FESTIVAL SALES ANALYSIS

1. INTRODUCTION

This project, titled **"Diwali Sales Analysis,"** aims to analyze the sales data collected during the Diwali festival season. The primary goal is to explore customer purchase behavior, identify top-selling products and regions, and uncover patterns that can help the company make better marketing and sales decisions in future festive seasons. The analysis includes data cleaning, exploration, and visualization to generate meaningful business insights that support data-driven decision making.
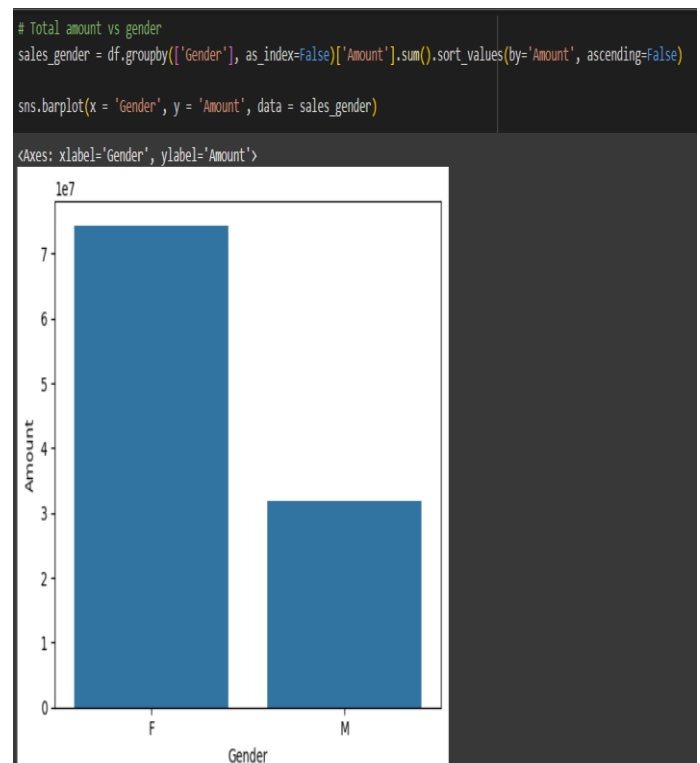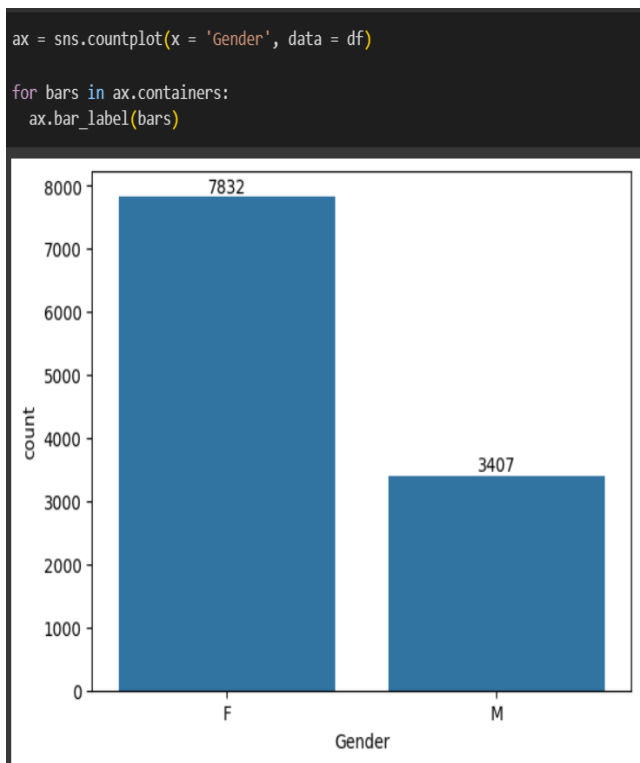
2. DATASET

LINK- https://www.kaggle.com/datasets/prajwal6362venom/diwali-sales

3. IMPORT LIBRARIES

4. LOAD DATASET

5. DATA CLEANING
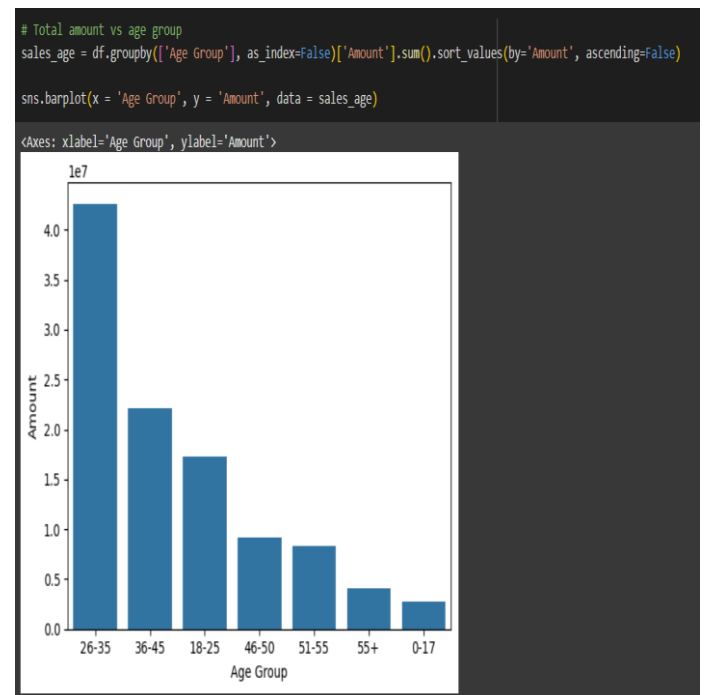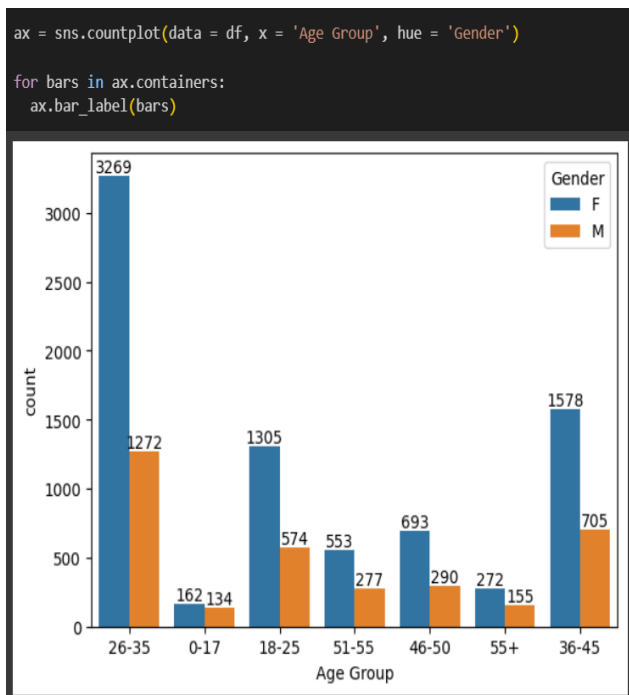
6. ANALYZE THE DATA

Exploratory Data Analysis- EDA

a. Gender

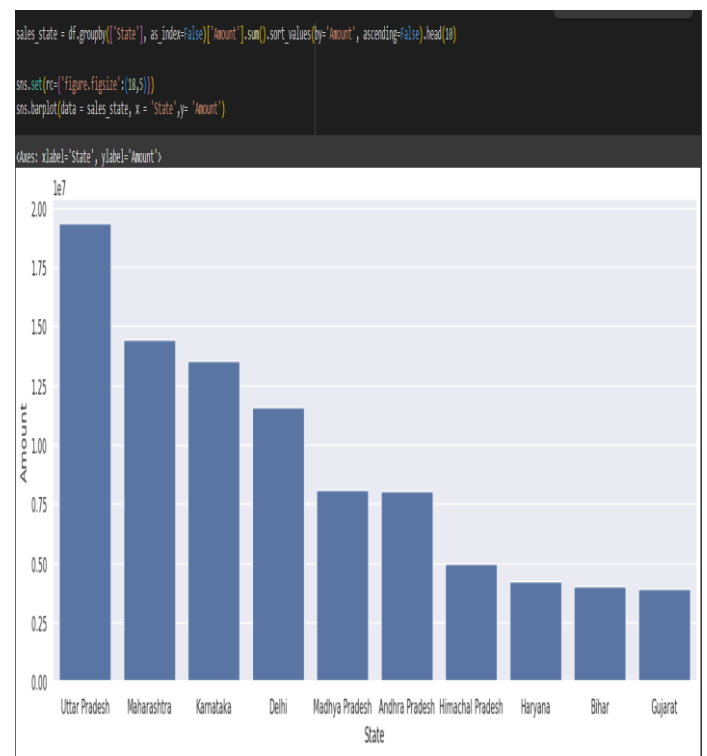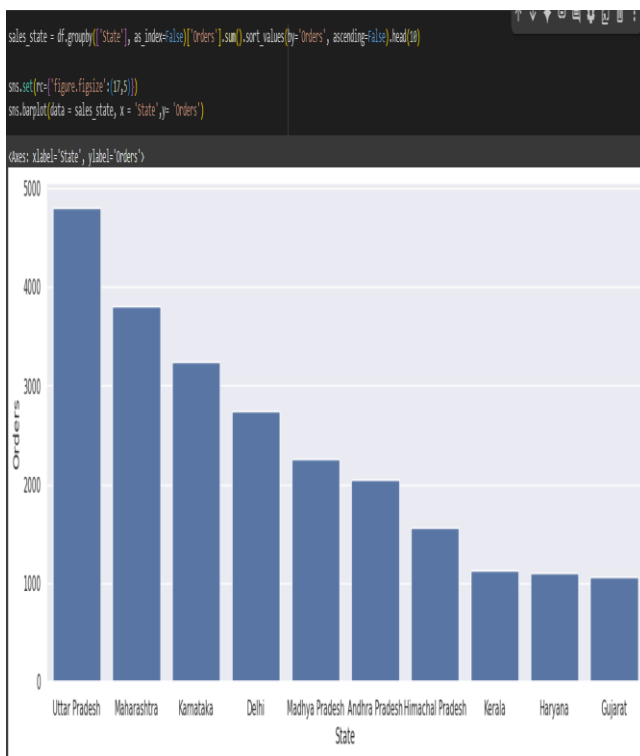**From above graphs we can see that most of the buyers are females and even the purchasing power of females are greater than men**
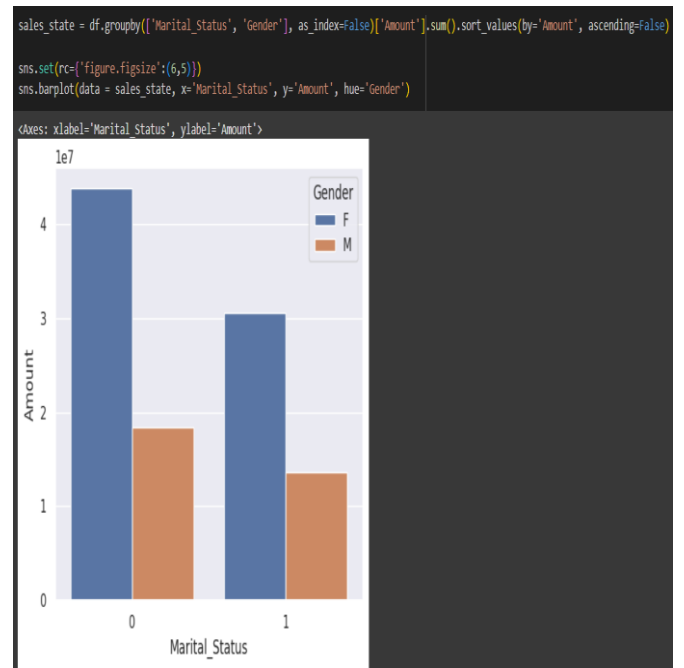
b. Age

```
ax = sns.countplot(data = df, x = 'Age Group', hue = 'Gender')

for bars in ax.containers:
  ax.bar_label(bars)
```



```
# Total amount vs age group
sales_age = df.groupby(['Age Group'], as_index=False)['Amount'].sum().sort_values(by='Amount', ascending=False)

sns.barplot(x = 'Age Group', y = 'Amount', data = sales_age)
```

```
<Axes: xlabel='Age Group', ylabel='Amount'>
```


From above graphs we can see that most of the buyers are age group between 26-35 years female

c. state

```
sales_state = df.groupby(['State'], as_index=False)['Orders'].sum().sort_values(by='Orders', ascending=False).head(10)

sns.set(rc={'figure.figsize':(17,5)})
sns.barplot(data = sales_state, x = 'State', y= 'Orders')
```

```
<Axes: xlabel='State', ylabel='Orders'>
```



```
sales_state = df.groupby(['State'], as_index=False)['Amount'].sum().sort_values(by='Amount', ascending=False).head(10)

sns.set(rc={'figure.figsize':(18,5)})
sns.barplot(data = sales_state, x = 'State', y= 'Amount')
```

```
<Axes: xlabel='State', ylabel='Amount'>
```


From above graphs we can see that most of the orders and total sales/amount are from Utter Pradesh, Maharashtra and Karnataka respectively

d. Marital State

```
ax = sns.countplot(data = df, x = 'Marital_Status')

sns.set(rc={'figure.figsize':(7,5)})
for bars in ax.containers:
  ax.bar_label(bars)
```
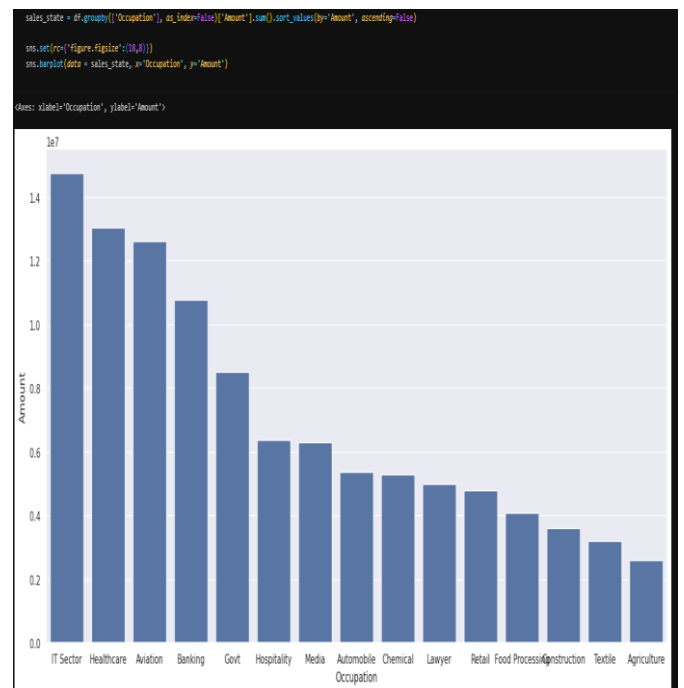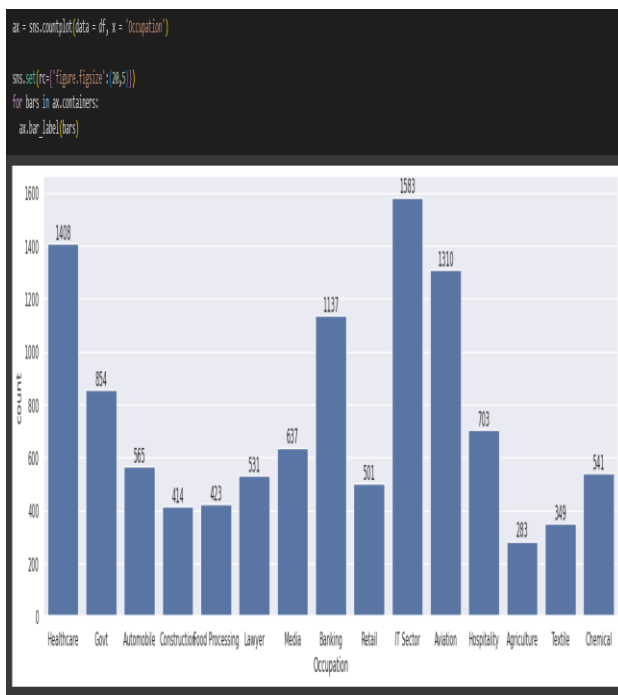
```
sales_state = df.groupby(['Marital_Status', 'Gender'], as_index=False)['Amount'].sum().sort_values(by='Amount', ascending=False)

sns.set(rc={'figure.figsize':(6,5)})
sns.barplot(data = sales_state, x='Marital_Status', y='Amount', hue='Gender')
```

```
<Axes: xlabel='Marital_Status', ylabel='Amount'>
```
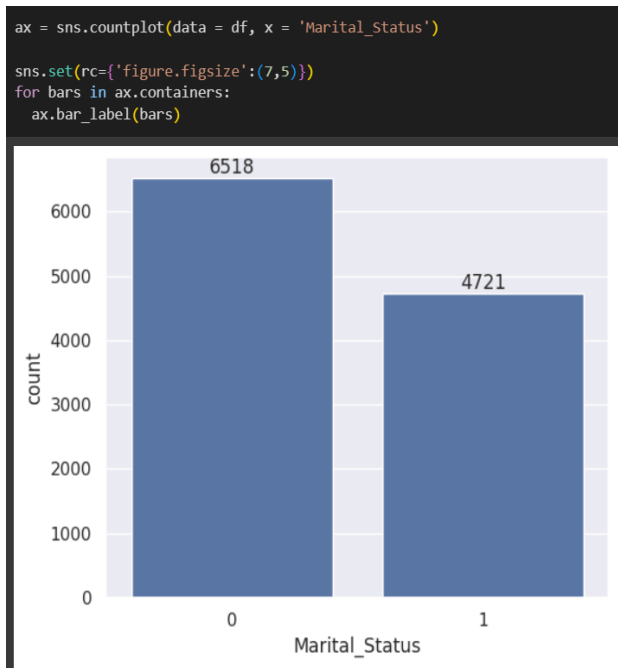

From above graph we can see that most of the buyers are working in IT, Aviation and Healthcare sector

e. Occupation

```
ax = sns.countplot(data = df, x = 'Occupation')

sns.set(rc={'figure.figsize':(20,5)})
for bars in ax.containers:
  ax.bar_label(bars)
```

```
sales_state = df.groupby(['Occupation'], as_index=False)['Amount'].sum().sort_values(by='Amount', ascending=False)

sns.set(rc={'figure.figsize':(18,0)})
sns.barplot(data = sales_state, x='Occupation', y='Amount')
```

```
<Axes: xlabel='Occupation', ylabel='Amount'>
```


From above graph we can see that most of the buyers are working in IT, Healthcare and Aviation sector

7. CONCLUSION

Married women age group 26-35 yrs form UP, Maharashtra and Karnataka working in IT, Healthcare ana Aviation are more likely to buy product from Food, Clothing and Electronics category