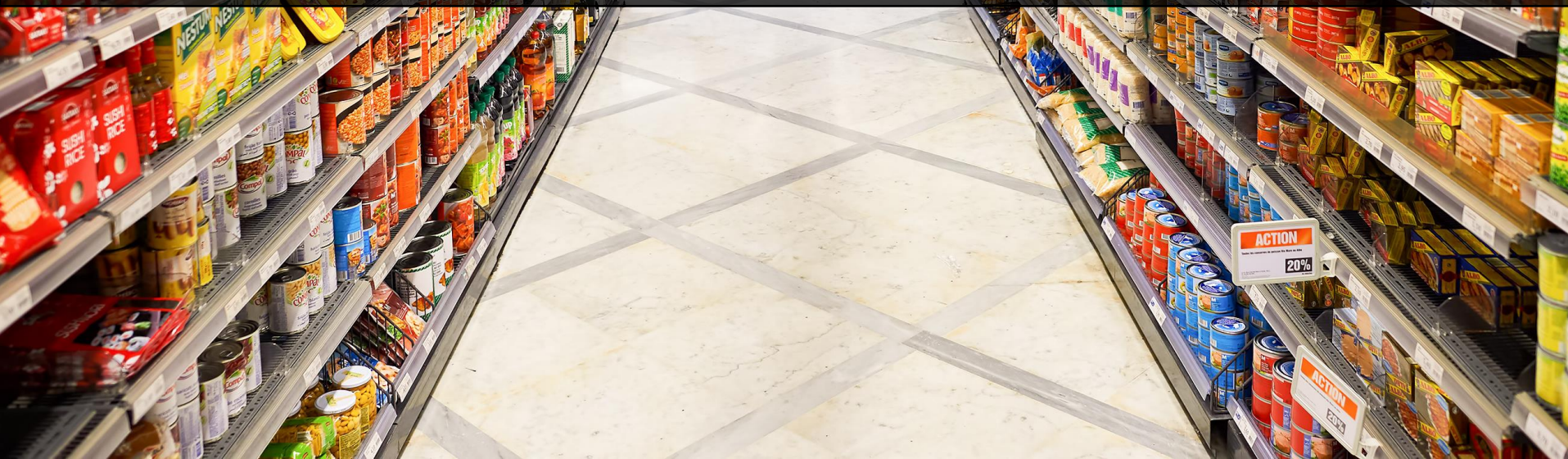




PANDAS ANALYTICS CAPSTONE PROJECT

Data manipulation, Processing and Analytics on XYZ Supermarket





PROJECT OVERVIEW

Company XYZ owns a supermarket chain across, with its major branch located in three(3) cities **Abuja, Lagos and Port Harcourt** with the following product lines; electrical accessories, Electronic accessories, Home and lifestyle, Food and beverages, Sports and travel, Health and beauty, Fashion accessories.

PROJECT OBJECTIVES

Descriptive data manipulation, processing, and analysis to be performed on the gathered data from all three branches and insights are to be generated to help understand trends and determine growth across the data from the locations

PROJECT STANDOUT SESSION

- Group-by categorization of the areas with the most rating
- Distribution visualisation of sales by gender to see the volume of the total amount contributed by respective gender.
- Comparison of other features to an important column '**RATING**' which was not given in the project requirement. It is crucial to know the reaction of the branches to respective products and services provided, rating based on product line.

PROJECT METHODOLOGY



In the course of the analysis the following steps were taken before and during the analysis;

- First, performed a deep study of each description of features in order to understand the concept of the supermarket and point out probably areas of interest of the enterprise towards its growth, development, and deployment of efficient business management and strategies.
- Secondly, a short statistical summary of the data that was gathered from the three main branches and drew a brief insight of it.
- Thirdly, checking for missing values and noted the observation and also checked the current state (data-type) of all the columns using the `info()` attribute
- Moving on, I performed few breakdowns of some columns that were meant to have been in DateTime format, hence the process of datatype conversion was initiated after which hours, days, weeks, and months were extracted from the DateTime columns to the end that more detailed information about periodical trends are obtained
- Next, I obtained the unique values in all categorical features.
- I also performed groupby type categorization and aggregation on the data, other columns by city, and other columns
- Lastly I used matplotlib and seaborn library for the visualization of the data to get inferences from the data. The visualization is compressed into three parts: a) Trend Identification b) Correlation c) Distribution

ANALYSIS AND INSIGHT FROM EXPLORATION



▪ STATISTICAL SUMMARY INSIGHT

The data properly collated as there were no missing values or outliers. Starting with the statistical summary, interesting insight about the approximate average unit price and quantity are #20,000 and 6 units respectively making our total sales/revenue to be around #120,000 without incurring tax.

Note: The tax on the supermarket is about 4.5% of the total sales. After accounting for tax the average total value is found to be around #116,000. Knowing that the average cost of goods sold is approximately #110,700 and the gross income is approximately #5,537 it informs us that the average sales made within the timeframe are about #115,000

The average rating is 6.97 indicating that the supermarket is doing slightly better above-average implying that more work needs to be done to improve the companies services and products available to customers

We can say inferentially that Port Harcourt has the highest total gross income with Abuja being the second and then Lagos.

▪ RATING INSIGHT

The city with the highest average rating is and the City with the highest average rating is Port Harcourt with Lagos being the next in line followed by Abuja.

The branch with the highest average rating is and the City with the highest average rating is branch C with branch B being the next in line followed by branch A.



▪ PAYMENT CHANNEL USAGE INSIGHT

A large percent of the electrical accessories were paid for using the cash channel

For home and lifestyle product lines it was majorly purchased using Epay

There is a fine distribution of the payment channels used for the food and beverages product line

Sports and travel have the majority of their products paid in cash

Most of the customers who purchase health and beauty products were paid with using Epay

A very large amount of the Fashion accessories were paid for using Epay

▪ SALES CHANNEL INSIGHT

Food And Beverages are sold most in Port-Harcourt

Fashion accessories are also sold most in Port-Harcourt

Lagos has the highest purchase for electronic equipment

Sports and travel products were mostly sold in Abuja

Home and lifestyle commodities are heavily purchased in Lagos

All three states have an approximately equal number of sales for health and beauty commodities

The supermarket should focus on increasing its marketing strategy for sales of health and beauty products especially in Lagos which has a high potential yield rate

In other for the Branch at Port-Harcourt to do better the sales management should work on generating more incoming through sports and travel, home, and lifestyle commodities.



More efforts need to be put into the Lagos fashion accessory fraction

Relative to Lagos and Port Harcourt Abuja branch hasn't been doing well sales-wise as regards the yield in sales for the food, beverage, home, lifestyle, health and beauty hence more attention should be given to these areas

CORRELATION ANALYSIS

This indicates that there is a high variance in the correlation between the cost of goods sold and the total amount generated

Dominant female activity in branch A and B while the men seem to be slightly higher than the females at Branch C

CONCLUSION

The summary of this analysis has proven to show that the supermarket across the three (3) branches and cities is operative on the same relative threshold with not very high rating from the customers though not bad as there is a fine proportion of the contribution of both male and female to the market growth. So in just a matter of time and with sales, marketing, and service development the XYZ restaurant would be better than it is currently.