**SALES DATA ANALYSIS**

Project Description:

In this comprehensive sales data analysis project, Tableau software played a pivotal role in visualizing intricate sales patterns, regional trends, profit analysis, and seasonal variations. Through an array of detailed visualizations, including bar graphs, scatter plots, world maps, and interactive dashboards, the project adeptly highlighted the distribution of profits across various product types, uncovering seasonality trends in sales data, and identifying the top item types, regions, and countries by revenue and profit. It also sheds light on the distribution dynamics between online and offline sales channels, offering a nuanced understanding of consumer behaviour. Alongside Tableau, SQL was instrumental in data manipulation, enabling the extraction of valuable insights through sophisticated queries. These insights were then elegantly presented in Tableau, enhancing the visual appeal and interpretability of the data. This strategic combination of SQL and Tableau not only facilitated a deeper comprehension of the dataset but also empowered decision-making processes, highlighting potential areas for operational improvement and growth opportunities, thereby laying the groundwork for informed strategic planning and optimization in sales and marketing efforts.

Data Set:

Dataset Description: The dataset, available as the Sample Country Sales Dataset on Kaggle, captures a wide spectrum of sales-related details from various countries, product categories, and sales channels. It serves as a comprehensive repository of transactional insights, providing a holistic view of sales performance and market dynamics for strategic decision-making.

Featuring an array of information, the dataset includes:

* Region: Categorical segmentation of sales territories across distinct geographical areas like Sub-Saharan Africa, Europe, Asia, the Middle East, North Africa, Central America, and the Caribbean.
* Country: Character-based representation of countries involved in the sales, showcasing a diverse global market landscape.
* Item type: Categorical classification of products into different categories such as Meat, Fruits, Cosmetics, Vegetables, and Personal Care.
* Sales Channel: Character-based differentiation of sales mediums, distinguishing between online and offline channels.
* Order Priority: Classification indicating the priority levels of orders, categorized as High, Medium, or Low.
* Order date & Ship date: Date variables in MM/DD/YYYY format detailing order placement and shipment timelines.
* Order ID: Integer-based unique identifiers for each order, aiding in tracking and management.
* Units sold: Quantitative data representing the quantity of products sold in individual transactions.
* Unit Price, Unit Cost, Total Revenue, Total Cost, Total Profit: Decimal variables revealing financial aspects such as unit prices, costs, revenues, incurred expenses, and overall profits.

This dataset offers a comprehensive array of sales insights, allowing in-depth analysis of product performance, financial profitability, and regional market trends. Its diverse nature supports extensive exploratory analysis, predictive modelling, and informed decision-making in sales strategies and business operations.

Data Cleaning and Data Inspection:

The data cleaning process involves identifying and rectifying inconsistencies and errors within the dataset to ensure its accuracy and reliability for analysis. This involves several critical steps:

Handle Missing Data: Initially, the dataset was reviewed for any missing values in each column. No null values were found, ensuring completeness, and eliminating the need for further data imputation or deletion.

Categorical Data Validation: Categorical data about columns like 'Sales Channel', 'Region', 'Order Priority', and 'Item Type' were visualized using bar charts. This allowed a comprehensive understanding of category distributions and aided in data validation.

Address Inconsistent Data: Any inconsistencies or errors in the dataset were investigated and rectified to maintain data consistency and accuracy throughout the analysis.

Categorical Data Validation: Categorical data about columns like 'Sales Channel', 'Region', 'Order Priority', and 'Item Type' were visualized using bar charts. This allowed a comprehensive understanding of category distributions and aided in data validation.

**RESEARCH QUESTIONS:**

Question 1:

*Identification of products which are creating losses?*

Answer:

The assessment revealed that no products exhibit losses (profits less than zero), as depicted in the accompanying graph. However, certain products notably yield lower median profits. These products are represented in a light pink shade in the R-generated graph while varying shades of purple in the Tableau-generated graph indicate the same set of lower-profit products.

This analysis helps visualize and comprehend the distribution of profits across various product types, highlighting those with lower median profits. Although no products operate at a loss, these insights aid in identifying items contributing comparatively less to the overall profitability.

A graph with different colored bars

Description automatically generated

Figure 1

A graph of a bar graph

Description automatically generated with medium confidence

Figure 2

Question 2:

*Identification of Seasonality Patterns in sales data and purchasing behavior?*

Answer:

The analysis of sales data revealed consistent purchasing behavior among consumers across various regions. Sub-Saharan Africa and Europe maintained their dominance in sales, representing the most prominent markets. However, a notable shift occurred by the year 2017, indicating a decline in profits not only within these regions but also across others. Consequently, this decline had a direct impact on the company's overall profits, reducing them below the 1.5 billion mark.

A graph showing a number of green bars

Description automatically generated with medium confidence

Figure 3

A red and yellow bar chart

Description automatically generated

Figure 4

Question 3:

*What are the top item types, top regions and top countries with higher revenue and highest profit?*

Answer:

Household Items and Office Supplies are the primary contributors to revenue, led by top revenue-generating countries: China, Libya, Congo, Gabon, Mexico, and Guinea. These same countries are also responsible for the company's highest profits.

A graph with orange and red bars

Description automatically generated

Figure 5

A map of the world

Description automatically generated A map of the world

Description automatically generated

Figure 6 Figure 7

Question 4:

*What is the distribution of sales between online and offline channels, and how can I devise distinct strategies to enhance sales for each order type?*

Answer:

Although revenue varies across countries, both online and offline sales contribute equally. To enhance sales in both channels, strategies like targeted marketing campaigns, improving the online user experience, expanding physical store presence, and providing unique promotions for each platform can be implemented.

A graph showing a number of columns

Description automatically generated

Figure 8

A screenshot of a map

Description automatically generated A screenshot of a map

Description automatically generated

Figure 9 Figure 10

Question 5:

*Identify the regions with growth potential in terms of revenue and profit.*

Answer:

Examining the earlier findings, it's evident that Sub-Saharan Africa and Europe have notably boosted company profits. Nevertheless, unexplored growth opportunities lie in other regions, particularly Asia. With high internet penetration in Asia and the Middle East, it's advisable for the company to expand its reach in these areas, reducing reliance on Europe and Sub-Saharan Africa.

**Data Analysis**

Software’s Used-

1. Tableau Software:

Employed Tableau for data visualization purposes. Created insightful visualizations, including charts, graphs, and dashboards, to represent sales patterns, regional trends, profit analysis, and seasonal variations. Leveraged Tableau's interactive features for intuitive and comprehensive data representation, aiding in easier comprehension and strategic decision-making processes.

2. MySQL workbench:   
MySQL was utilized for its robust data manipulation capabilities. By writing and executing sophisticated SQL queries, you were able to sift through extensive datasets, extracting specific insights critical to the analysis. This process involved filtering, sorting, and aggregating data to identify trends, patterns, and anomalies within the sales information. For instance, queries might have been crafted to calculate the total revenue and profit per product type, identify seasonal sales peaks, or determine the performance of different sales channels. The results from these queries provided a structured dataset that served as the foundation for your analysis. This pre-processed data was then imported into Tableau, where it was transformed into a series of compelling visualizations. Through this methodical use of MySQL for data preparation and Tableau for visualization, you effectively bridged the gap between raw data and actionable insights, enabling a deeper understanding of the sales dynamics at play and informing strategic decision-making.