



Bharatiya Vidya Bhavan's
Sardar Patel Institute of Technology
(Autonomous Institute Affiliated to University of Mumbai)
Bhavan's Campus, Munshi Nagar, Andheri (West), Mumbai-400058-India

NAME: Ayush Kedare

ROLL NUMBER: 2021300058

AIM:

Create basic charts using Tableau / Power BI / R / Python / D3.js to be performed on the dataset of Ecommerce field.

- Basic Plots - Bar chart, Pie chart, Histogram, Timeline chart, Scatter plot, Bubble plot
- Calculate Product wise sales, region wise sales or any other reports
- Write observations from each chart

DATASET:

[E-Commerce Dataset](#)

- Order_Date: The date the product was ordered.
- Aging: The time from the day the product is ordered to the day it is delivered.
- Customer_id: Unique id created for each customer.
- Gender: Gender of customer.
- Device_Type: The device the customer uses to actualize the transaction (Web/Mobile).
- Customer_Login_Type: The type the customer logged in. Such as Member, Guest etc.
- Product_Category: Product category
- Product: Product
- Sales: Total sales amount
- Quantity: Unit amount of product
- Discount: Percent discount rate
- Profit: Profit
- Shipping_cost: Shipping cost
- Order_Priority: Order priority. Such as critical, high etc.
- Payment_method: Payment method

PROCEDURE DESCRIPTION:

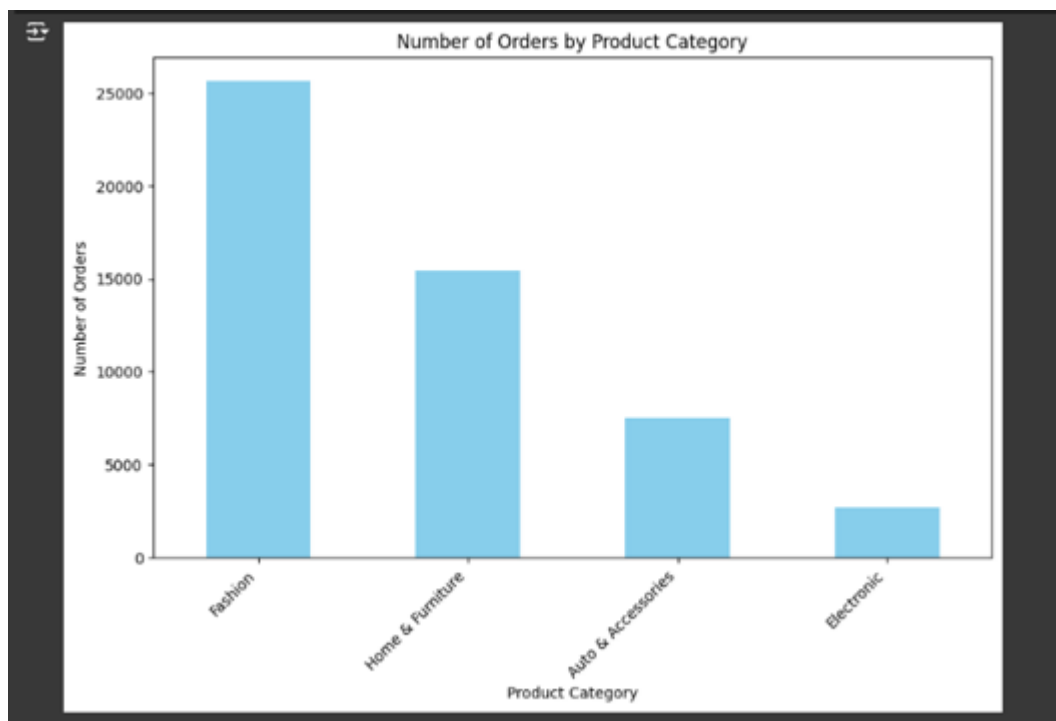
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np

df = pd.read_csv('E-commerce Dataset.csv')

[ ] df['Order_Date'] = pd.to_datetime(df['Order_Date'])
```

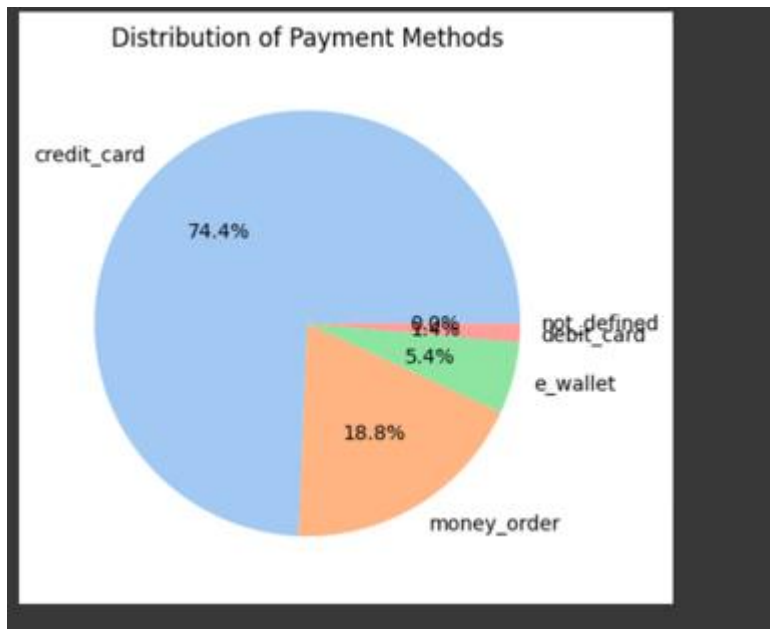
1. Bar Chart

```
[ ] plt.figure(figsize=(10, 6))
    df['Product_Category'].value_counts().plot(kind='bar', color='skyblue')
    plt.title('Number of Orders by Product Category')
    plt.xlabel('Product Category')
    plt.ylabel('Number of Orders')
    plt.xticks(rotation=45, ha='right')
    plt.show()
```



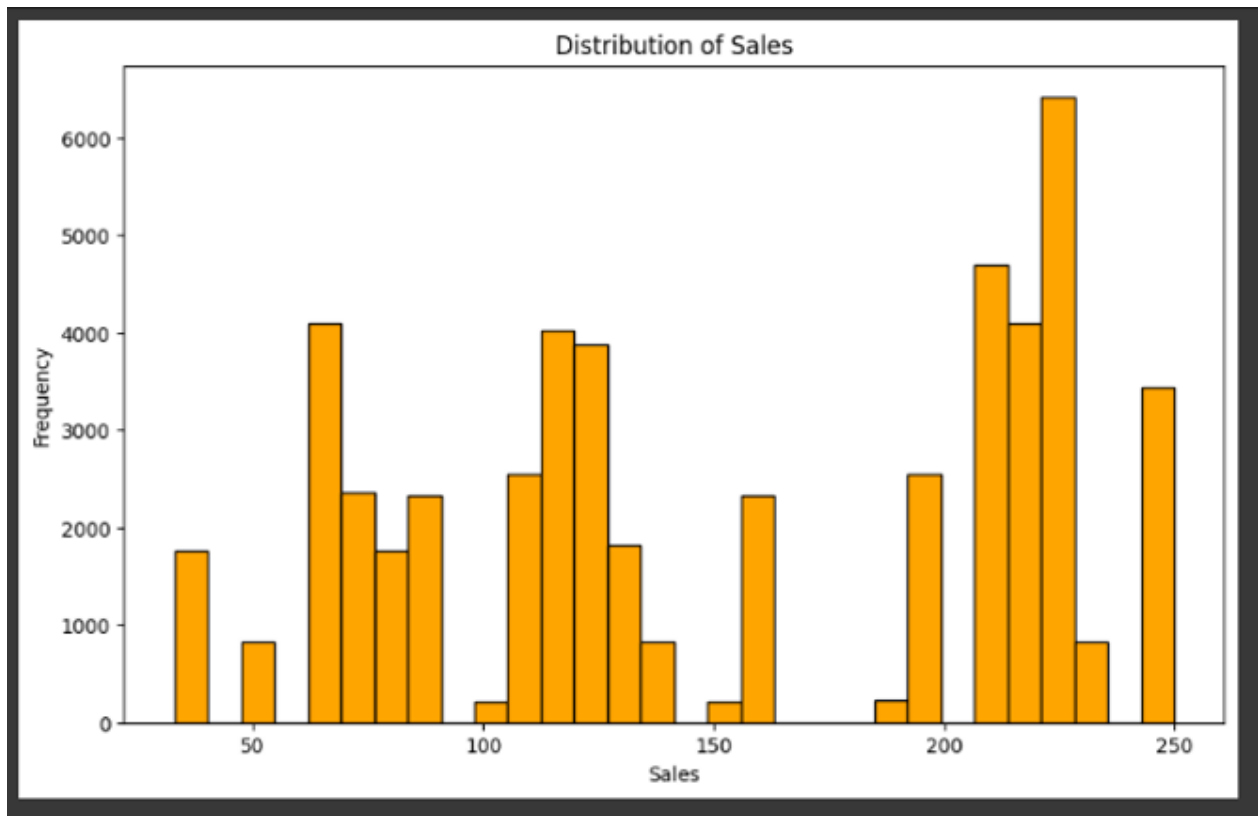
2. Pie Chart

```
[ ] df['Payment_method'].value_counts().plot(kind='pie', autopct='%1.1f%%', colors=sns.color_palette("pastel"))
plt.title('Distribution of Payment Methods')
plt.ylabel('')
plt.show()
```



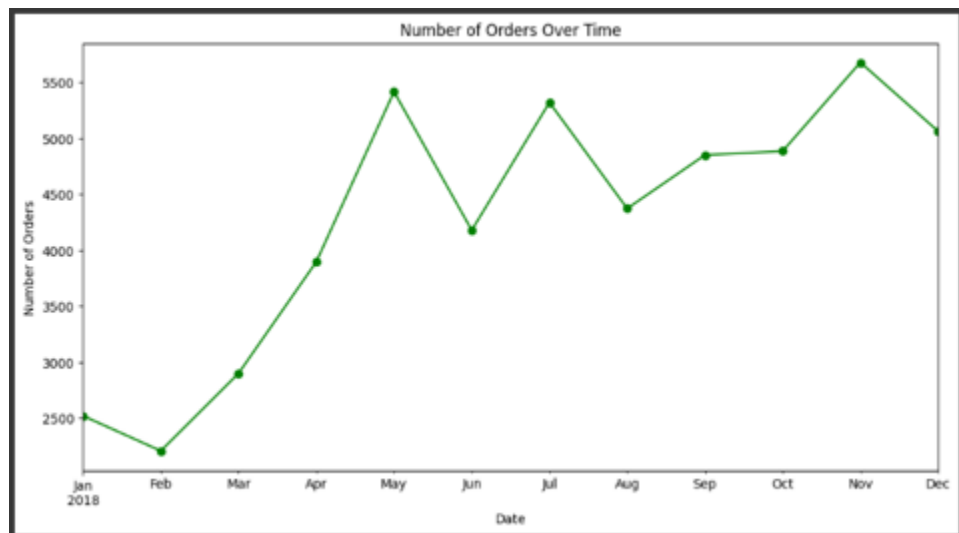
3. Histogram

```
[ ] plt.figure(figsize=(10, 6))
plt.hist(df['Sales'], bins=30, color='orange', edgecolor='black')
plt.title('Distribution of Sales')
plt.xlabel('Sales')
plt.ylabel('Frequency')
plt.show()
```



4. Timeline Chart

```
df.set_index('Order_Date').resample('M').size().plot(figsize=(12, 6), marker='o', color='green')
plt.title('Number of Orders Over Time')
plt.xlabel('Date')
plt.ylabel('Number of Orders')
plt.show()
```



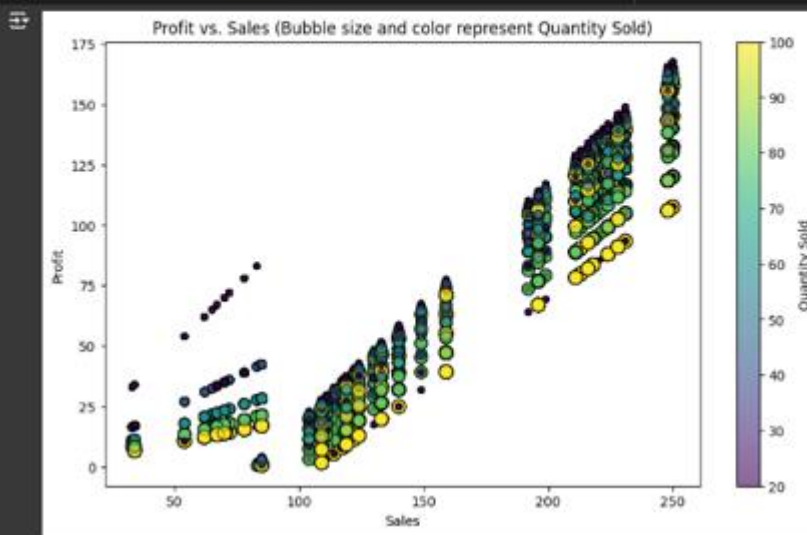
5. Scatter Plot

```
plt.figure(figsize=(10, 6))
plt.scatter(df['Sales'], df['Profit'], alpha=0.6, color='purple')
plt.title('Relationship between Sales and Profit')
plt.xlabel('Sales')
plt.ylabel('Profit')
plt.show()
```



6. Bubble Plot

```
plt.figure(figsize=(10, 6))
bubble_size = df['Quantity'] * 20
plt.scatter(df['Sales'], df['Profit'], s=bubble_size, c=bubble_size, cmap='viridis', alpha=0.6, edgecolor='black')
plt.title('Profit vs. Sales (Bubble size and color represent Quantity Sold)')
plt.xlabel('Sales')
plt.ylabel('Profit')
plt.colorbar(label='Quantity Sold')
plt.show()
```



Based on the visualizations and analysis of the e-commerce dataset, several important insights emerge:

1. **Product Category Performance:**
The bar graph indicates that certain product categories outperform others in terms of popularity. This insight can be leveraged to prioritize marketing efforts toward high-demand categories or explore strategies to boost sales in less popular areas.
2. **Customer Payment Preferences:**
The pie chart highlights the distribution of customer payment methods, showing a strong preference for certain payment options. This insight can guide the selection of payment gateway integrations or create promotions for less commonly used methods to broaden customer payment choices.
3. **Sales Value Distribution:**
The histogram reveals that the majority of transactions are of lower value, with some high-value exceptions. This skewed pattern suggests that while premium products are sold, the bulk of the revenue is driven by more affordable items. There is potential to implement strategies to raise the average transaction value.
4. **Seasonality in Orders:**
The time-series chart illustrates fluctuations in order volumes, with noticeable peaks during specific months. This insight hints at potential seasonality or the influence of promotions and events, which can be utilized for future sales planning.
5. **Sales-Profit Relationship:**
The scatter plot demonstrates a positive relationship between sales and profit, though it also reveals that not all high sales translate to high profits. This underscores the need for effective management of costs, pricing, and discounts to maintain profitability.
6. **Holistic Analysis via Bubble Plot:**
The bubble plot integrates sales, profit, and quantity sold into a single visualization. It shows that while higher quantities generally result in greater profits, the relationship is not strictly linear, indicating that factors like discounts or shipping costs may influence overall profitability.

Overall Conclusion:

This analysis provides a comprehensive overview of e-commerce performance across various dimensions, including product popularity, customer behavior, sales distribution, and profitability. The insights gained can inform the refinement of marketing strategies, the optimization of product offerings, and the enhancement of operational efficiency, ultimately driving better business results.

By pinpointing the strengths and weaknesses of current operations, the business can more effectively allocate resources, personalize customer experiences, and strategize for future growth. This data-driven approach enables the company to make informed decisions that align with both consumer preferences and evolving market trends.

Experiment submission by google form:

[https://docs.google.com/forms/d/e/1FAIpQLScStf2etXAC5WqB7d3ysvN5HF1UyvjFs832Y6-zUYAFguGc8Q/viewform?usp=sf link](https://docs.google.com/forms/d/e/1FAIpQLScStf2etXAC5WqB7d3ysvN5HF1UyvjFs832Y6-zUYAFguGc8Q/viewform?usp=sf_link)