ONLINE RETAILS DATA ANALYSIS EXCEL PROJECT

Objective: The primary objective of this project is to generate a comprehensive annual sales report for the year 2022, specifically tailored for multiple online stores. The purpose of this project is to provide valuable insights into customer behavior and purchasing patterns, enabling the stores to gain a deep understanding of their customer base and leverage these insights to drive growth and success in the year 2023.

By analyzing and interpreting the sales data from 2022, the project aims to uncover key trends, identify customer preferences, and highlight growth opportunities for the online stores. The report will delve into various aspects, including customer demographics, purchasing habits, popular product categories, preferred channels, and regional sales performance.

Through this data-driven approach, the online stores will gain actionable insights that will facilitate strategic decision-making, targeted marketing campaigns, and the development of customer-centric strategies for sustainable growth. By understanding their customers better and aligning their business efforts accordingly, the stores can optimize their operations, enhance customer satisfaction, and foster long-term loyalty, ultimately leading to increased revenue and market competitiveness in 2023.

Datasheet used in the project: https://bit.ly/3X381ok

Columns In Datasheet:

Here is a short description for each of the columns in the datasheet.

Index: An index column that assigns a unique identifier to each row in the dataset.

Order ID: A unique identifier assigned to each order made by customers.

Cust ID: A unique identifier for each customer who placed an order.

Gender: Indicates the gender of the customer (e.g., Male, Female).

Age: Represents the age of the customer.

Date: The date when the order was placed.

Status: The status of the order (e.g., pending, processing, shipped, delivered, etc.).

Channel: The channel through which the order was received (e.g., Amazon, Ajio, Myntra etc.).

SKU: Stock Keeping Unit, a unique identifier for a specific product.

Category: The category or type of the product (e.g., electronics, clothing, accessories, etc.).

Size: The size or dimensions associated with the product (if applicable).

Qty: The quantity of a particular product ordered in an order.

Currency: The currency used for the transaction.

Amount: The total amount or cost associated with the order.

Ship-city: The city where the order is being shipped to.

Ship-state: The state or province where the order is being shipped to.

Ship-postal-code: The postal code or ZIP code of the shipping address.

Ship-country: The country where the order is being shipped to.

B2B: Indicates whether the order is a business-to-business transaction (B2B).

DATA CLEANING

Gender: The "Gender" column contains inconsistent values such as "Women," "Men," "W," and "M." To standardize the values, the occurrences of "W" will be replaced with "Women," and "M" will be replaced with "Men." This will create a more uniform and understandable representation of gender in the dataset.

Qty: The "Qty" column exhibits inconsistencies with values such as "1," "One," "2," and "Two." To standardize the representation, "One" will be replaced with "1," and "Two" will be replaced with "2." This will ensure a consistent numerical representation of quantities.

It is important to note that the rest of the columns in the dataset do not contain any null values, indicating a good quality of data.

By performing these data cleaning operations, we enhance the consistency and uniformity of the dataset, facilitating accurate analysis and interpretation of the data.

DATA PROCESSING

Age Group Column Creation:

A new column called "Age Group" was created to categorize individuals based on their age. The categorization criteria are as follows:

If the age is greater than or equal to 50, the value in the "Age Group" column is set as "Senior."

If the age is greater than or equal to 25, the value in the "Age Group" column is set as "Adult."

For all other cases, indicating ages below 25, the value in the "Age Group" column is set as "Teenage."
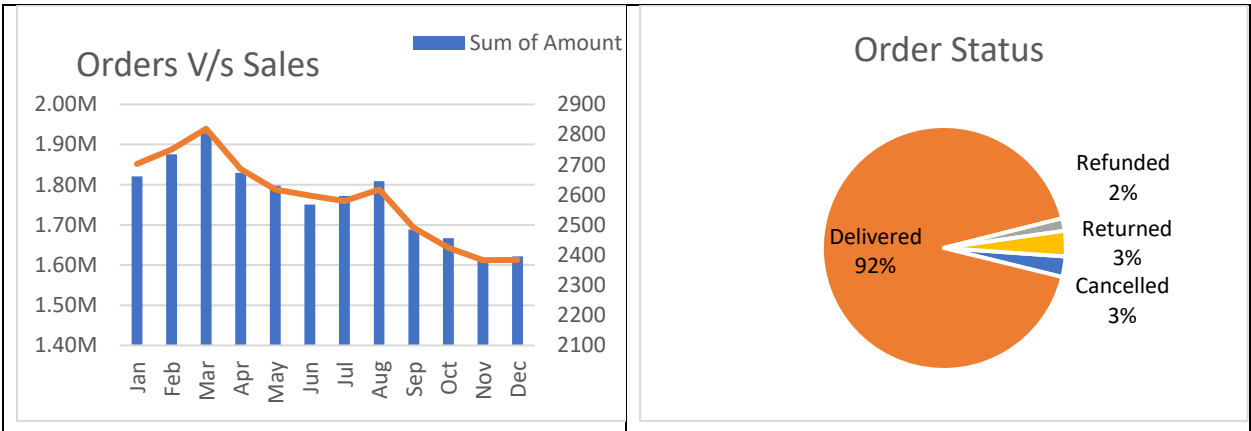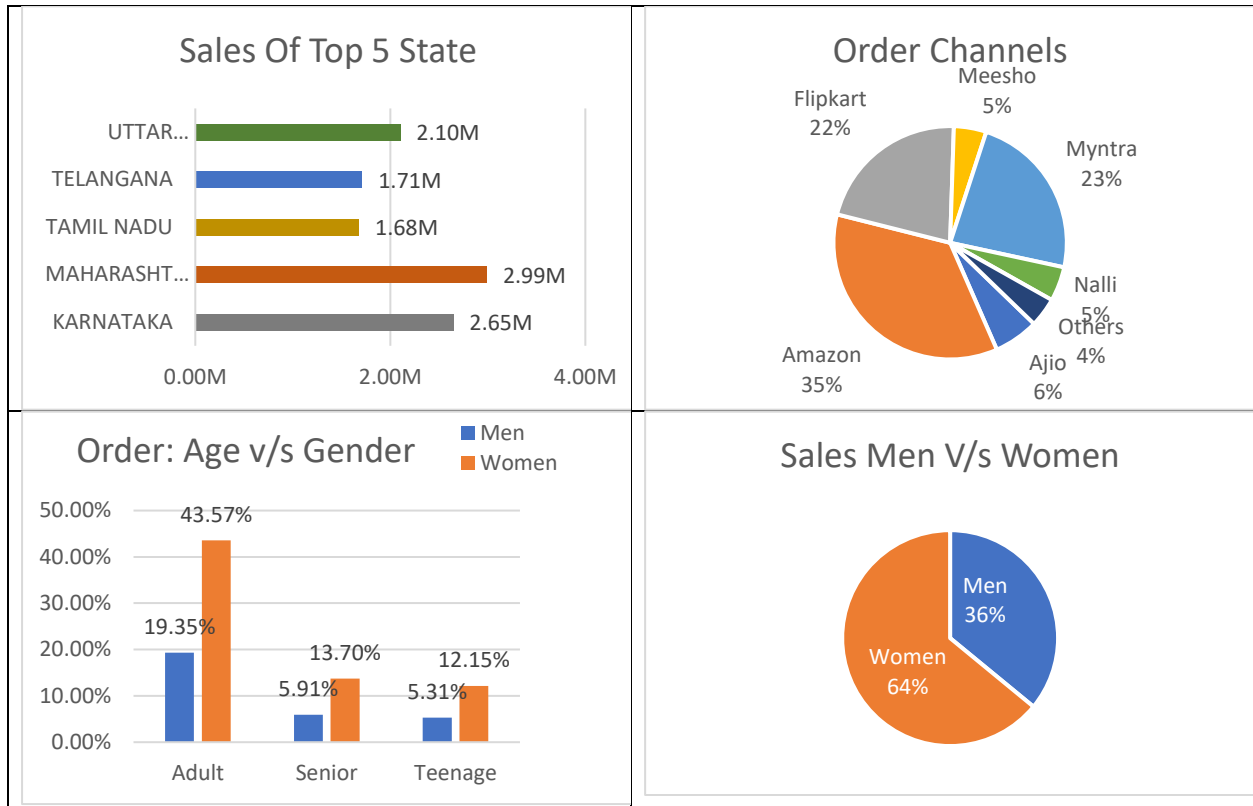
Month Extraction:

The month component was extracted from the date column to facilitate further analysis. This extraction enables the evaluation of data based on a monthly basis, providing insights into seasonal patterns and trends.

These data processing steps aim to enhance the analysis and segmentation of the dataset, enabling more nuanced exploration and understanding of the information at hand.
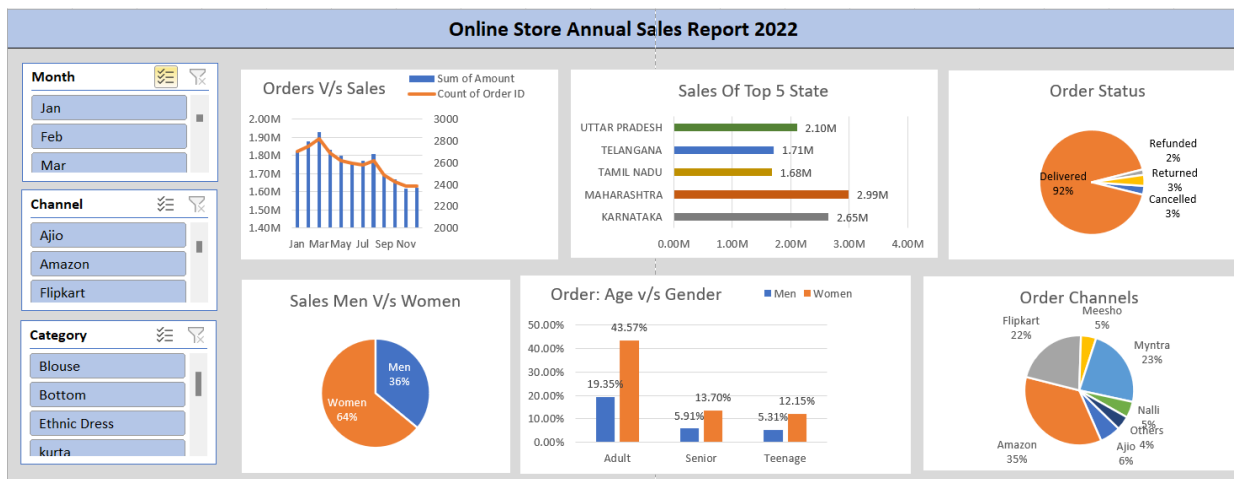
DATA ANALYSIS

In order to conduct comprehensive data analysis, several pivot tables were generated. Pivot tables are powerful analytical tools that enable efficient summarization, grouping, and calculation of data based on various dimensions and measures. The creation of these pivot tables allowed for in-depth exploration and interpretation of the dataset, yielding **valuable insights, and facilitating data-driven decision-making**.

Sales Of Top 5 State

State	Sales
UTTAR...	2.10M
TELANGANA	1.71M
TAMIL NADU	1.68M
MAHARASHT...	2.99M
KARNATAKA	2.65M

Order Channels

- Flipkart 22%
- Meesho 5%
- Myntra 23%
- Nalli 5%
- Others 4%
- Ajio 6%
- Amazon 35%

Order: Age v/s Gender

- Adult: Men 19.35%, Women 43.57%
- Senior: Men 5.91%, Women 13.70%
- Teenage: Men 5.31%, Women 12.15%

Sales Men V/s Women

- Men 36%
- Women 64%

FINAL INSIGHTS

Online Store Annual Sales Report 2022

Based on the final insights derived from the data analysis, the following conclusions can be drawn:

- **Women are more likely to make purchases compared to men**, indicating a higher potential for customer acquisition and retention by targeting female customers specifically.
- The states of **Maharashtra, Karnataka, and Uttar Pradesh** emerge as the top three contributors to sales, highlighting the significance of focusing marketing efforts and resources in these regions for maximum impact and market penetration.
- The adult age group, specifically **individuals aged between 30 and 50 years, contributes the most to sales**, suggesting that tailoring marketing campaigns and product offerings to cater to the preferences and needs of this age segment can yield favorable results.
- Among the various channels, **Amazon, Flipkart, and Myntra are the primary contributors** to sales. Therefore, allocating resources and promotional activities on these platforms can effectively reach the target audience and maximize sales potential.

FINAL CONCLUSION:

- To optimize business growth and increase sales, it is recommended to target women customers within the age range of 30 to 50 years residing in Maharashtra, Karnataka, and Uttar Pradesh.
- By strategically advertising and offering tailored promotions on platforms such as Amazon, Flipkart, and Myntra, the potential to attract and retain these target customers can be significantly enhanced.
- By focusing efforts on these key segments and channels, the business can position itself for success and capitalize on the identified market opportunities.

Datasheet used in the project: https://bit.ly/3X381ok

Project GitHub link: https://github.com/niyanta02/Online-Store-Data-Analysis-Project