**Megastore Sales Dashboard (Fictional Company)**

**By: Jillian Ireland**

*Project Overview*

Customer Transaction Analysis: Insights into Buying Patterns and Customer Segmentation

**Introduction:**

This project analyzes a dataset of customer transactions to uncover buying patterns, customer segmentation, and product performance. The goal is to provide actionable insights that can inform business decisions, such as targeting high-value customers and optimizing product offerings. The analysis was conducted using Tableau, and the project is documented here for sharing on GitHub.

**Source:**

The dataset was sourced from Kaggle, consisting of transaction records from an online retail store. It contains a wide range of information, from customer details to product purchases.

*Dataset*

**Dataset Details:**

The dataset includes the following key columns:

- Transaction\_ID: Unique identifier for each transaction.

- Customer\_ID: Unique identifier for each customer.

- Name: Customer's name.

- Email: Customer's email address.

- Phone: Customer's phone number.

- Address: Customer's address.

- City: City of residence.

- State: State of residence.

- Zipcode: Postal code.

- Country: Country of residence.

- Age: Customer's age.

- Gender: Customer's gender.

- Income: Annual income of the customer.

- Customer\_Segment: The segment to which the customer belongs.

- Date: Date of the transaction.

- Year: Year of the transaction.

- Month: Month of the transaction.

- Time: Time of the transaction.

- Total\_Purchases: Total number of purchases made by the customer.

- Amount: Amount spent in the transaction.

- Total\_Amount: Total amount spent by the customer.

- Product\_Category: Category of the product purchased.

- Product\_Brand: Brand of the product purchased.

- Product\_Type: Type of product purchased.

- Feedback: Customer feedback on the product.

- Shipping\_Method: Method used for shipping.

- Payment\_Method: Method used for payment.

- Order\_Status: Status of the order (e.g., completed, pending).

- Ratings: Customer rating for the transaction.

*Data Cleaning and Preparation*

**Cleaning Process:**

To ensure the accuracy of the analysis, the dataset was cleaned to handle missing values, remove duplicates, and format columns. Specific steps included:

- Missing Values: Any rows with missing essential data (such as `Customer\_ID` or `Amount`) were removed.

- Data Formatting: Columns such as `Date`, `Year`, `Month`, and `Time` were formatted to ensure consistency.

- Data Integrity: Duplicates were identified and removed to maintain the integrity of the dataset.

**Preparation for Analysis:**

The data was prepared by creating new calculated fields in Tableau, such as:

- Total Revenue: Calculated as `Amount` multiplied by the number of purchases.

- Profit: Calculated as SUM([Total Sales])- SUM([Amount])

- Custom ID: Calculated the Country Abbreviation and the Transaction ID : [Country Abr] + "-" + STR(INT([Transaction ID]))

- Customer Lifetime Value (CLV): An estimate of the total value a customer will bring to the business over time.

*Analysis and Visualization*

**Tableau Dashboard:**

The Tableau dashboard created for this analysis includes the following components:

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated

1. Summary:

- A demographic breakdown of the customer base by location.

- Insights into which countries are most active and have the highest spending.

- KPI cards of Total Sales, Profit, Products Sold, Customers and Average Ratings

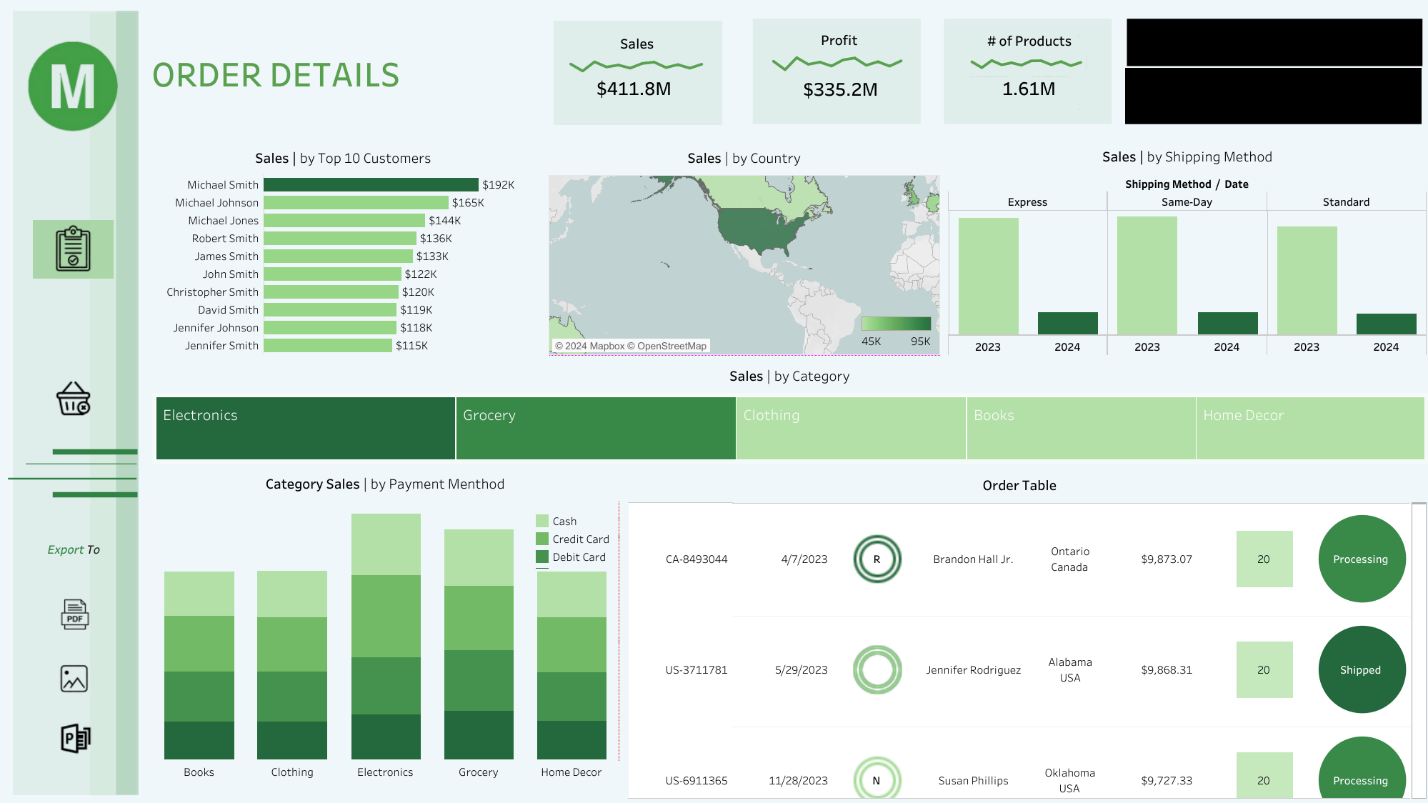
- Top ten customers by sales revenue

- A time series analysis of transactions over two years, segmented by month.

- Visualizations showing peak shopping times and months with the highest sales.

- Sales by Top 10 brands

A screenshot of a computer

Description automatically generated

2. Order Details:

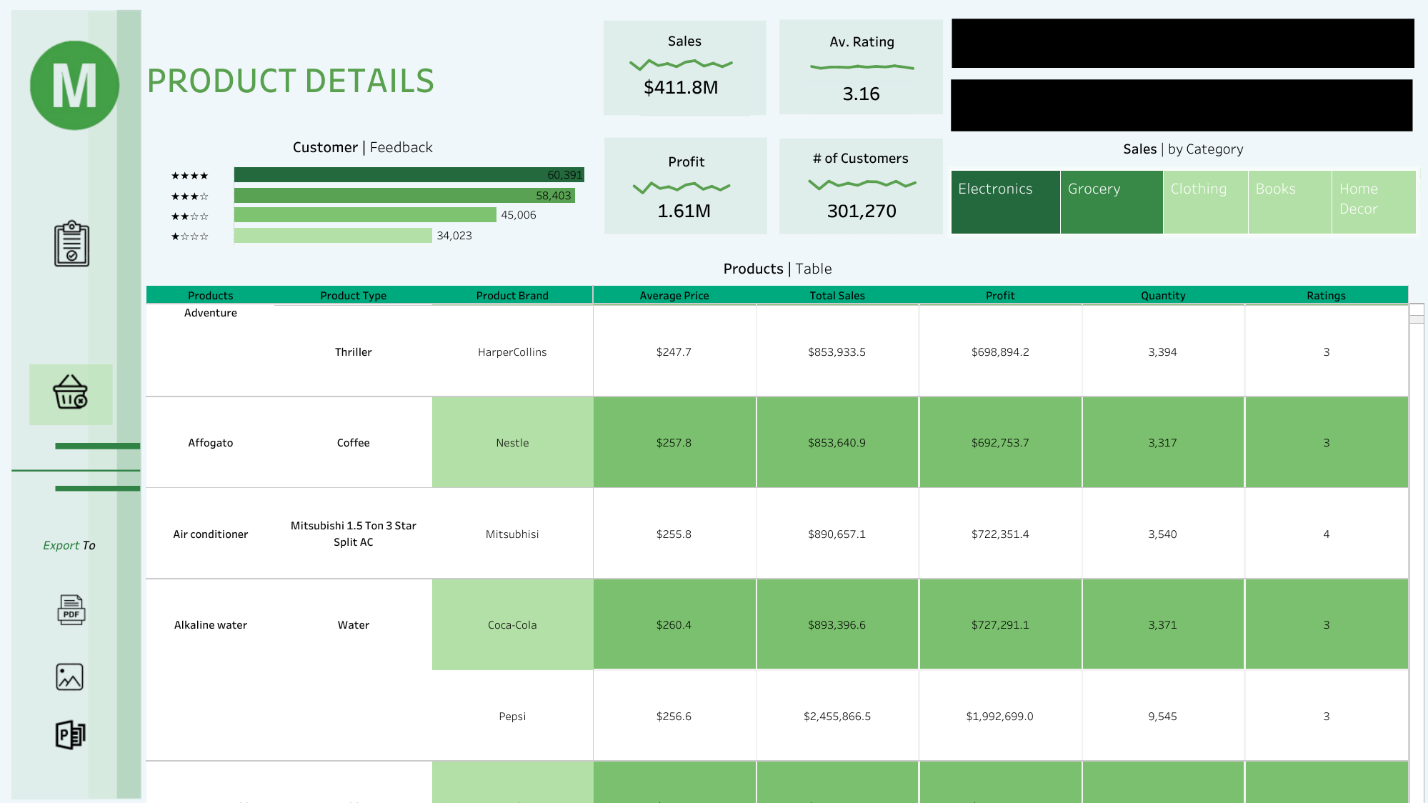
- KPI cards of Total Sales, Profit and Products Sold

- Order Table depicting customer demographics, segment , purchases and order status

- A financial breakdown of Category Sales by Payment Method

- An overview of Sales by Shipping Methods in two years

A screenshot of a computer

Description automatically generated

3. Product Performance:

- A comparison of product categories, brands, and types in terms of sales volume and revenue.

- KPI cards of Total Sales, Profit, Customers and Average Ratings

- Identification of top-performing products and underperforming categories.

**Key Metrics**

- Total Revenue

- Total Profit gained

- Top Product Categories

- Average Ratings

- Total Customers

Filters and Interactivity:

- The dashboard includes filters for `Year` and 'Country', allowing users to drill down into specific areas of interest.

*Insights and Conclusions*

**Key Findings**

- High-Value Customers: The analysis revealed a segment of customers who contribute a significant portion of the total revenue resides in the United States. These customers are primarily interested in premium products and have a higher average order value.

A screenshot of a computer

Description automatically generated

- Seasonal Trends: Sales peak during January, April, and August can be attributed to post-holiday sales, tax refunds, and back-to-school shopping, respectively, suggesting that targeted marketing during these times could boost revenue further.

A screenshot of a computer dashboard

Description automatically generated

- Product Insights: Electronics and Groceries were identified as the top-performing categories, with Pepsi, consistently leading in sales.

A screenshot of a computer

Description automatically generated

**Business Implications:**

- Targeted Marketing: Focus marketing efforts on high-value customer segments, especially during peak seasons, to maximize revenue.

- Product Strategy: Invest in popular product categories and consider phasing out underperforming products to optimize inventory.

**Challenges and Limitations**

Challenges:

- Data Quality: Some customer data had missing fields or inconsistencies, which required extensive cleaning.

- Sample Size: The dataset may not be representative of all customer behavior, limiting the generalizability of the findings.

Limitations:

- Limited Historical Data: The dataset only covers a limited timeframe, which may not fully capture long-term trends.

- Feedback and Ratings: Customer feedback and ratings data were sparse, making it difficult to draw definitive conclusions about customer satisfaction.

**References**

- Source: [Kaggle - https://www.kaggle.com/datasets/sahilprajapati143/retail-analysis-large-dataset]