**Analyze Supermarket Data Across the Country - Company XYZ**

**Introduction**

In this project, I analyze and report findings on the Sales data for Company XYZ which owns a supermarket chain across the country. Each major branch of the supermarket is located in 3 cities across the country and the data consist of a record of sales information for 3 months. The aim of this project is to help the company understand its sales trends and determine its growth, and possibly make a few recommendations to improve the business.

* The city with the highest total gross income: Port Harcourt.
* The city with the highest Customer Satisfaction rating on average on their overall shopping experience is Port Harcourt.
* The city with the highest Total Quantity of Purchases is Lagos.
* The most expensive product by unit price is the Fashion accessories line, which registers the least patronage from customers across all the product lines.
* Though the Sports and travel product line is one of the most expensive products, customers patronage is not as low when compared to the Fashion accessories product line. This could mean that customers value Sports and travel more than they do Fashion accessories.

Other findings are captured in the body of this report.

**Approach**

Each dataset (file) per branch which is available in a .csv file is combined into a single file, exported into a .csv and read into a pandas dataframe. A data exploration was done on the combined dataframe to confirm its shape, object type, missing value, and to get a statistical summary of the data. The Date and Time datatype were updated to the datetime datatype; and then features of Date and Time were extracted into separate additional columns for further analysis.

**Data Overview**

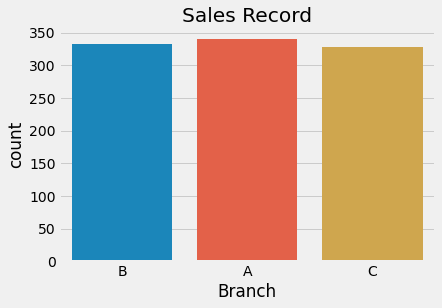
Each branch of the supermarket in the 3 cities have submitted similar datasets containing the same information across a 3-month period. Some of these information include:

* City
* Customer type
* Product line
* Gender
* Payment type
* Details of goods purchased like quantity and amount paid
* Time and Date of purchase
* Customer Satisfaction rating
* Etc

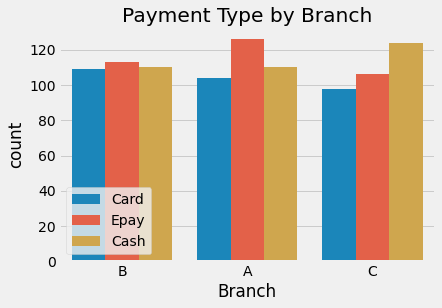
**Analysis**

With the intent of analyzing the supermarket data across the 3 branches, an Aggregation with Groupby method across the 3 branches of interest was done on the dataset; a number of questions were generated and answered in this project, some of which are listed below:

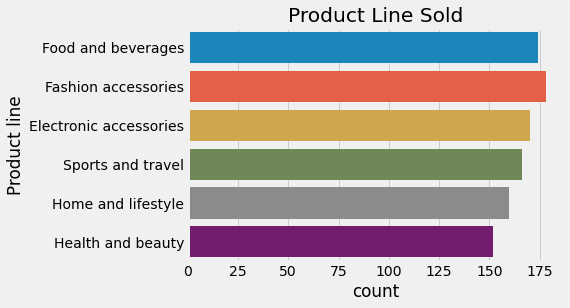
1. The sum and mean of the gross income across the branches.
2. The time of the day with the most shopping payments across the cities.
3. The City (branch) with the highest rating by Customer Satisfaction.
4. The City (branch) with the highest Total Quantity of purchases.
5. Determined the branch with the highest sales record using a countplot.



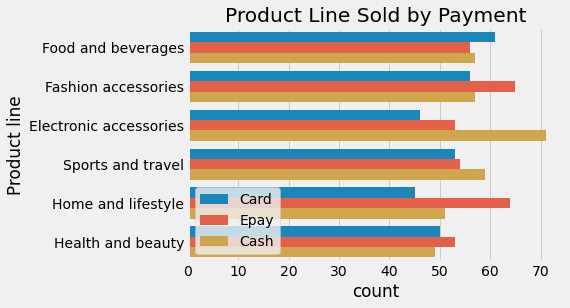
1. Determined the payment type breakdown per branch, using a countplot.



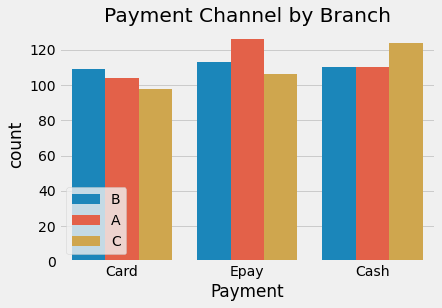
1. Determined the highest & lowest sold product line, using Countplot.



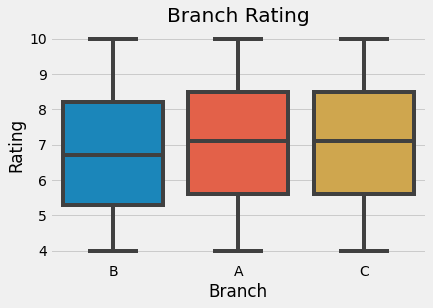
1. Determined the Payment channel used by most customer to pay for each product line.



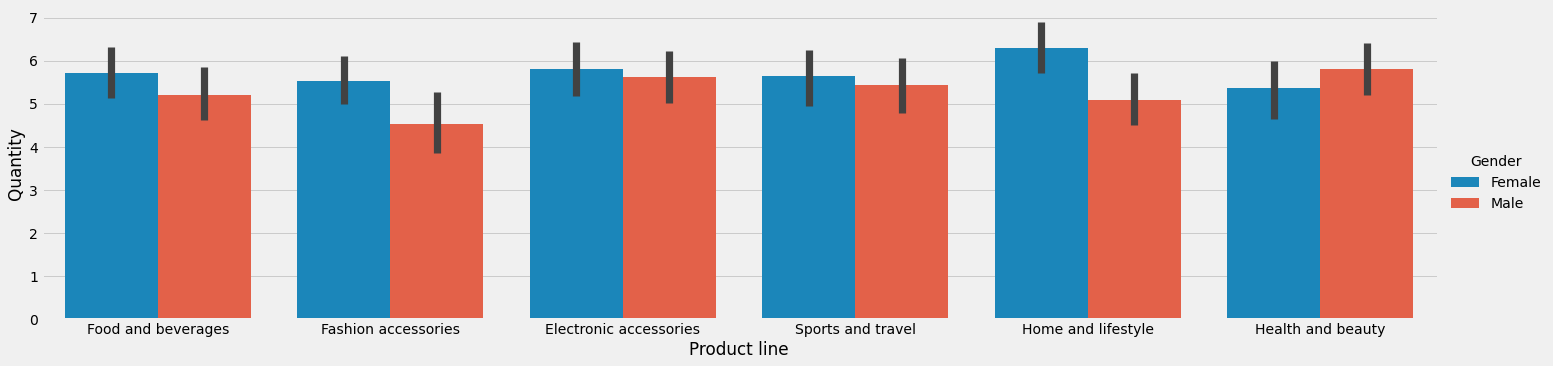
1. Determined the Payment channel for each branch.



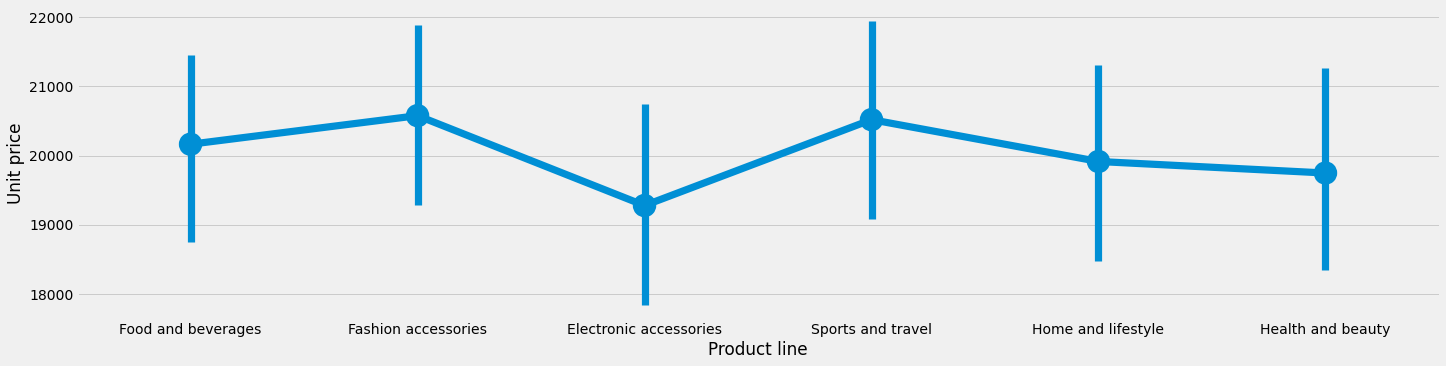
1. Determined the branch with the lowest rating using a box plot.

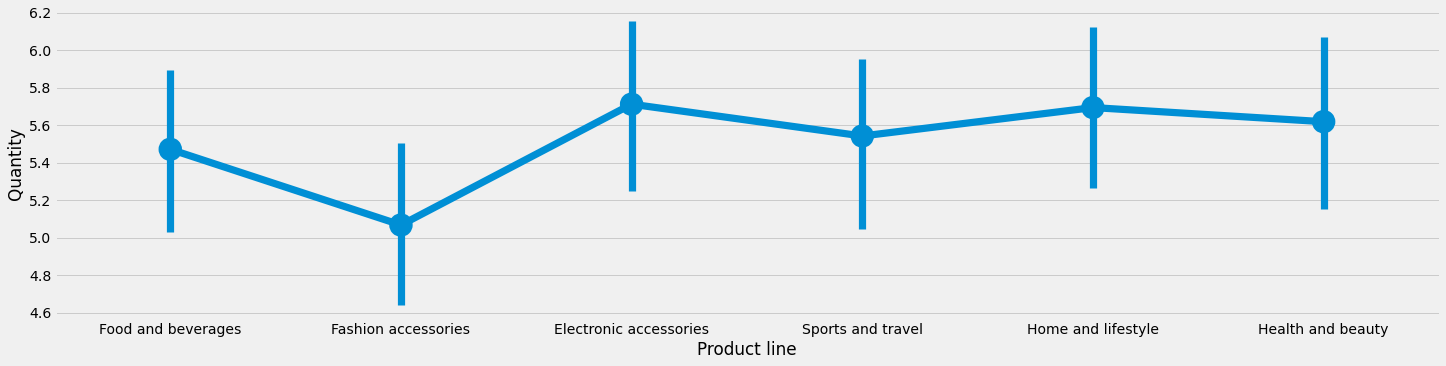


1. An analysis of the gender type versus the kind of products being purchased at the supermarket was done using the catplot.

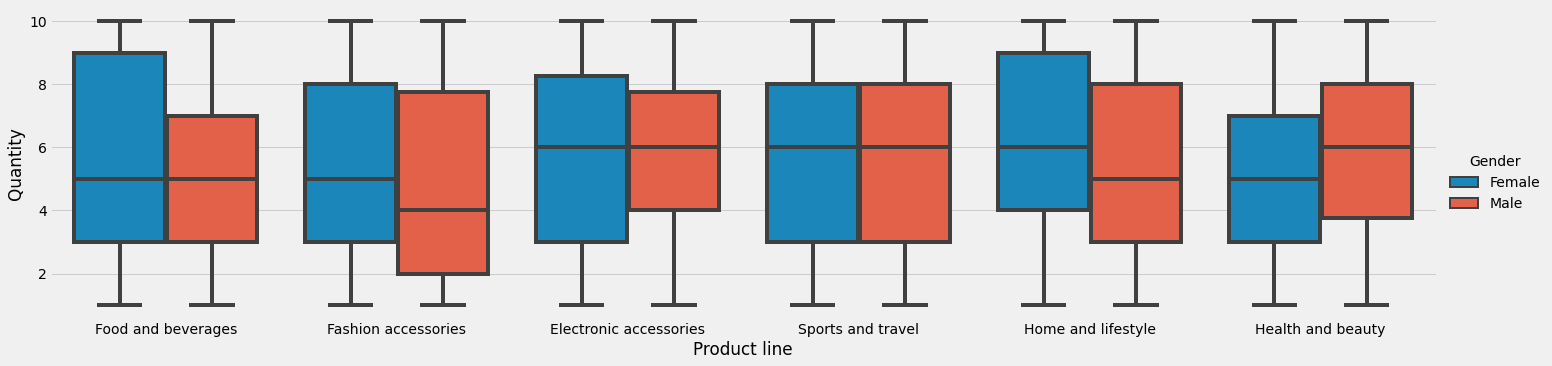


1. The interaction of Unit price on the Quantity of goods purchased.





1. The interaction between Customer type and Quantity of Product line purchased using a catplot.



**Conclusion**

Following the analysis of the available data, a number of findings and conclusions have been inferred and are captured below:

1. The city with the highest total gross income: Port Harcourt
2. The time of day with most shopping payments in Port-Harcourt is the 10th hour.
3. The time of day with most shopping payments in Abuja is the 19th hour.
4. The time of day with most shopping payments in Lagos is the 10th hour.
5. The city with the highest Customer Satisfaction rating on average on their overall shopping experience is Port Harcourt.
6. The city with the highest Total Quantity of Purchases is Lagos.
7. Branch A, which corresponds to Lagos has the highest sales record.
8. Epay ranks as the most used form of payment in Branch A and Branch B, while Cash payment tops the charts for Branch C.
9. The Fashion accessories product line registered the highest number of sales, while the Health and beauty line was the least sold across the Branches.
10. Cash stands out as the most preferred payment for Electronic accessories, while Epay is preferred for Home and lifestyle.
11. The branch with the lowest rating is Branch B, which corresponds to the city of Abuja.
12. Across the Branches, females purchased more of the all the available Product line except for Health and beauty, than males.
13. Although the analysis shows that Males spent more money than Females in the Sports and travel Product line, even though Females purchased more in quantity.
14. Males also spent more money in the Health and beauty Product line, and purchased more quantities than Females here as well.
15. The most expensive product by unit price is the Fashion accessories line, which registers the least patronage from customers across all the product lines.
16. Though the Sports and travel product line is one of the most expensive products, customers patronage is not as low when compared to the Fashion accessories product line. This could mean that customers value Sports and travel more than they do Fashion accessories.
17. The cheapest product by unit price is Electronic accessories which shows a corresponding high patronage by Quantity.
18. The average quantity purchased by either male or female is the same for the following Product lines:

* Food and beverages
* Electronic accessories
* Sports and travel

1. The average quantity purchased by Female is more than that purchased by their male counterpart for the following Product lines:

* Fashion accessories
* Home and lifestyle

1. The average quantity purchased by Male is more than that purchased by their female counterpart in the Health and beauty Product line.