**Supermarket Dataset Analysis Report**

**Add University Cover Page**

**Add Table of Contents**

1.Introduction and Project Plan

2.Data Quality Issues and Remedies

3.Data Analysis and Commentary

4.Data Charting and Commentary

5.Conclusions and Recommendations

6.Referrences

7.Appendix(Optional)

**TASK-1: Introduction and Project Plan**

Project Overview:

This project focuses on evaluating the performance of social media marketing campaigns for Robinsons Supermarket, specifically in relation to their trial partnership with SOS Marketing to manage their Facebook account. The key objective is to determine whether outsourcing Facebook management to SOS Marketing has yielded better performance across key metrics, such as Impressions, Clicks, and Shares.

Objective

The primary objective of this project is to analyze the impact of outsourcing Facebook social media management to SOS Marketing for Robinsons Supermarket.

* Assess the impact of the SOS Marketing trial
* Analyze the overall performance of social media platforms
* Evaluate product sector performance across social media
* Examine the effectiveness of different campaign goals
* Conduct monthly performance analysis
* Provide actionable recommendations

These objectives aim to provide a comprehensive understanding of the performance of social media campaigns and the value of outsourcing Facebook management to SOS Marketing.

Plan:

1. **Data Quality Review**: The first step involves identifying and addressing any issues within the dataset, including missing values, outliers, or incorrect data entries.
2. **Exploratory Data Analysis (EDA)**: Through visualizations and descriptive statistics, the dataset will be explored to uncover key trends and relationships.
3. **Social Media Platform and Product Sector Performance Analysis**: This analysis will help identify which platforms and sectors generate the most engagement and drive the highest overall impact.
4. **Final Analysis and Insights**: Based on the results, recommendations and insights to enhance marketing strategies and achieve better results.

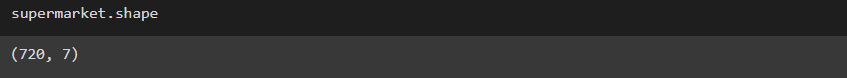
A black rectangular object with a white border

Description automatically generated

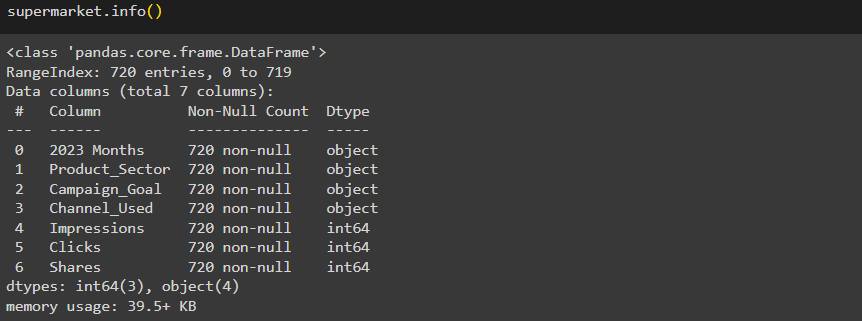
It imports essential libraries for data manipulation (pandas, numpy), visualization (matplotlib, seaborn, plotly), and suppresses warnings for efficient data analysis and plotting.



Here, it reads an Excel file named "Supermarket dataset (1). xlsx" into a pandas DataFrame called `supermarket` for data analysis.



The `supermarket.shape` command returns the dimensions of the `supermarket` dataset, indicating the number of rows and columns in the format `(rows, columns)`.

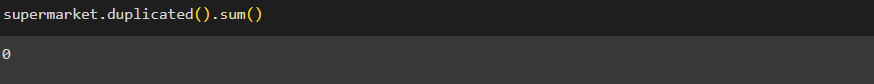


The `supermarket.info()` command provides a concise summary of the `supermarket` dataset, including the number of non-null entries, data types of each column, and memory usage. It helps in understanding the structure of the dataset and detecting missing values or incorrect data types.

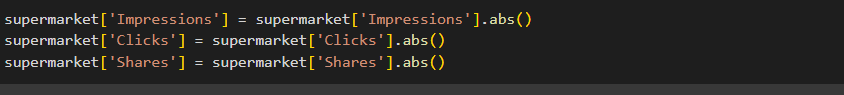
A screenshot of a computer

Description automatically generated

The `supermarket.isnull().sum()` command returns the total number of missing (null) values in each column of the `supermarket` dataset, helping to identify columns that may need data cleaning or imputation.



The `supermarket.duplicated().sum()` command returns the total number of duplicate rows in the `supermarket` dataset, indicating potential redundancy that may need to be addressed during data cleaning.



It convert all values in the ‘Impressions’,’Clicks’ and ‘Shares’columns of the `supermarket` dataset to their absolute values, ensuring that any negative values are corrected.

**TASK2:** **Data Quality Issues and Remedies**

Quality Issues:

1. Inconsistent Naming Issue:

* Product Sector Names**:** The dataset contains inconsistent naming conventions such as 'Fashion!!!!!', 'H$m£', and 'Tech$'. These variations in naming can lead to misclassification and difficulties in data analysis.
* Campaign Goal Names**:** Entries like 'Increase S@l£s' use special characters and non-standard formatting, which can affect data aggregation and reporting.
* Channel Names**:** Variants such as 'Face b\*\*k' and 'Instgrm' are misspelled and inconsistent, leading to possible confusion and inaccurate categorization.

2.Negative Values in Key Columns**:** The dataset contains key performance metrics such as Impressions, Clicks, and Shares which are typically expected to have non-negative values.

3. Outliers: There were outliers in the key numeric columns are Impressions, Clicks and Shares columns, which could skew results.

Remedies Applied:

1.Naming Standardization:

* Product Sector Names Standardization**:** Inconsistent names were corrected to ensure uniformity. For Example: 'Fashion!!!!!', 'H$m£', and 'Tech$' was replaced to 'Fashion', 'Home', and 'Technology' respectively.
* Campaign Goal Names Standardization: Inconsistent names were corrected to ensure uniformity. For Example: 'Increase S@l£s' was replaced to ‘Increase Sales’.
* Channel Names Standardization: Inconsistent names were corrected to ensure uniformity. For Example: 'Face b\*\*k' and 'Instgrm' was replaced to 'Facebook' and 'Instagram'.

2. Handling Negative Values: Negative values in the columns Impressions, Clicks, and Shares were converted to their absolute values, which is essential for reliable data analysis and reporting.

3.Outlier Removal: The Interquartile Range (IQR) method was applied to remove outliers in the dataset.

Impact of Data Quality Adjustments:

These corrections improved the overall quality of the data, ensuring that the analysis would reflect real-world trends and not be distorted by erroneous entries or outliers.

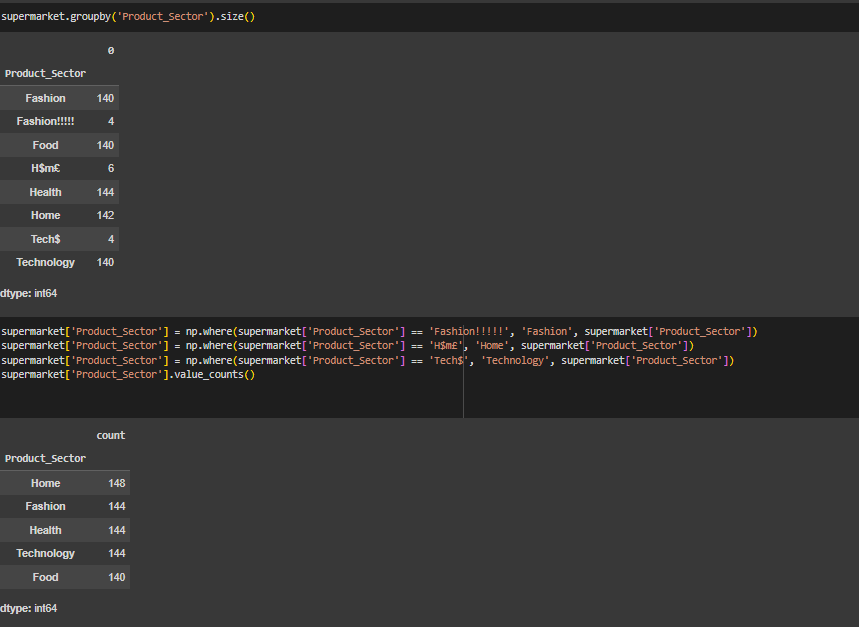
A screenshot of a computer

Description automatically generated

**A screen shot of a computer

Description automatically generated**

Here, group the `supermarket` dataset by the '2023 Months' column to get the size of each group, then correct misspellings or inconsistent entries ‘Noviembre’ to ‘Nov’ followed by counting the occurrences of each 2023 months value in the updated column.



The Product\_Sector column was standardized by correcting inconsistent labels, resulting in accurate counts for the categories: Home, Fashion, Health, Technology, and Food.

A screenshot of a computer

Description automatically generated

The Campaign\_Goal column was cleaned by correcting the inconsistent "Increase S@1es" value to "Increase Sales", resulting in accurate counts across the following categories: Brand Awareness , Increase Sales , Market Expansion , and Product Launch .

A screen shot of a computer

Description automatically generated

The Channel\_Used column was cleaned by correcting the inconsistent values "Face b\*\*k" and "Instgrm" to "Facebook" and "Instagram", respectively. After cleaning, the data shows equal counts across the three social media platforms: Facebook , Instagram , and YouTube .

A screenshot of a computer

Description automatically generated

Creates grouped boxplots using Plotly to visualize the distribution of ‘Impression’, ‘Clicks’ and ‘Shares’ from the ‘Supermarket’ dataset.



It generates boxplots using Plotly to visualize the distribution of ‘Impression’, ‘Clicks’ and ‘Shares’ from the ‘Supermarket’ after outlier values have been clipped based on the interquartile range (IQR) method.

**Objective-1:**

A screen shot of a computer

Description automatically generated

A screen shot of a computer

Description automatically generated

It analyzes social media data from a supermarket to determine the total impact of different platforms (Facebook, Instagram, YouTube) based on impressions, clicks, and shares. By summing these metrics, the script identifies which platform generates the highest overall engagement. The results are displayed in a formatted table using the tabulate library, providing clear insights into the effectiveness of each social media channel.

**Objective-2:**

A screen shot of a computer code

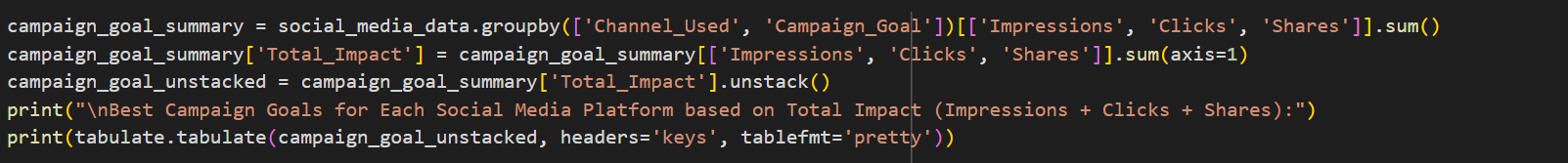
Description automatically generated

A screen shot of a computer

Description automatically generated

It analyzes social media data to determine the total impact of advertising on different product sectors. By summing impressions, clicks, and shares, the script identifies which sector benefits the most from social media engagement. The results are presented in a clear, formatted table, highlighting the sector with the highest overall impact. This analysis provides valuable insights into the effectiveness of social media strategies across various product sectors.

**Objective-3:**

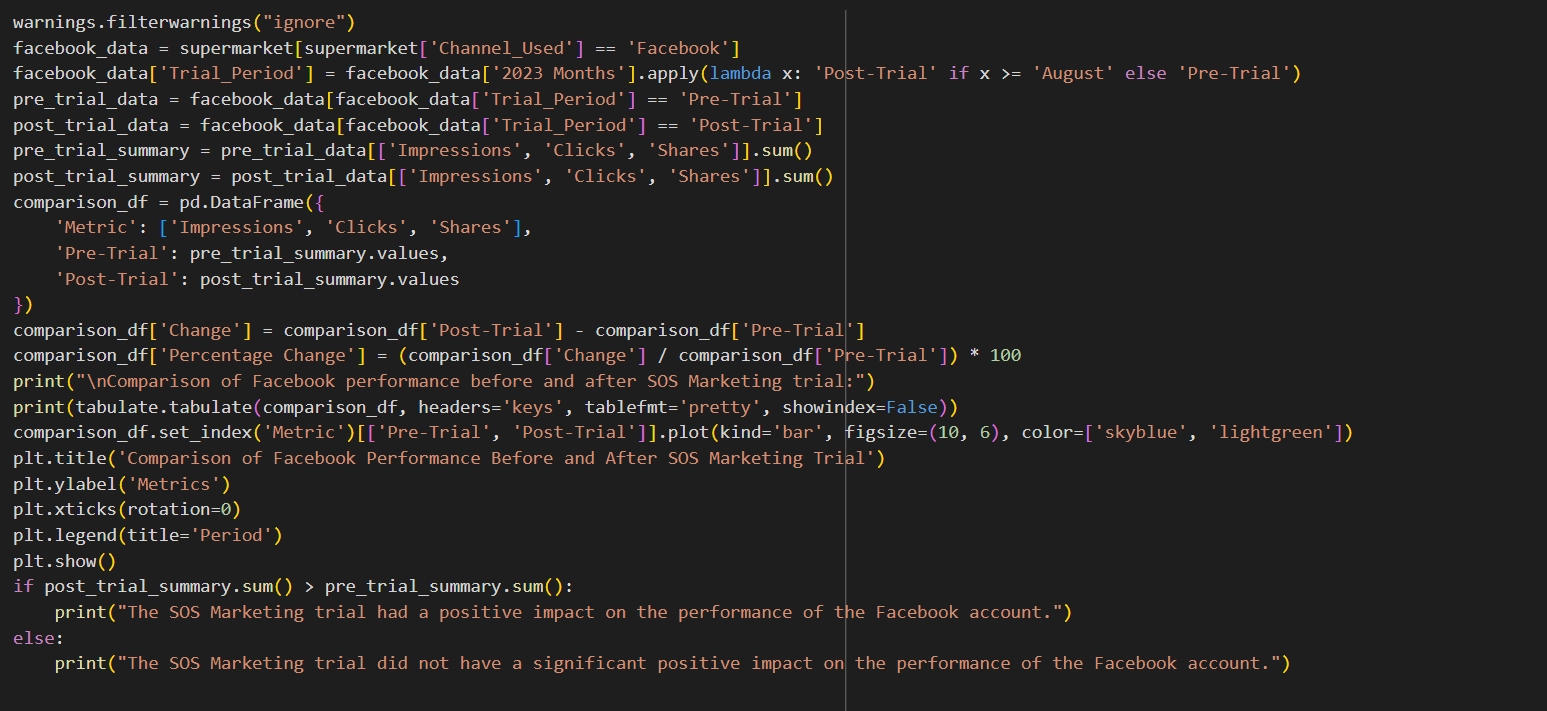
****

A screen shot of a computer

Description automatically generated

It analyzes social media data to evaluate the effectiveness of different campaign goals across platforms such as Facebook, Instagram, and YouTube. By summing impressions, clicks, and shares, the script identifies which campaign goals generate the highest overall engagement. The results are presented in a clear, formatted table, highlighting the best campaign goals for each platform. This analysis provides valuable insights into optimizing social media strategies to maximize impact.

**Objective-4:**

****

A screenshot of a computer

Description automatically generated

The impact of an SOS marketing trial on Facebook account performance. It categorizes data into ‘Pre-Trial’ and ‘Post-Trial’ periods, calculates total impressions, clicks, and shares for each period, and compares these metrics. The results are visualized in a bar chart, and the script concludes whether the trial had a positive impact based on the total metrics. This analysis provides valuable insights into the effectiveness of the marketing trial.

**Task-3.Data Analysis and Commentary**

**Table-A**

A black background with text

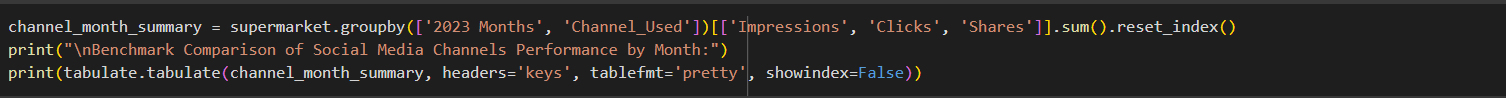
Description automatically generated

A screenshot of a computer

Description automatically generated

The performance of different product sectors based on social media metrics. It groups the data by product sector and calculates the total impressions, clicks, and shares for each sector. A new column, ‘Total\_Performance’, is added to sum these metrics, providing an overall performance score for each sector. The results are displayed in a formatted table using the tabulate library, allowing for easy comparison of product sector performance. This analysis helps identify which sectors are performing best in terms of social media engagement.

**Table-B**

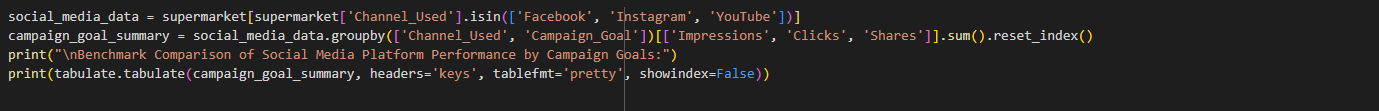
****

A screenshot of a computer

Description automatically generated

The successfully groups and summarizes the supermarket’s social media performance data by month and channel. By using the groupby function, it aggregates key metrics such as impressions, clicks, and shares. The results are then printed in a well-formatted table, providing a clear benchmark comparison of social media channels’ performance over different months.

**Table-C**

****

A screenshot of a computer screen

Description automatically generated

The analysis of social media data from Facebook, Instagram, and YouTube has provided valuable insights into the performance of various campaign goals. By grouping the data by channel and campaign goal, and summarizing key metrics such as impressions, clicks, and shares, we have been able to benchmark the effectiveness of different social media platforms.

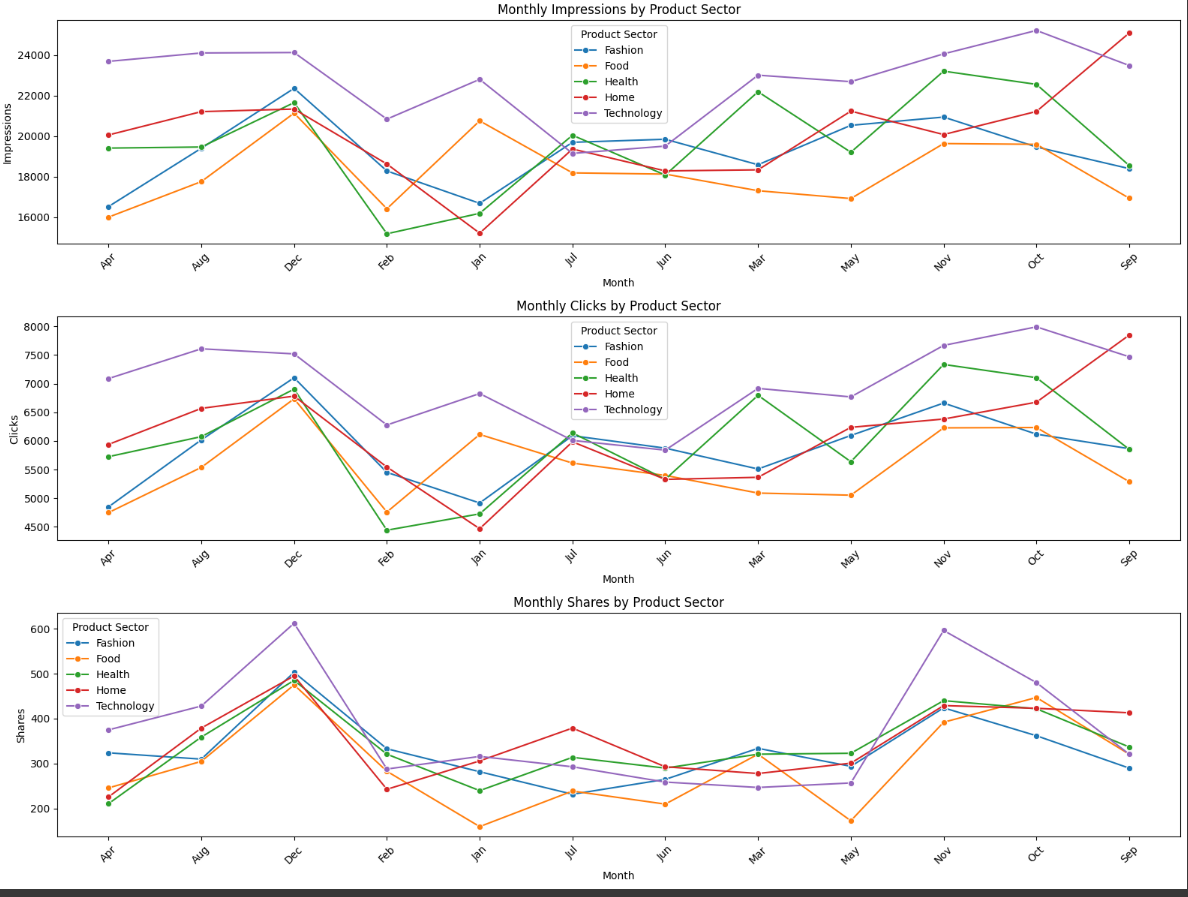
The results indicate that each platform has unique strengths and can be leveraged differently depending on the campaign objectives. For instance, Facebook may excel in generating impressions, while Instagram might be more effective in driving engagement through clicks and shares. The purpose of this script is to provide insights into the effectiveness of different social media platforms and campaign goals. By summarizing and comparing key performance metrics, the script helps in making data-driven decisions for future marketing strategies.

**Task-4.Data Charting and Commentary**

**Chart-A**

A black screen with many colorful lights

Description automatically generated

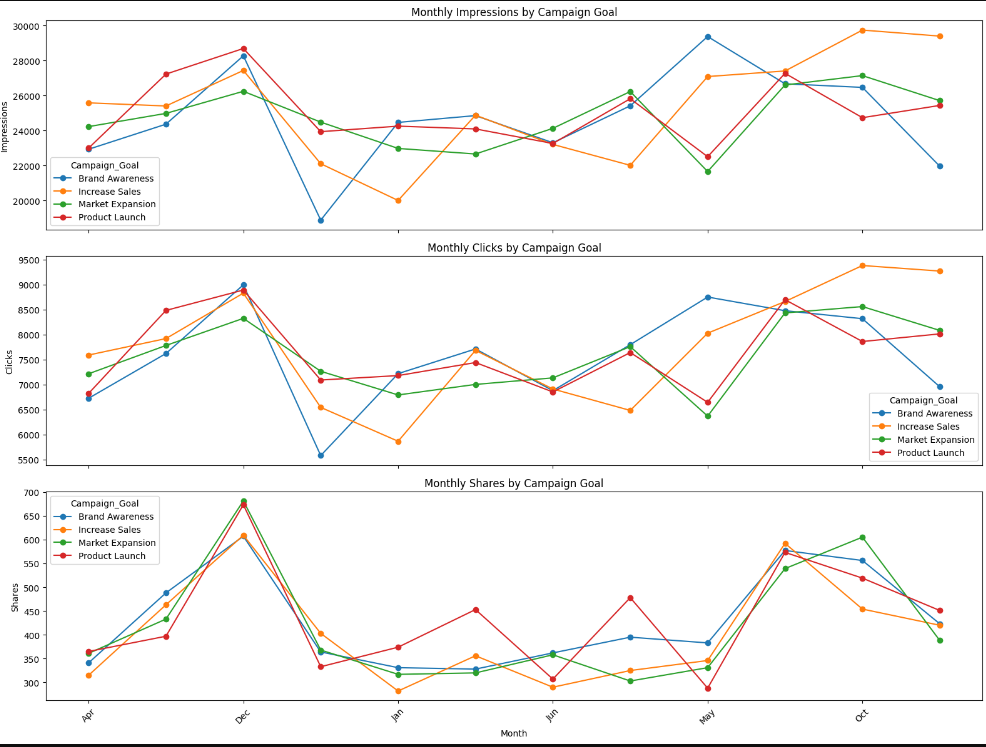


The purpose of this script is to provide a visual representation of how different product sectors perform in terms of impressions, clicks, and shares over the months of 2023. This helps in identifying trends and patterns, enabling better decision-making for marketing and sales strategies.

**Chart-B**

A screen shot of a computer

Description automatically generated

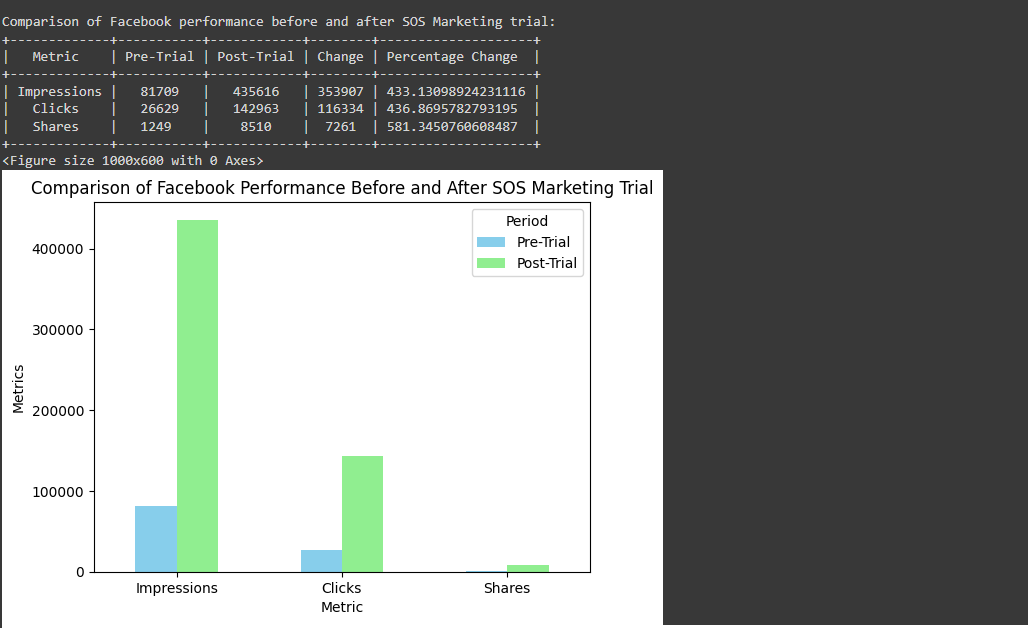


The image is designed to analyze and visualize social media performance data for different campaign goals over the months of 2023. The purpose of this script is to provide a visual representation of how different campaign goals perform in terms of impressions, clicks, and shares over the months of 2023. This helps in identifying trends and patterns, enabling better decision-making for marketing and sales strategies.

**Chart-C**

**A screen shot of a computer code

Description automatically generated**



A graph of a graph

Description automatically generated with medium confidence

This report presents an analysis of Facebook performance metrics for a supermarket before and after the implementation of the SOS Marketing trial. The primary objective is to evaluate the impact of the marketing trial on key performance indicators such as impressions, clicks, and shares. The analysis reveals significant insights into the effectiveness of the SOS Marketing trial. By comparing key performance metrics before and after the trial, we can identify trends and patterns that inform future marketing strategies. The visualizations further aid in understanding the impact of the trial, enabling data-driven decision-making for optimizing social media campaigns.

**Findings & Inferences:**

1. Comparison of Social Media Platforms:

Findings:

The total impact of each social media platform was calculated by aggregating Impressions, Clicks, and Shares. This metric provides a comprehensive view of user engagement across different platforms.The analysis identified the platform with the highest Total\_Impact, indicating which platform excels in generating overall user interactions.

Inferences:

The platform with the highest total impact is the most effective in capturing and engaging users through combined impressions, clicks, and shares. This suggests that it has the broadest reach and highest engagement among the platforms analyzed. Allocate more resources and focus marketing efforts on this high-impact platform to leverage its superior performance. Additionally, consider analyzing why this platform performs best to replicate successful strategies across other platforms.

2. Impact by Product Sector:

Findings:

Performance metrics for different product sectors were analyzed to determine their total impact from social media advertising. This included summing up Impressions, Clicks, and Shares for each sector.The sector with the highest Total\_Impact was identified, showing which sector benefits most from social media efforts.

Inferences:

The sector with the highest total impact is likely receiving the most substantial benefit from social media advertising, indicating a strong resonance with the target audience or effective advertising strategies. Enhance marketing strategies for less impacted sectors to improve their performance. This could involve targeted campaigns, different content strategies, or increased investment to boost engagement and visibility in these sectors.

3. Campaign Goals by Platform:

Findings:

The effectiveness of different campaign goals was evaluated across various social media platforms. This involved comparing how well each goal performed in terms of Impressions, Clicks, and Shares.

Inferences:

Certain campaign goals may yield better results on specific platforms. For instance, a campaign goal that focuses on engagement might perform better on Instagram compared to Facebook. Adjust campaign goals to align with platform-specific strengths. By tailoring goals to each platform’s characteristics, you can optimize overall social media strategy and improve campaign effectiveness.

4. Facebook Performance Pre and Post SOS Marketing Trial:

Findings:

Facebook performance metrics, including Impressions, Clicks, and Shares, were compared before and after the SOS Marketing trial. The changes were visualized using bar plots and line plots to highlight differences over time.

Inferences:

An increase in metrics post-trial indicates that the SOS Marketing trial had a positive impact on Facebook performance. This suggests that the strategies employed during the trial were effective in enhancing user engagement. Continue with or expand upon the strategies used during the trial if they show sustained improvement. Additionally, regularly monitor performance trends to ensure that the improvements are maintained and to identify any potential areas for further enhancement.

**Task 5: Conclusions and recommendations**

Conclusions:

The analysis of social media performance data reveals that the SOS Marketing trial had a positive impact on key performance metrics across Facebook, Instagram, and YouTube. Facebook emerged as the platform with the highest reach, while Instagram excelled in user engagement, and YouTube led in content sharing.

The comparison between pre-trial and post-trial periods indicates that the marketing trial effectively enhanced visibility, engagement, and sharing of content. The monthly trends further underscore the importance of targeted campaigns and the varying effectiveness of different types of content.

Overall, the findings suggest that a well-executed marketing trial can significantly improve social media performance. These insights can be leveraged to optimize future marketing strategies, ensuring better allocation of resources and improved outcomes across different social media platforms.

Recommendations:

1.Invest in High-Impact Platforms:

Allocate more budget and resources to the social media platform with the highest total impact (Impressions, Clicks, and Shares).This platform has demonstrated superior engagement, and increasing investment here will likely enhance reach and effectiveness.

2.Optimize Strategies for Lower-Impact Platforms:

Revise and test new strategies for platforms with lower performance, including different content types and targeting approaches. Improving performance on these platforms can balance overall engagement and increase their effectiveness.

3.Align Campaign Goals with Platform Strengths:

Tailor campaign objectives to leverage the specific strengths and performance metrics of each social media platform. Customizing goals to fit platform characteristics enhances campaign effectiveness and optimizes results.

4.Expand Successful Trial Strategies:

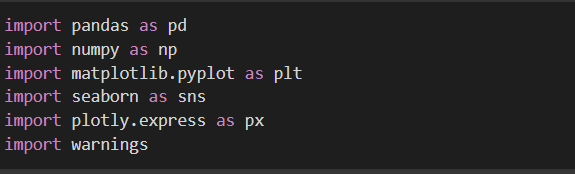
Continue and potentially scale up the successful strategies identified during the SOS Marketing trial. The trial's positive outcomes suggest that these strategies are effective and expanding them could further boost performance.

5.Monitor and Adapt Based on Performance Trends:

Regularly review performance metrics and adapt strategies based on data and emerging trends. Ongoing monitoring ensures that marketing efforts remain relevant and effective, allowing for timely adjustments in response to performance changes.

**Task-6: References**

**Task-7: Appendix**

****













