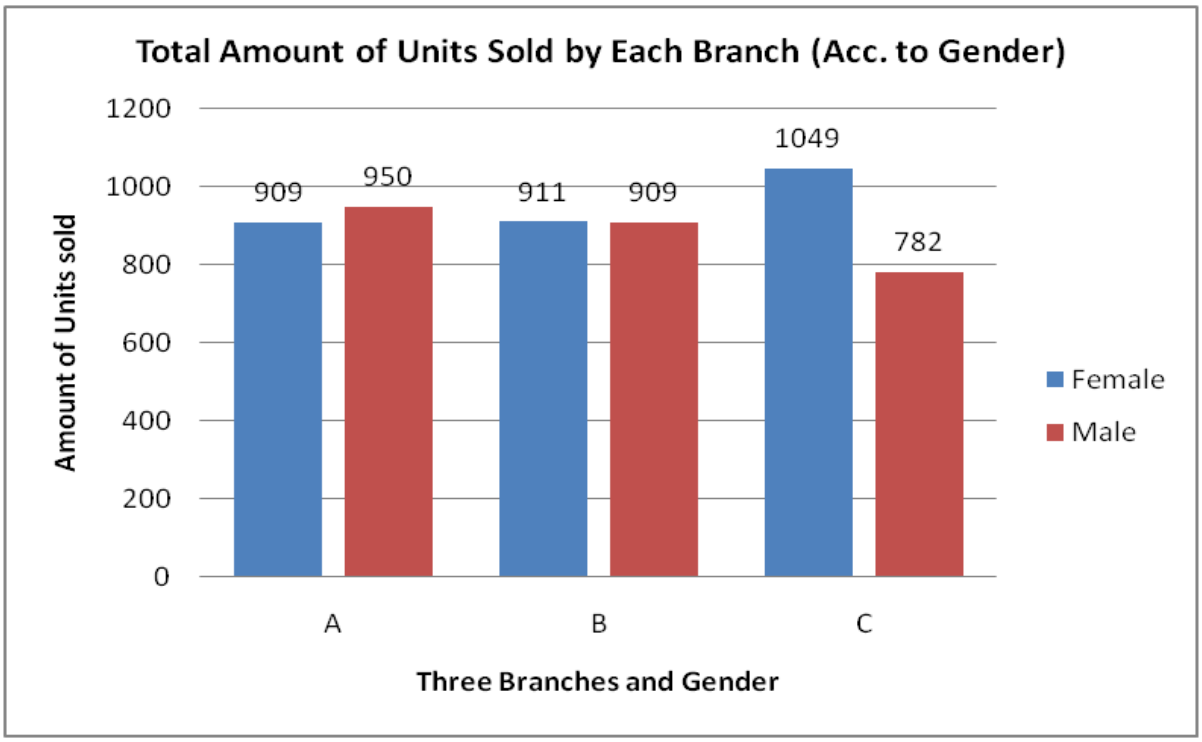
Analytics:

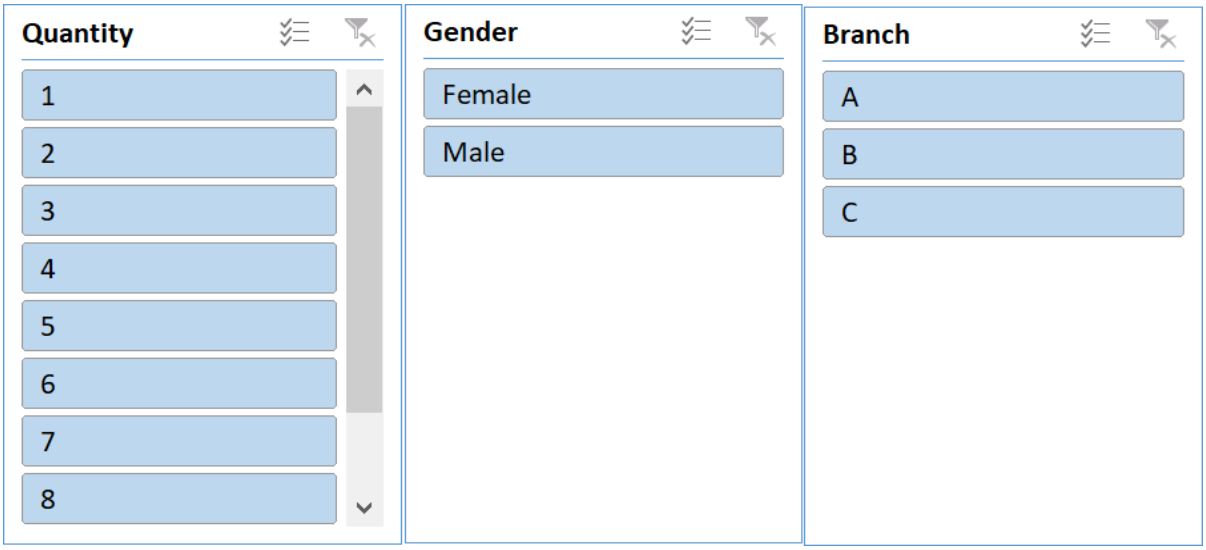
Q1. Which of the given cities having tax 5% slab performed better than all the others?



Based on the data analyzed, the city that outperformed all is **Mandalay**. This conclusion is drawn from superior performance in total sales/revenue generation compared to the other cities in the same tax slab of 5%.

Q2. Which customer gender ordered most items from all the three branches?

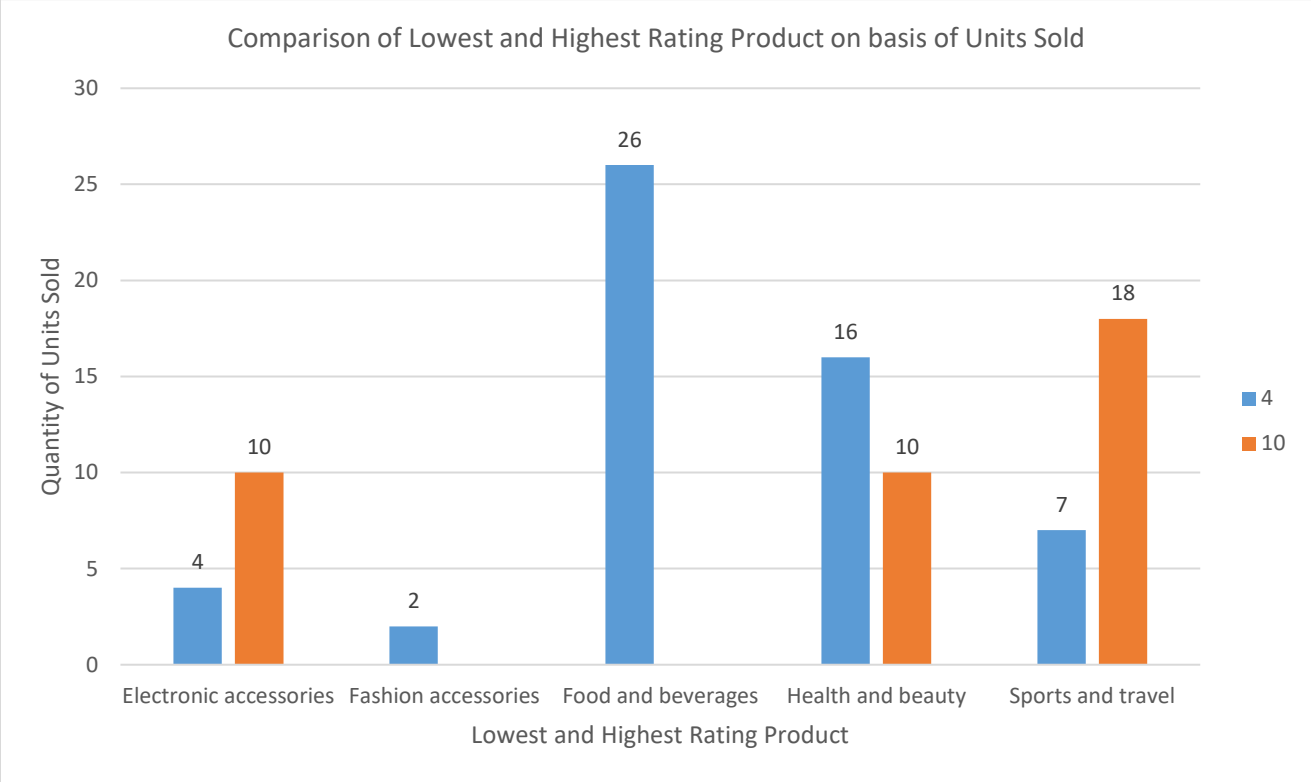




Answer. Our analysis of the Supermarket Sales Data revealed the following:

a. At Branch A, females placed the highest number of orders.
b. Branch B saw higer number of orders placed by Females
c. Meanwhile, at Branch C, males placed the most orders.

Q3. Compare highest and lowest rating products on the basis of units sold.



Answer. Upon analyzing the Supermarket Sales Data, we discovered that product ratings ranged from a minimum of 4 to a maximum of 10.

a) Electronic Accessories with higher ratings garnered more customer purchases, indicating a preference for quality in this category.
b) Fashion accessories and food and beverages mainly comprised lower-rated products in customer purchases.
c) Health and beauty products also leaned towards lower-rated items in customer preferences.
d) However, in the Sports and Travel category, customers showed a tendency to purchase higher-rated products.

Q4. Analyzing units sold and unit price data answer the following sub questions

a) What is the degree of freedom?
b) Co-relation of Unit price and revenue generated
c) What result you can draw from regression of the two data

SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0.010777564					
R Square	0.000116156					
Adjusted R Square	-0.000885732					
Standard Error	2.924724997					
Observations	1000					
ANOVA						
	df	*SS*	*MS*	*F*	*Significance F*	
Regression	1	0.9917274	0.991727	0.115937	0.733555221	
Residual	998	8536.908273	8.554016			
Total	999	8537.9				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	5.443794599	0.215314544	25.28299	2.1E-109	5.021273429	5.86631577
Unit price	0.001189202	0.003492565	0.340495	0.733555	-0.005664411	0.008042815

Answer:

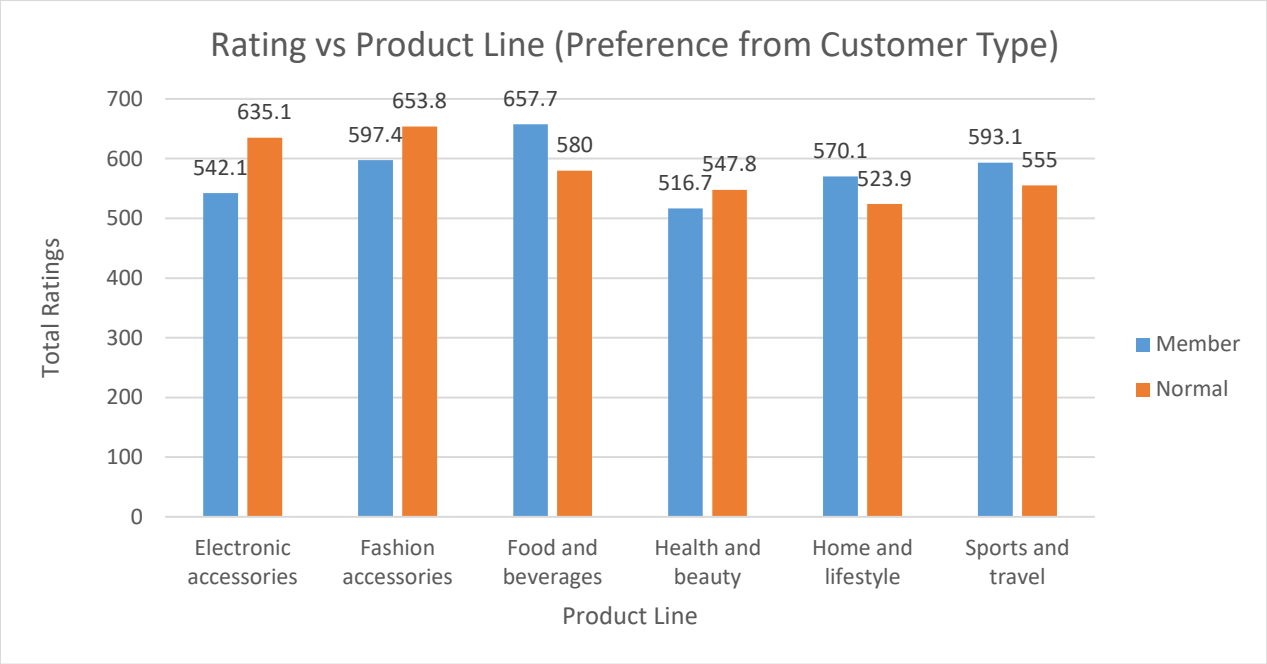
a. The degree of freedom of the analyzed data is 1.
b. The correlation between unit price and generated revenue was found to be 0.63392, indicating a moderate positive relationship. The analysis focused on the columns of unit price and total revenue, employing the CORREL function.
c. Upon examining the regression results, we aimed to discern the relationship between quantity and unit price, exploring how customers' purchasing quantity correlates with the unit price of a product.

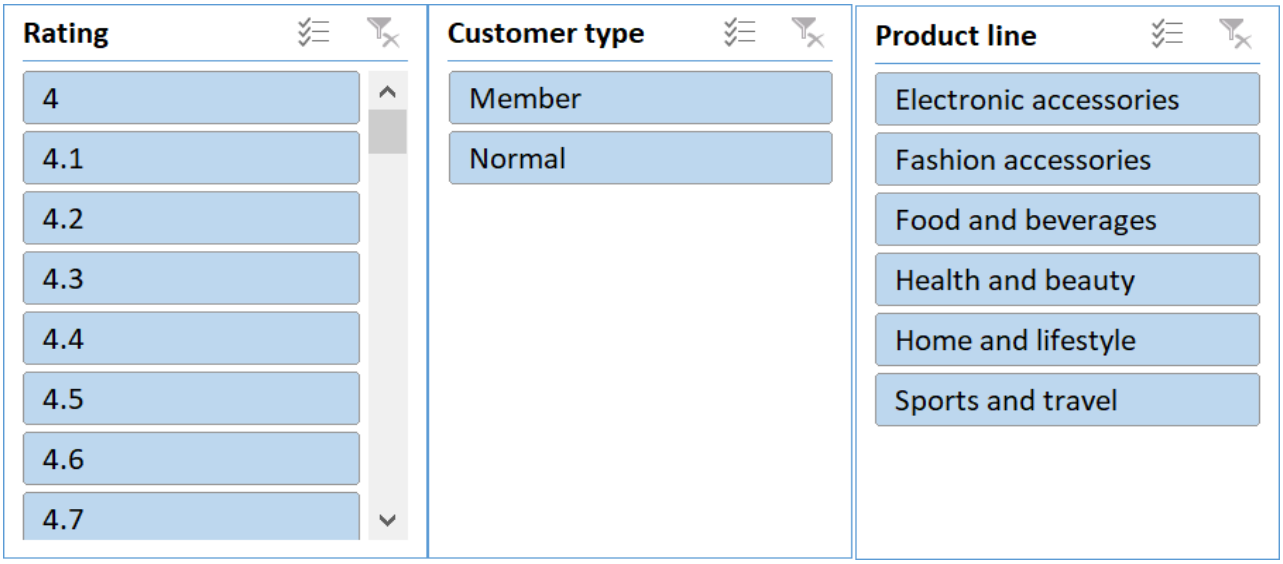
   However, from the regression analysis, it's evident that the observed trend lacks consistency. The expected outcomes derived from the trend deviate significantly from the actual outcomes.

   With a degree of freedom of 1, the trendline equation stands as

   Quantity = 0.0012x + 5.4438. Despite this equation, the coefficient of determination (R2) is merely 0.0001, highlighting the inconsistency in customer buying patterns solely based on unit price.

Q5. What product will you suggest as per the city data analysis to each type of customer





Answer. As per the city Data Analysis, **Food and Beverages** will be a good option for **Member** type customer and **Fashion Accessories** for **Normal** type of customers.

Conclusion and Reviews

The comprehensive analysis of supermarket sales dynamics provides valuable insights into consumer behavior, operational trends, and performance metrics. Here's a summary of the findings and reviews:

1. City Performance:

Mandalay emerged as the top-performing city among those with a 5% tax slab. Its superior sales/revenue generation signifies a potentially lucrative market for supermarket businesses.

2. Gender-based Ordering:

Female customers showed a higher propensity to order items from Branch A, while males dominated in Branch C. Branch B saw equal orders from both genders. This gender-specific trend highlights the importance of targeted marketing strategies.

3. Rating and Units Sold:

Further analysis is needed to compare products with the highest and lowest ratings based on units sold. Understanding the correlation between product ratings and sales volume can inform inventory management and marketing decisions.

4. Unit Price and Revenue Relationship:

The regression analysis revealed a weak correlation ($R2 = 0.0001$) between unit price and quantity sold. This suggests that customers' purchasing decisions may not be significantly influenced by unit price alone, indicating the need for deeper insights into consumer preferences and behavior.

5. Product Recommendations:

Based on city data analysis, Food and Beverages are recommended for member-type customers, while Fashion Accessories are suggested for normal customers. These recommendations align with the observed preferences and purchasing patterns in respective cities.

Reviews:

The report provides a thorough exploration of supermarket sales dynamics, covering various aspects such as city performance, gender-based ordering trends, and product recommendations.

The inclusion of regression analysis enhances the depth of insights, though further interpretation of the results could strengthen the analytical rigor.

Clear visuals, such as graphs and charts, would enhance the presentation of findings and aid in understanding complex relationships.

Overall, the report offers valuable insights for supermarket stakeholders, highlighting areas for strategic focus and improvement in marketing and operational strategies.