**Title: Exploratory Data Analysis Report**

**Prepared for: SCKU Datathon**

**Prepared by: [Catherine Auma, Eubert Moracha, Ezekiel Omagwa], [The Datathon.]**

**Date: [19th February, 2024]**

**Executive Summary:**

This report presents the findings of our exploratory data analysis (EDA) conducted on the Global Superstore dataset. The objective of this analysis was to understand customer behavior, product performance, and identify opportunities for social impact in alignment with Sustainable Development Goals (SDGs). By leveraging data analysis techniques, we aimed to provide actionable insights for businesses to drive social impact and sustainable development.

**Introduction:**

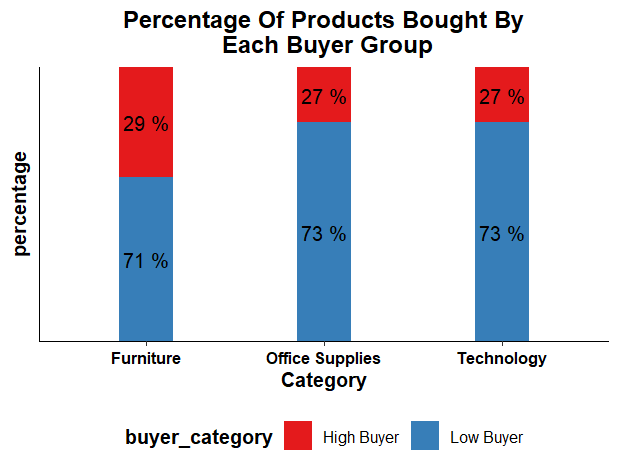
The objective of this report is to analyze the Global Superstore dataset to gain insights into customer behavior, product performance, and opportunities for social impact. Through this analysis, we aim to demonstrate the potential of data-driven insights in addressing societal challenges and contributing towards achieving Sustainable Development Goals.

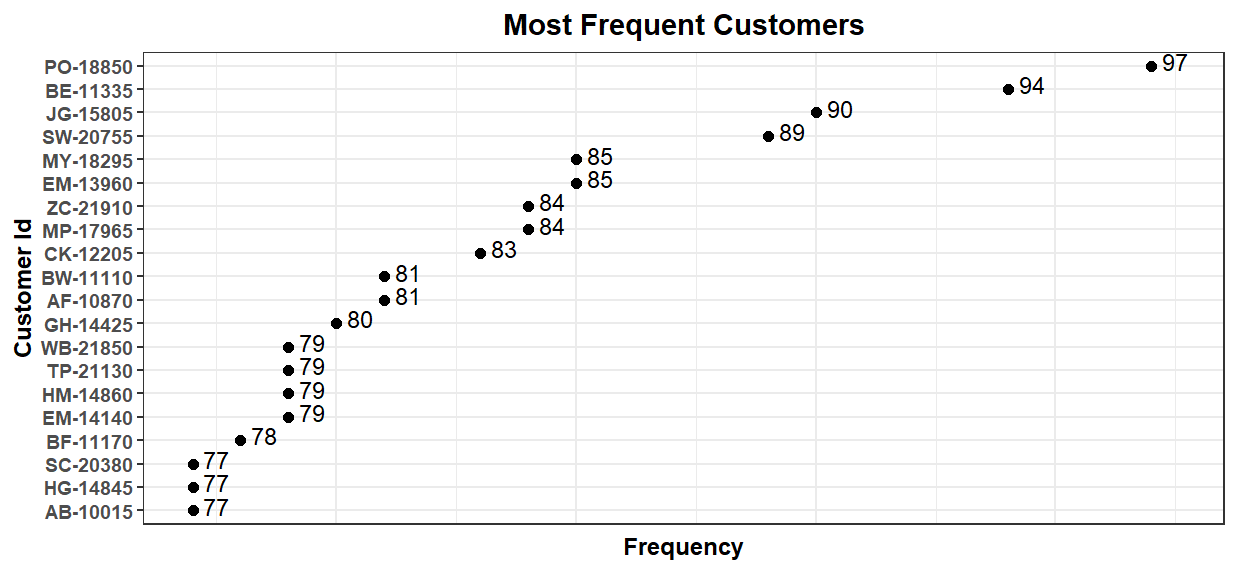
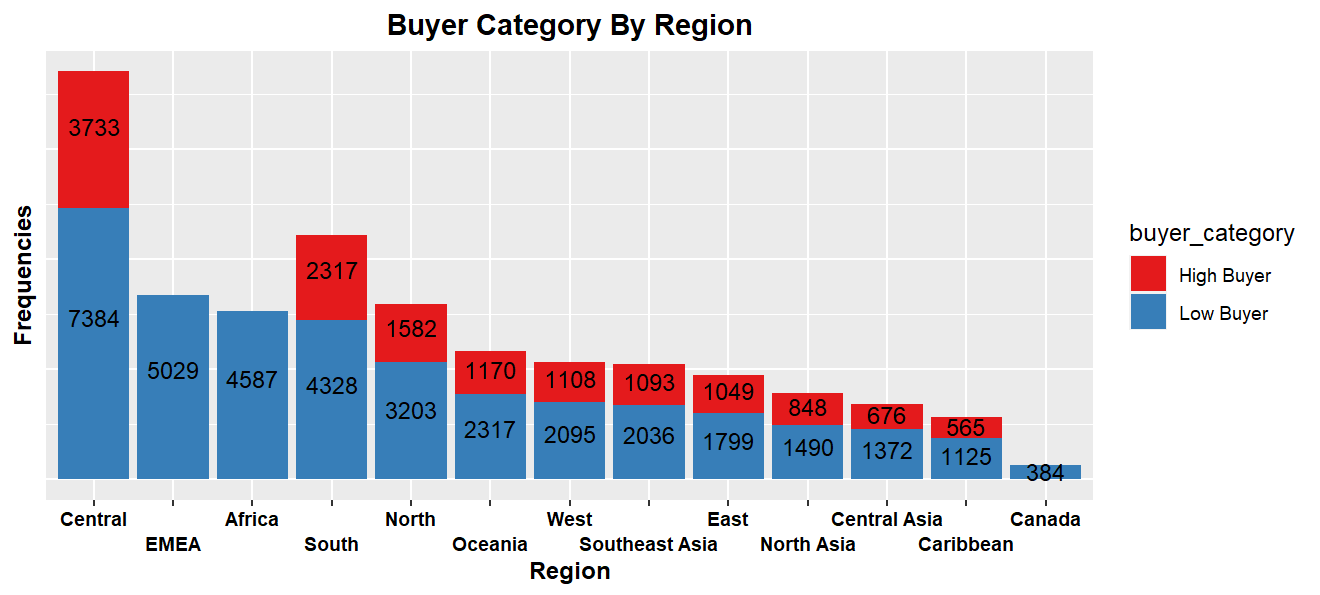
**Methodology:**

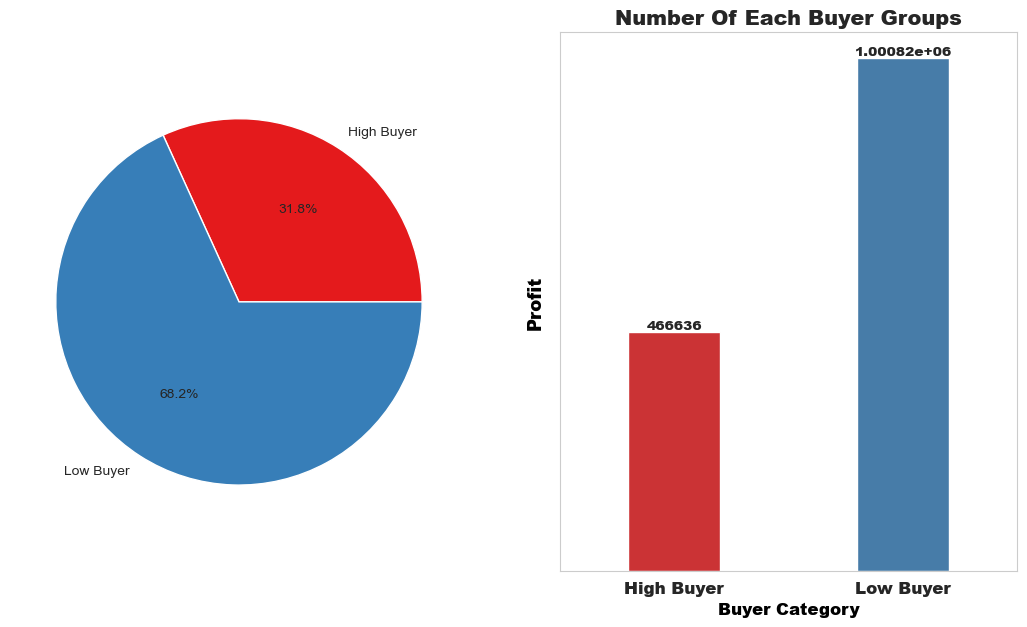
We imported the Global Superstore dataset into R Studio and conducted exploratory data analysis using R programming language. The dataset was cleaned and transformed to ensure accurate analysis. Various statistical and visualization techniques were employed to analyze customer and product data.

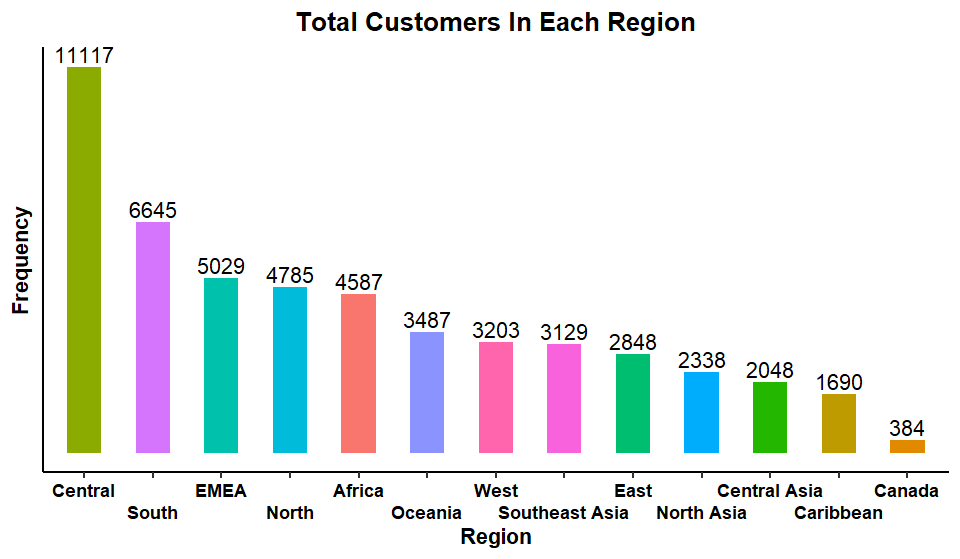
**Plots**

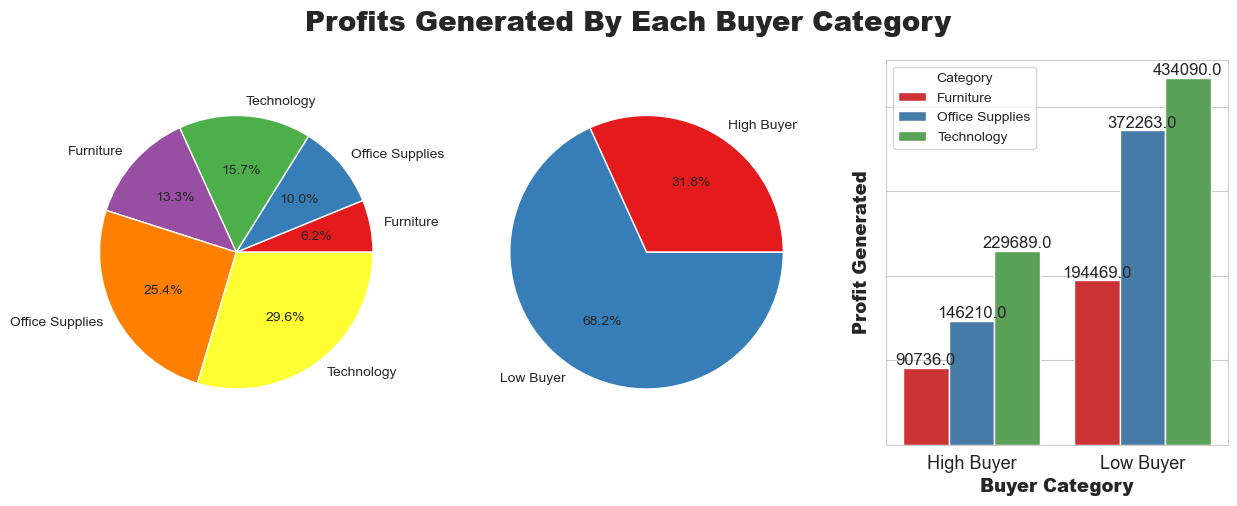
The exploratory data analysis revealed insights into customer behavior, geographical distribution, and product performance. High-frequency customers emerged as significant revenue drivers, while geographical analysis highlighted regions for targeted marketing. Profit margins between buyer groups were compared, and top-selling product types were identified annually. Trends in profit and sales provided valuable insights for business optimization. Overall, the analysis underscored the importance of data-driven strategies in understanding customers and driving business growth.

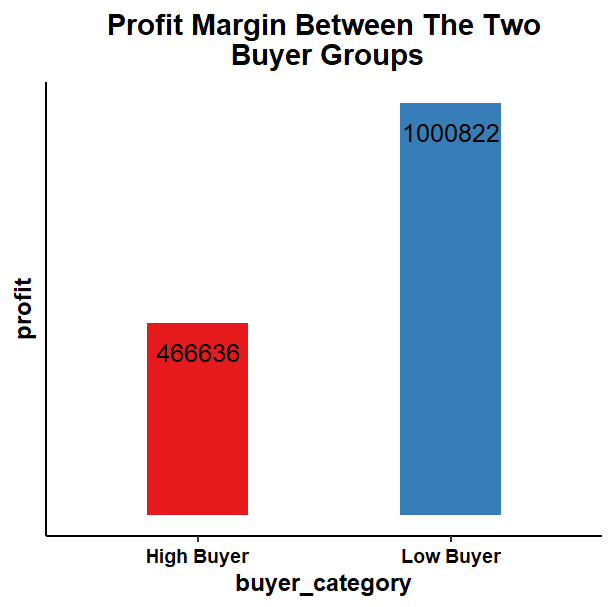


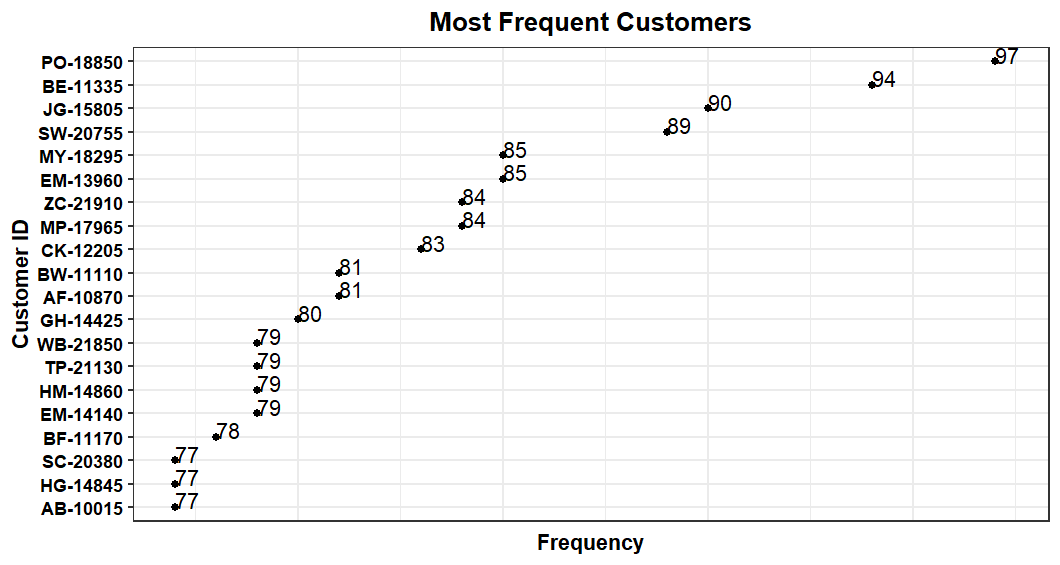












**Findings:**

***Customer Analysis:***

1. Profiled customers based on their frequency of purchase, identifying high-frequency customers contributing significantly to revenue.
2. Segmented customers to identify the most profitable customer segments across different years.
3. Analyzed the geographical distribution of customers to identify potential areas for social impact.

**Product Analysis:**

1. Identified the top sales country and the top 5 profit-making product types annually.
2. Explored the relationship between product price and sales volume to understand price elasticity.
3. Calculated the average delivery time across different countries to assess logistical efficiency.

**Social Impact and SDGs**

1. Evaluated the distribution of customers across regions to identify opportunities for social impact in underserved regions.
2. Identified product categories aligning with SDGs such as poverty alleviation, gender equality, and environmental sustainability.
3. Proposed strategies for businesses to integrate social impact initiatives into their operations and contribute towards achieving SDGs.

**Conclusion:**

In conclusion, our analysis of the Global Superstore dataset has provided valuable insights into customer behavior, product performance, and opportunities for social impact. By leveraging data-driven insights, businesses can not only drive profitability but also contribute meaningfully towards achieving Sustainable Development Goals. We believe our analysis and recommendations demonstrate the power of data in addressing societal challenges and advancing sustainable development agendas.

**Recommendations:**

Based on our analysis, we recommend businesses to:

1. Focus on retaining high-frequency customers through targeted marketing strategies.
2. Align product offerings with SDGs to contribute towards societal welfare.
3. Improve logistical efficiency to enhance customer satisfaction and reduce costs.

**Acknowledgements:**

We would like to express our gratitude to SCKU Datathon for providing us with the opportunity to undertake this assignment. Additionally, we acknowledge the support and guidance received from our peers throughout the analysis process.

***\*\*End of Report\*\****