**XYZ SUPERMARKET ANALYSIS USING PANDAS**

**Awolusi Oluwabukola Olajumoke**

****

**INTRODUCTION**

Company XYZ, a leading supermarket chain with branches across the country, seeks to uncover insights into its sales trends and growth patterns. With increasing competition from emerging supermarkets, understanding customer purchasing behavior and sales dynamics is crucial for informed decision-making. This project aims to analyze sales data from three major branches in Lagos, Abuja, and Port Harcourt, covering a three-month period. By leveraging data analytics and machine learning techniques, we will identify key factors influencing sales and provide actionable recommendations to drive business growth.

**Background:**

The supermarket industry in the country is characterized by:

1. Increasing competition from local and international players.

2. Shifting consumer behavior towards convenience, quality, and price.

3. Growing demand for online shopping and home delivery.

4. Fluctuating economic conditions impacting consumer spending.

**RESEARCH OBJECTIVES**

This project aims to:

1. Analyze sales trends and patterns across branches.

2. Identify key drivers of sales growth and decline.

3. Develop predictive models to forecast future sales.

4. Provide actionable recommendations to enhance sales performance and competitiveness.

**Research Questions:**

1. What are the primary factors influencing sales across different branches?

2. How do sales trends vary between Lagos, Abuja, and Port Harcourt?

3. Can we predict future sales based on historical data?

4. What strategies can Company XYZ implement to improve sales performance?

**Expected Outcomes:**

This project will provide:

1. A comprehensive understanding of sales trends and patterns.

2. Data-driven insights to inform business decisions.

3. Actionable recommendations to enhance sales performance.

**PROBLEM STATEMENT**

The rise of supermarket competition poses a significant threat to Company XYZ's market share. To maintain its competitive edge, the company needs to:

1. Identify key drivers of sales growth and decline.

2. Uncover regional sales patterns and trends.

3. Determine the impact of seasonal fluctuations on sales.

4. Develop predictive models to forecast future sales.

**METHODOLOGY**

This project will employ a combination of:

1. Descriptive analytics to summarize sales data.

2. Inferential statistics to identify relationships and trends.

3. Data visualization to communicate insights.

**DATA COLLECTION**

1. Imported necessary libraries and set the working directory.

2. Combined datasets from the three branches using pd.concat.

3. Saved the combined dataset to a new CSV file.

4. Loaded the combined dataset and checked for missing values.

**DATA CLEANING AND PREPARATION PROCESS**

1. Performed data manipulation and cleaning.
2. Df.describe was used to check the statistical summary of the data set **df.isnull was used to check for missing values  
   df.info() was used to check the information of the datase**

**CONVERTING DATA TO APPROPRIATE DATA TYPE**

1. Converted date columns to datetime format.

2. Added new rows for day, month, year, and hour.

3. Grouped data by city and aggregated.

**ANALYSIS AND VISUALIZATION**

Generated visualizations using seaborn.

Countplot was used to determine the branch with the highest sales record as shown below:

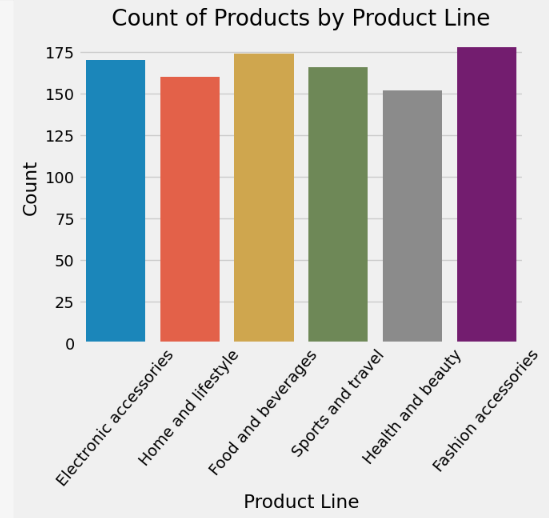


Fig1: Fashion and accessories has the highest sales record

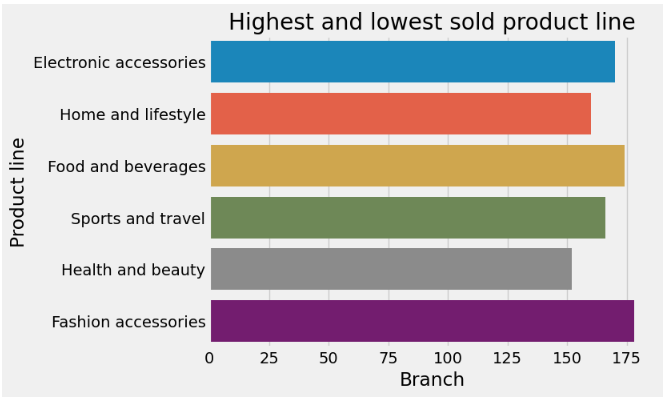


Fig2: This shows the highest and lowest sold product line

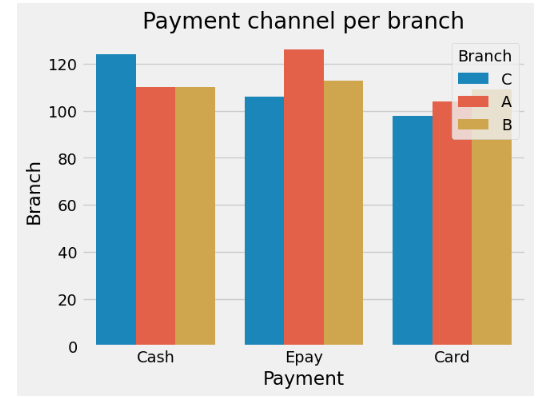


Fig 3: Payment method used per branch

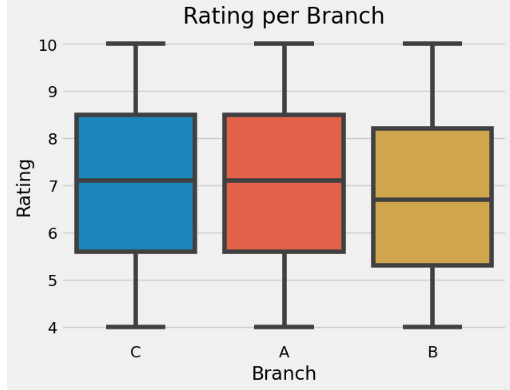


Fig4: Branch c has the highest rating

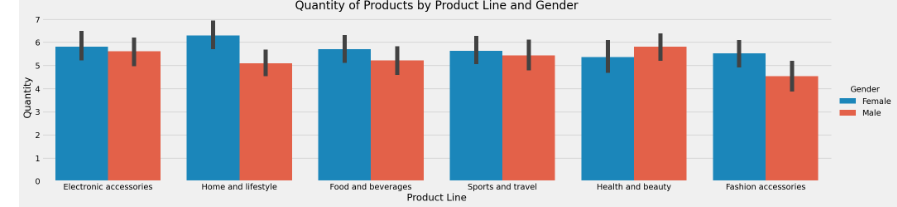


Fig 5: Quantity of product by product line and gender line

* Observation: It was observed that across all cities where the supermarket is located, women purchase;
* more fashion accessories, compared to males, while male have a higher habit of purchasing sports and leisure, and health and beauty women also record a high total purchase of home and life style products
* High quantity of food and beverages a slight difference is seen in the quantity of electronic accessories product between male and female



Fig 6: product line by unit price

* Observation: Fashion accessories has the highest unit price while health and beauty has the lowest unit price

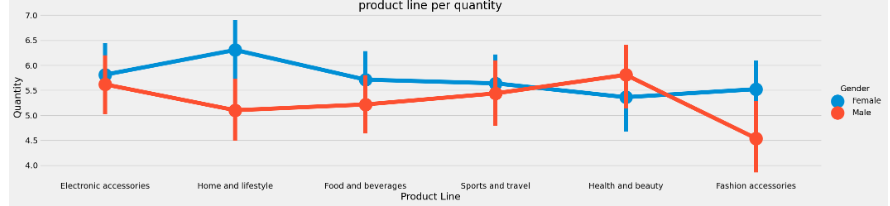


Fig 7: product line by quantity

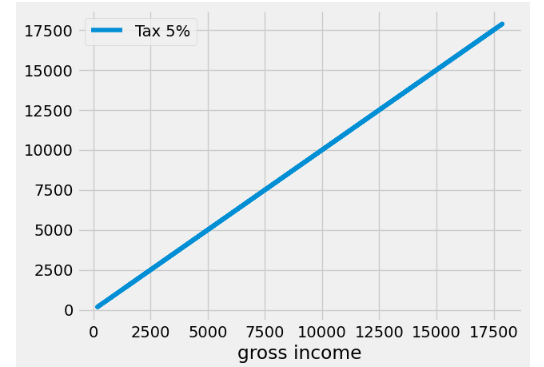


Fig 8: gross income by tax

### Observation: This shows that the higher the gross income, the higher the tax as shown

**CONCLUSION AND KEY FINDINGS**

Based on the analysis of sales data from Company XYZ's three branches in Lagos, Abuja, and Port Harcourt, several key insights have emerged, providing valuable recommendations for business growth and strategic decision-making.

1. Product Category Sales: Highest sales in Home and Lifestyle, Fashion Accessories, and Electronics (primarily among female customers).

2. Gender-Based Purchasing Habits: Female customers purchase more Fashion Accessories, while male customers prefer Sports and Leisure.

3. Product Quantity and Unit Price: High quantity of Food and Beverages sold, with Fashion Accessories having the highest unit price.

4. Payment Channels: Payment channels vary across branches, with Epay and Cash being the most common.

5. Gross Income and Tax: Higher gross income leads to higher tax.

By analyzing xyz supermarket data, we’ve been able to answer the research question listed below including providing actionable recommendation

**Research Questions:**

1. What are the primary factors influencing sales across different branches?

2. How do sales trends vary between Lagos, Abuja, and Port Harcourt?

3. What strategies can Company XYZ implement to improve sales performance?

**Recommendations**

- Female customers drive sales in Fashion Accessories and Home and Lifestyle products.

- Male customers dominate sales in Food and Beverages, Sports, and Travel.

- Optimize pricing strategies for Fashion Accessories and Health and Beauty products.

- Enhance payment channel options to improve customer convenience.

- Conduct Time Series Analysis, Sentiment Analysis, and develop interactive dashboards to visualize key insights and trends.

**Potential future work**

1. Segmentation Analysis: Perform customer segmentation based on demographics, behavior, and preferences to identify target audiences.
2. Predictive Modeling: Develop predictive models to forecast sales, customer churn, and response to marketing campaigns.
3. Product Recommendation: Implement a product recommendation system to suggest relevant products to customers based on their purchase history and preferences.