DIWALI SALES ANALYSIS PROJECT

OBJECTIVE: To increase sales at different states of INDIA according to different category of people's preferences and other factors...

DATASET TO BE USED: Real life data of people and products purchased by them at or around the time of Diwali season from different states of india in a CSV file...

```
# import python libraries

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt # visualizing data
%matplotlib inline
import seaborn as sns

# import csv file
df =
pd.read_csv('C:/Users/charu/Downloads/Python_Diwali_
Sales_Analysis/Python_Diwali_Sales_Analysis/Diwali
Sales Data.csv', encoding= 'unicode_escape')
df.shape
df.head()
df.info()

#drop unrelated/blank columns
```

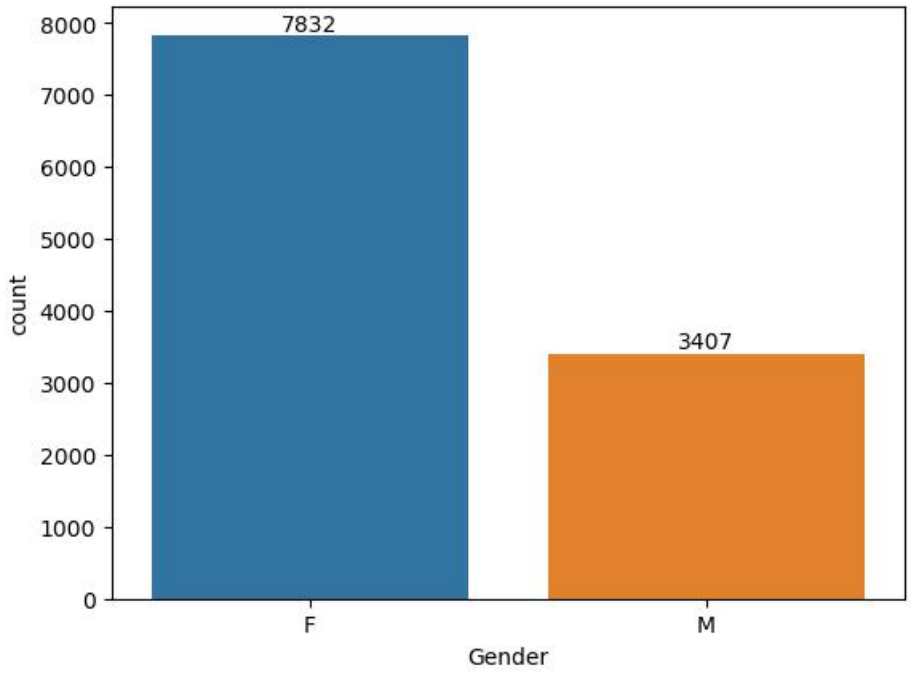
```
df.drop(['Status', 'unnamed1'], axis=1, inplace=True)
#check for null values
pd.isnull(df).sum()
# drop null values
df.dropna(inplace=True)
# change data type
df['Amount'] = df['Amount'].astype('int')
df['Amount'].dtypes
Df.columns

#rename column
df.rename(columns= {'Marital_Status':'Shaadi'})

# describe() method returns description of the data in the DataFrame (i.e. count, mean, std, etc)
df.describe()

# use describe() for specific columns
df[['Age', 'Orders', 'Amount']].describe()
```

#EXPLORATORY DATA ANALYSIS

#GENDER

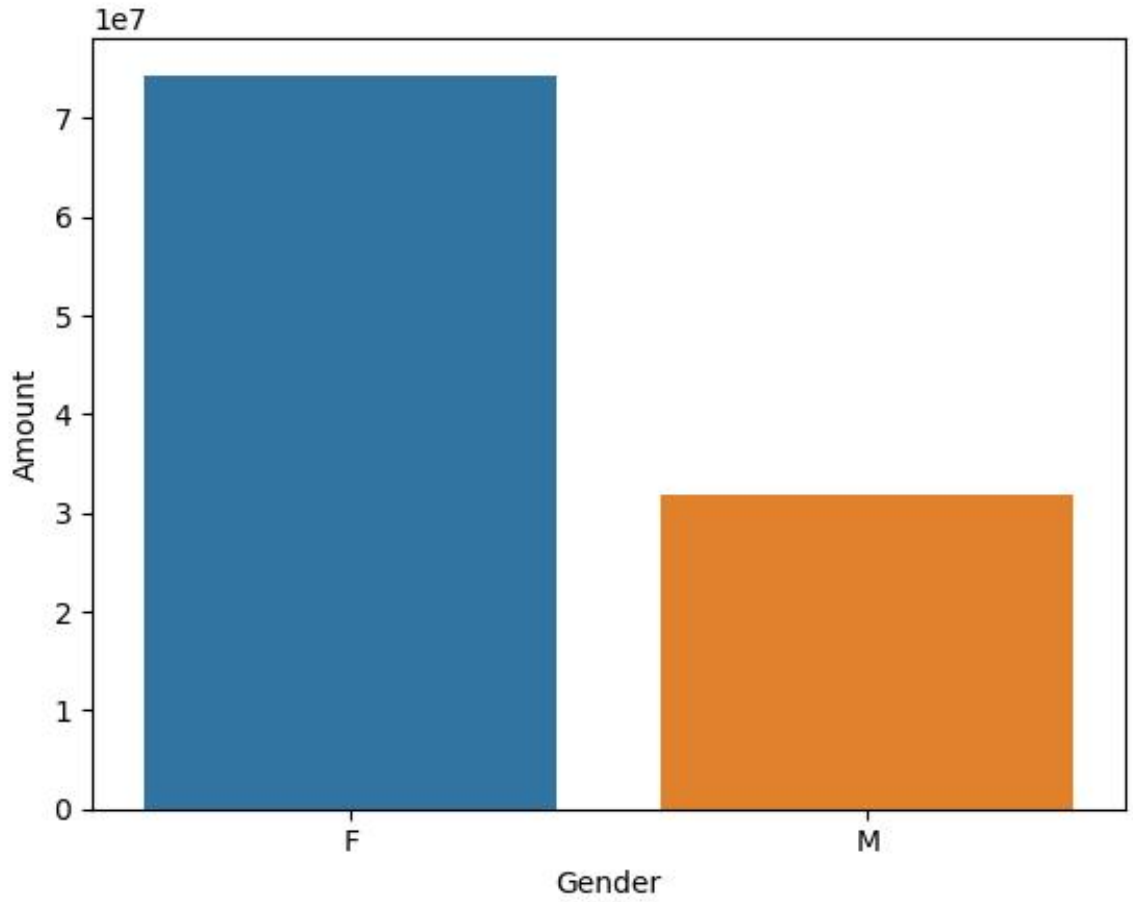
plotting a bar chart for Gender and it's count

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

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



plotting a bar chart for gender vs total amount

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

sns.barplot(x = 'Gender',y= 'Amount' ,data = sales_gen)
```

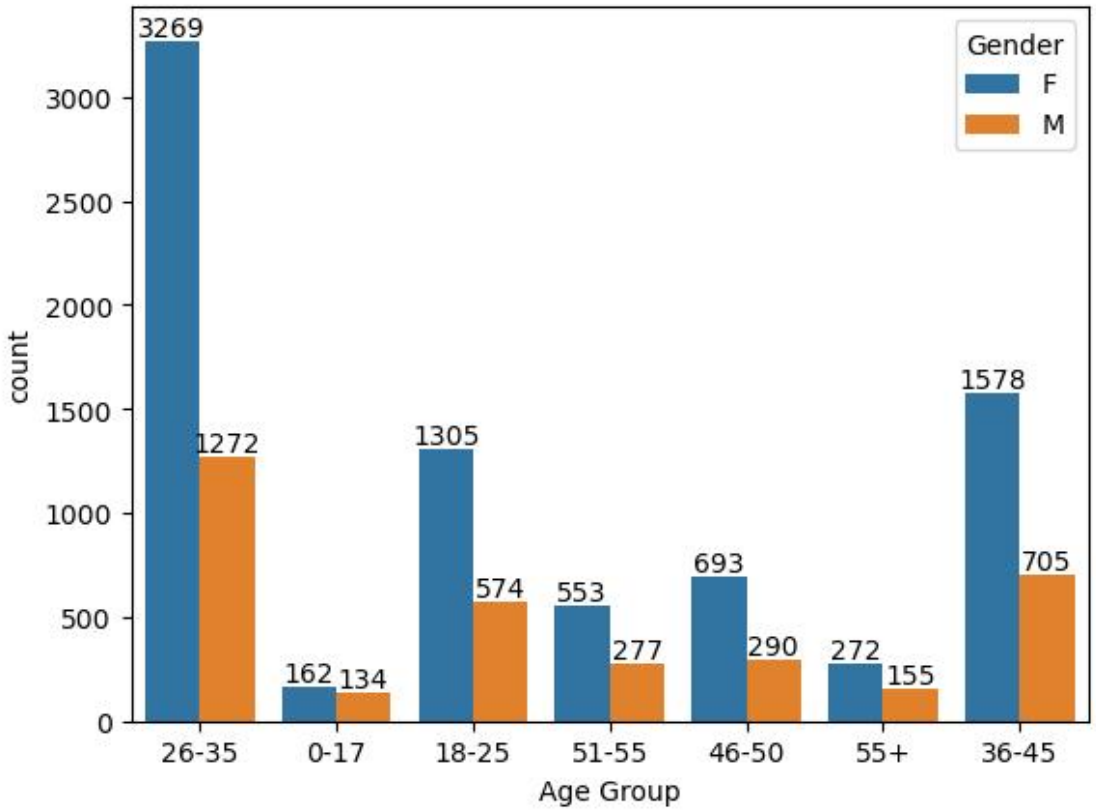


Q1. Which gender is more involved in Diwali Shopping?

Ans: From the above graphs, we can see that most of the buyers are female and even the purchasing power of females is greater than men.

#AGE

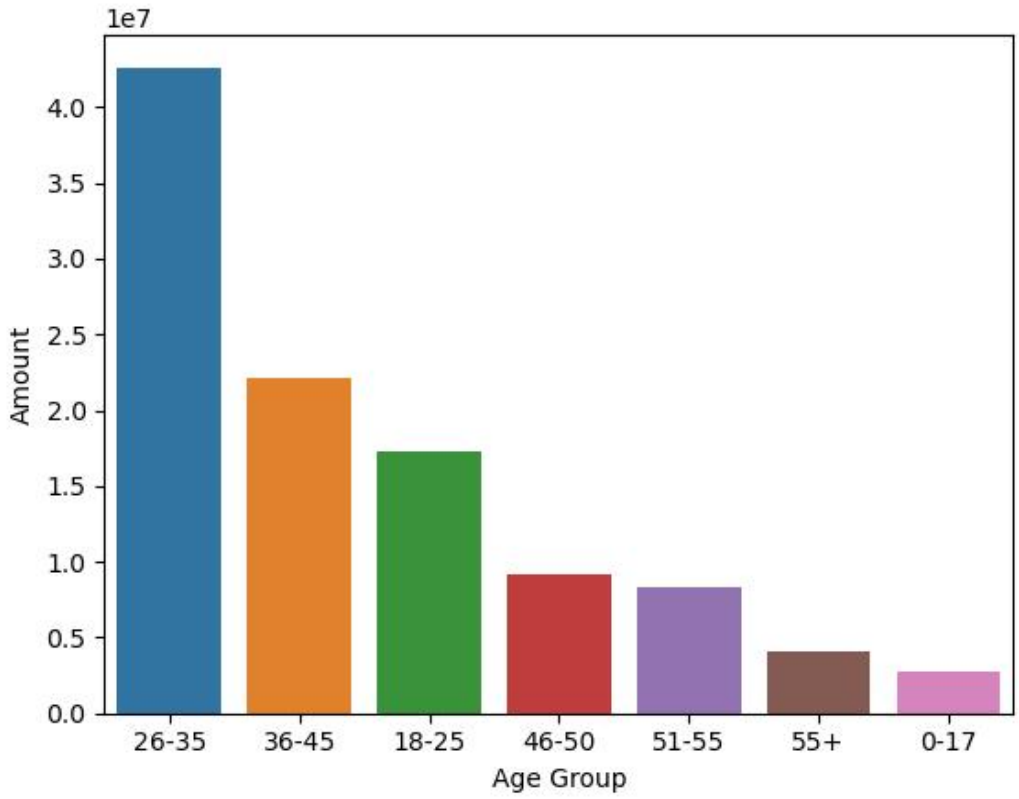
```
ax = sns.countplot(data = df, x = 'Age Group', hue = 'Gender')

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)
```



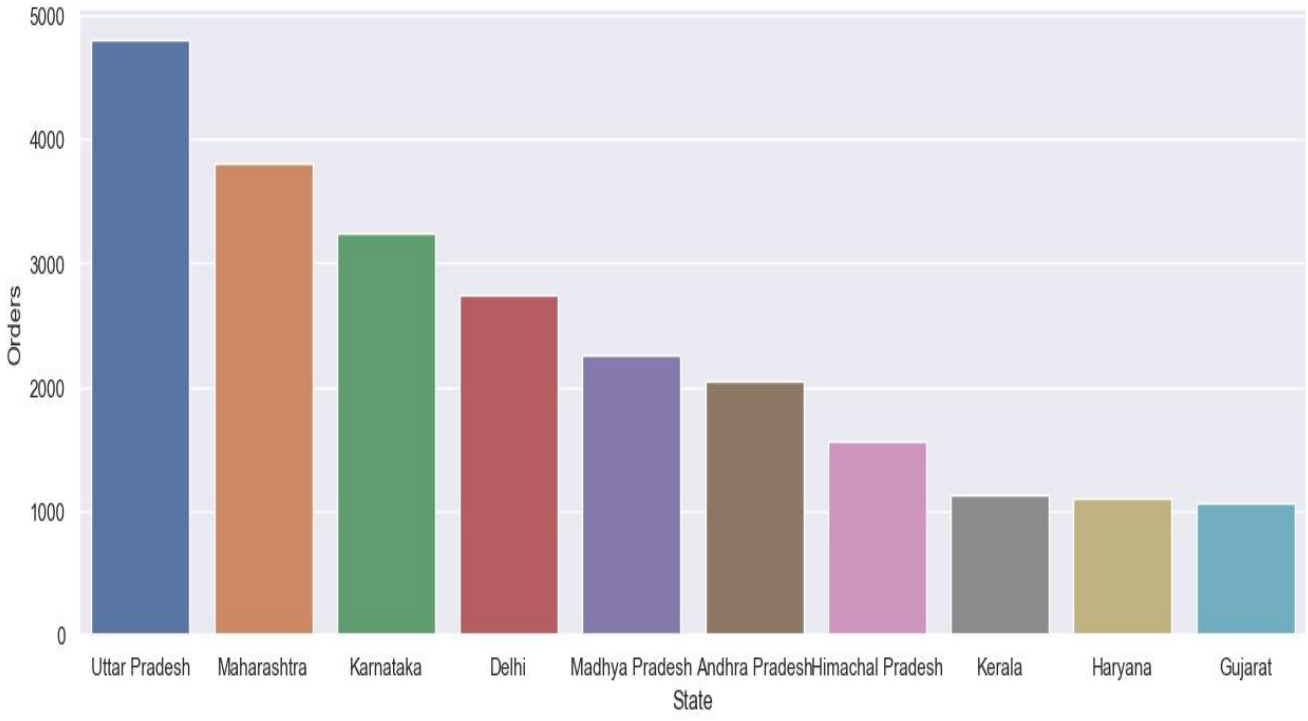
Q2. Which age group holds the most buyers?

Ans: From the above graphs, we can see that most of the buyers are of the age group 26-35 years female.

#STATE

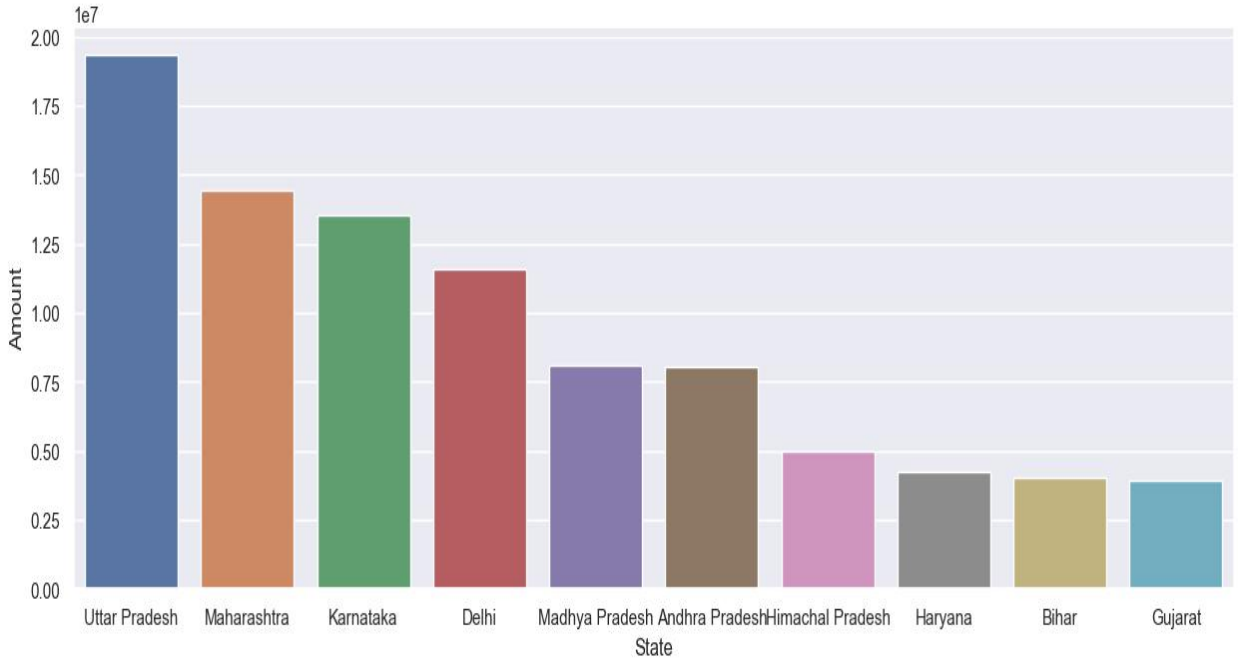
total number of orders from top 10 states

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


sns.set(rc={'figure.figsize':(15,5)})
sns.barplot(data = sales_state, x = 'State',y= 'Orders')
```




```
# total amount/sales from top 10 states
```

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


sns.set(rc={'figure.figsize':(15,5)})

sns.barplot(data = sales_state, x = 'State',y= 'Amount')
```

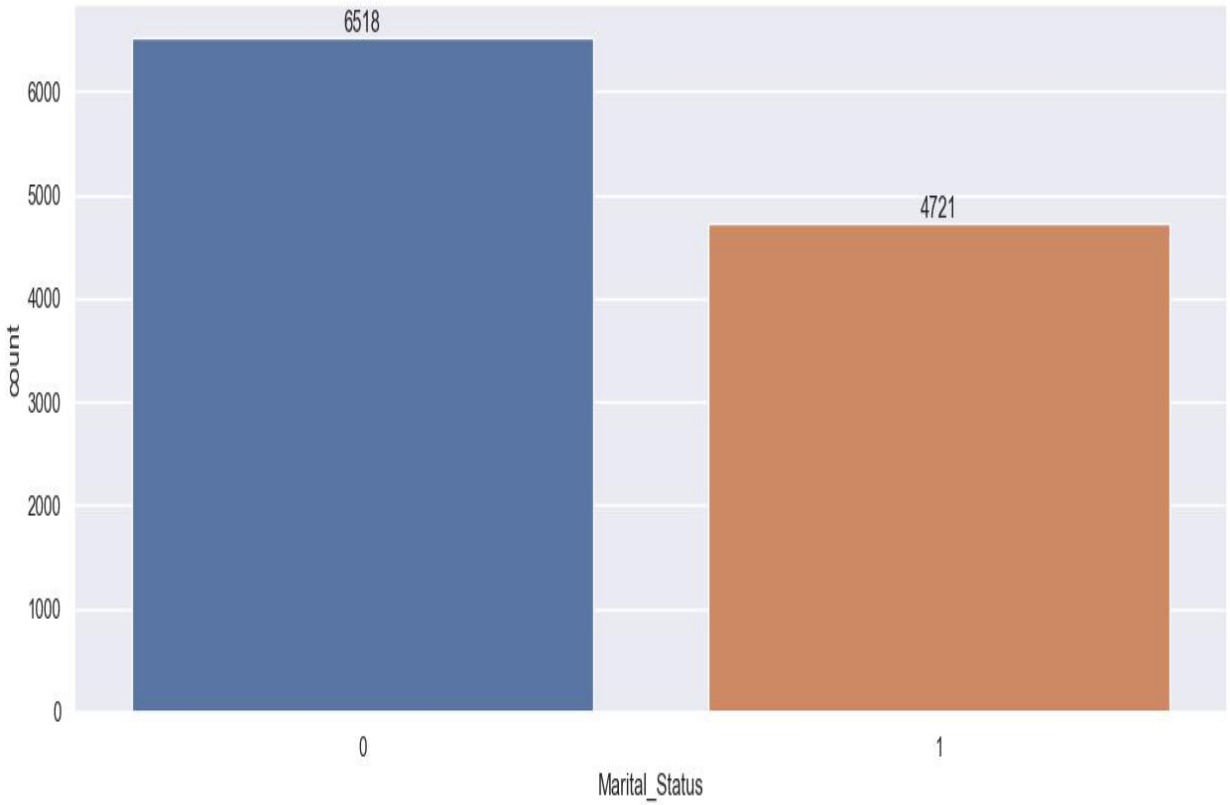


Q3. Majority of the buyers are from which state?

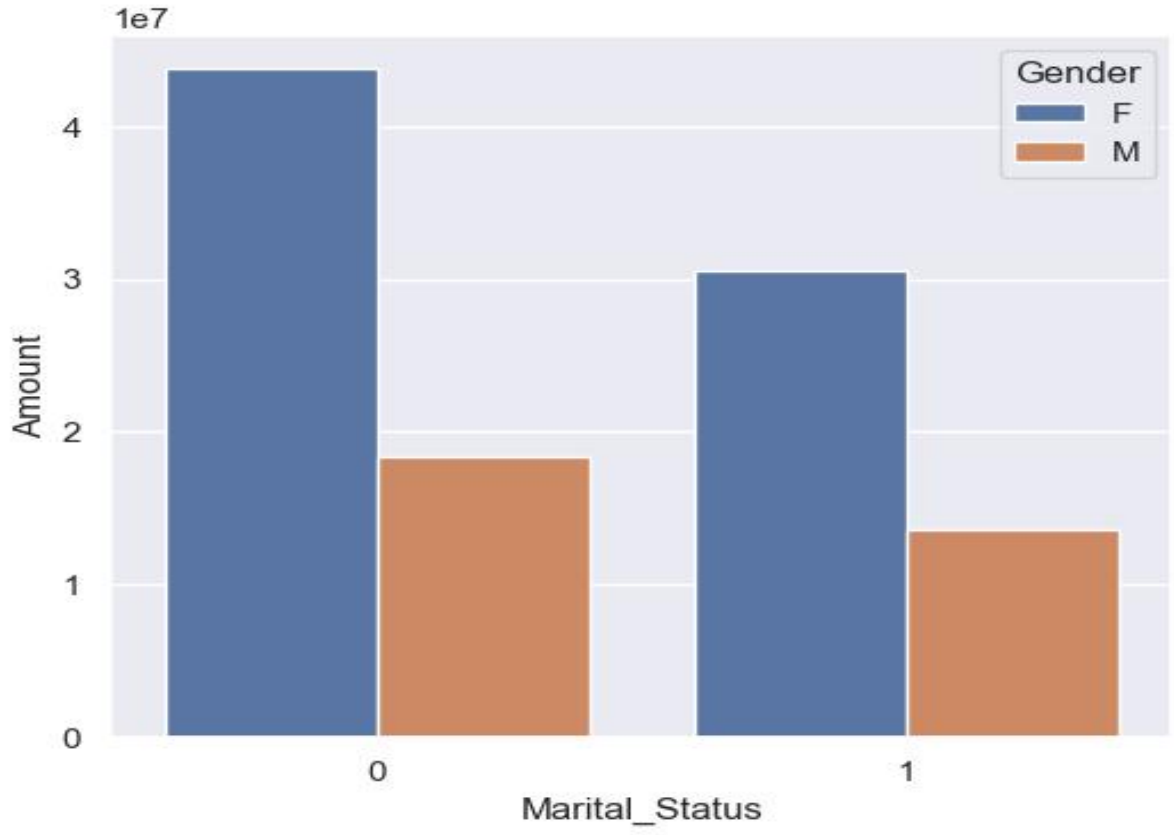
Ans: From the above graphs we can see that most of the orders and total sales/amount are from Uttar Pradesh, Maharashtra and Karnataka respectively.

#MARITAL STATUS

```
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')
```

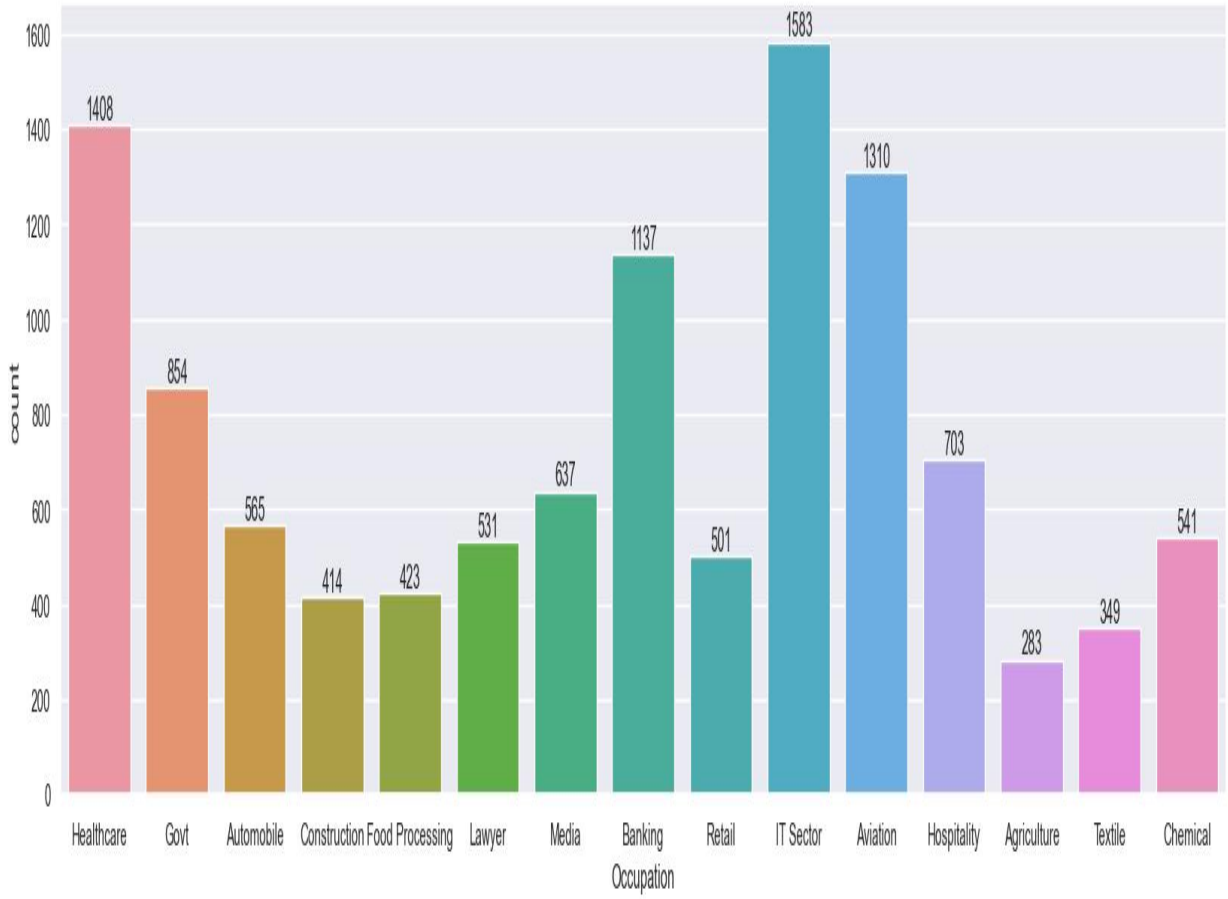


Q4. What is the marital status of most of the buyers?

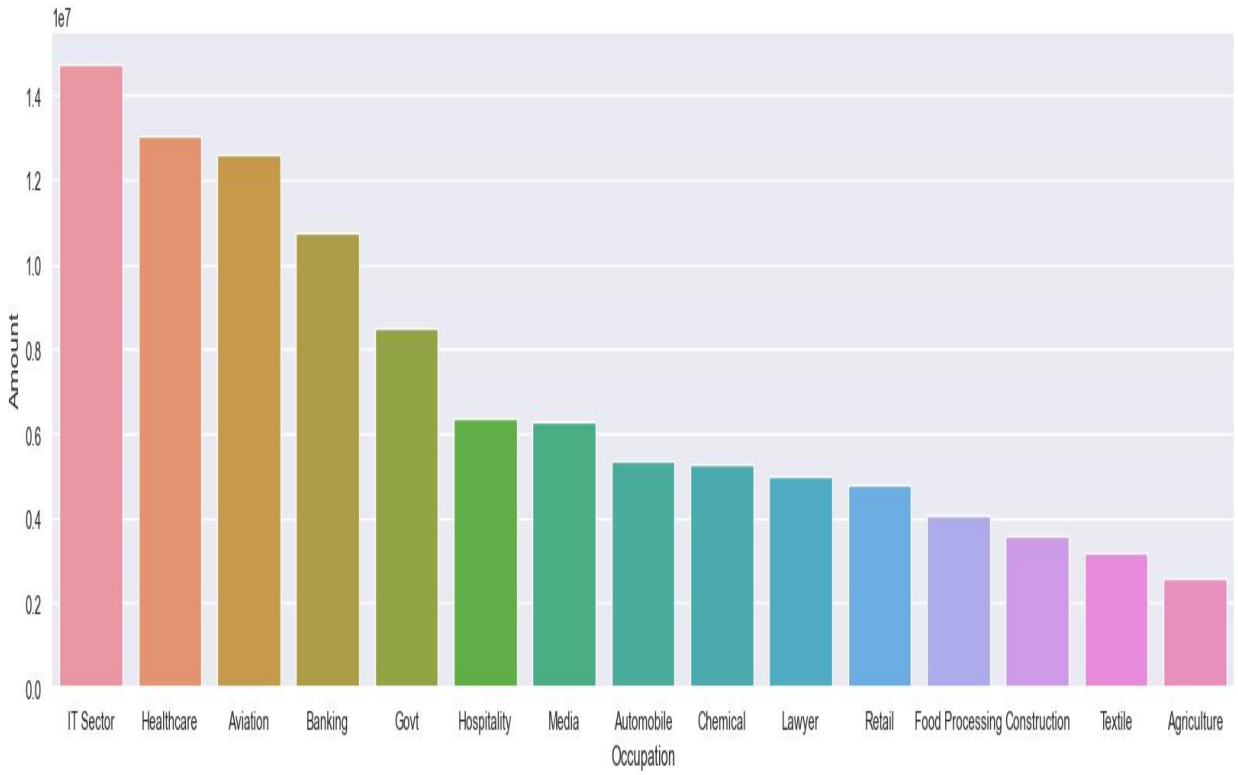
Ans: From the above graphs we can see that most of the buyers are married(women) and they have high purchasing power.

#OCCUPATION

```
sns.set(rc={'figure.figsize':(20,5)})

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


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':(20,5)})

sns.barplot(data = sales_state, x = 'Occupation',y= 'Amount')
```



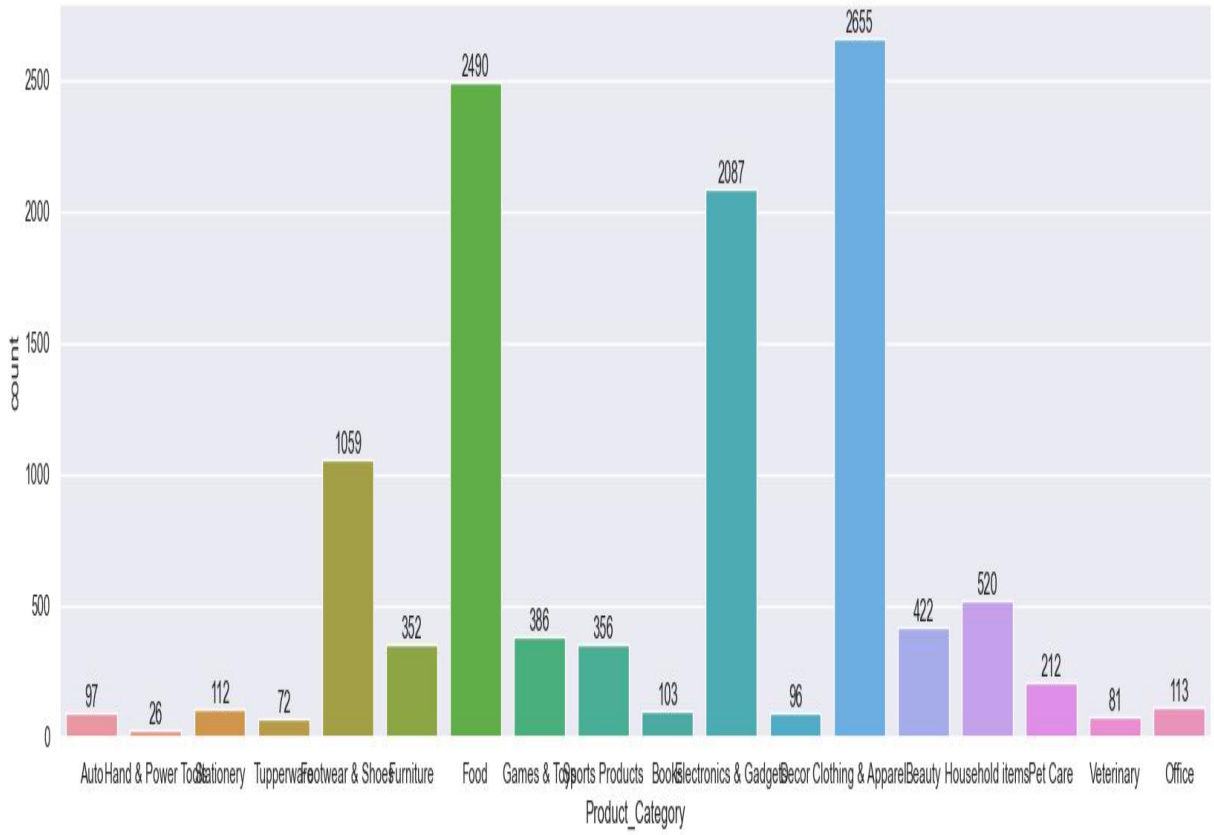
Q5. What is the occupation of majority of the buyers?

Ans: From the above graphs, we can see that most of the buyers are working IT, Healthcare and Aviation sector.

#PRODUCT CATEGORY

```
sns.set(rc={'figure.figsize':(20,5)})

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


for bars in ax.containers:
```

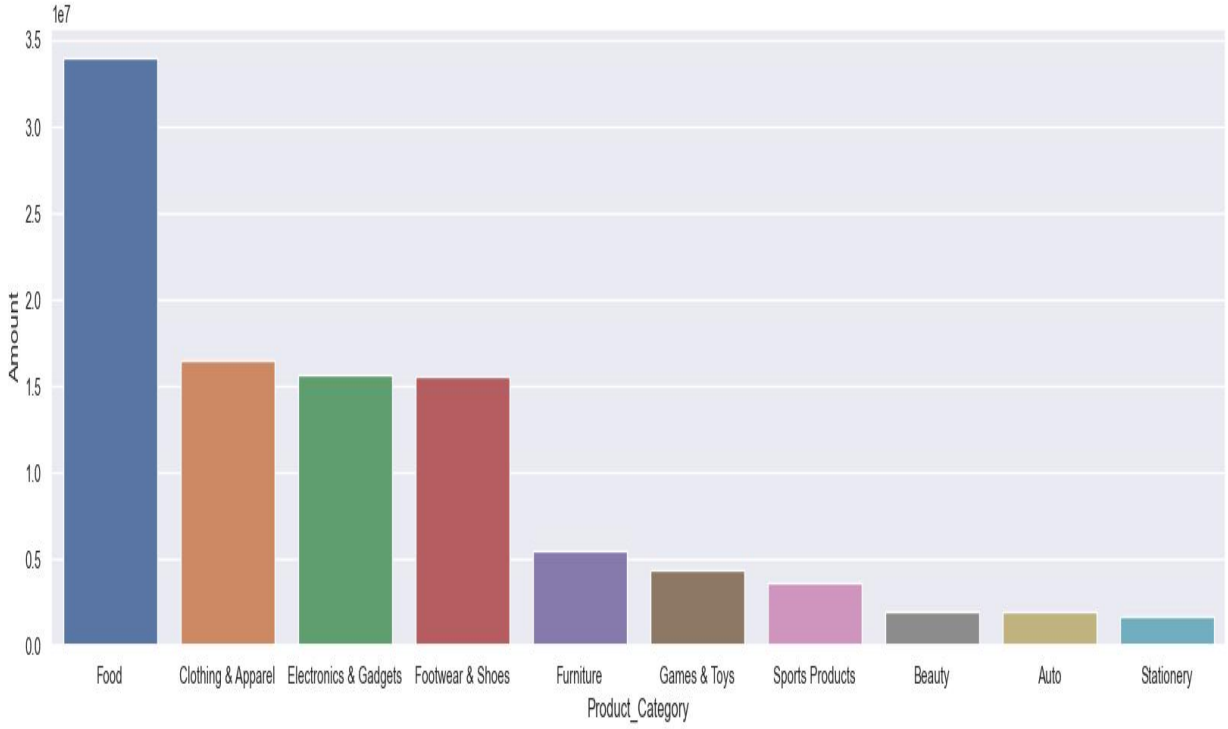
```
ax.bar_label(bars)
```




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

```
sns.set(rc={'figure.figsize':(20,5)})
```

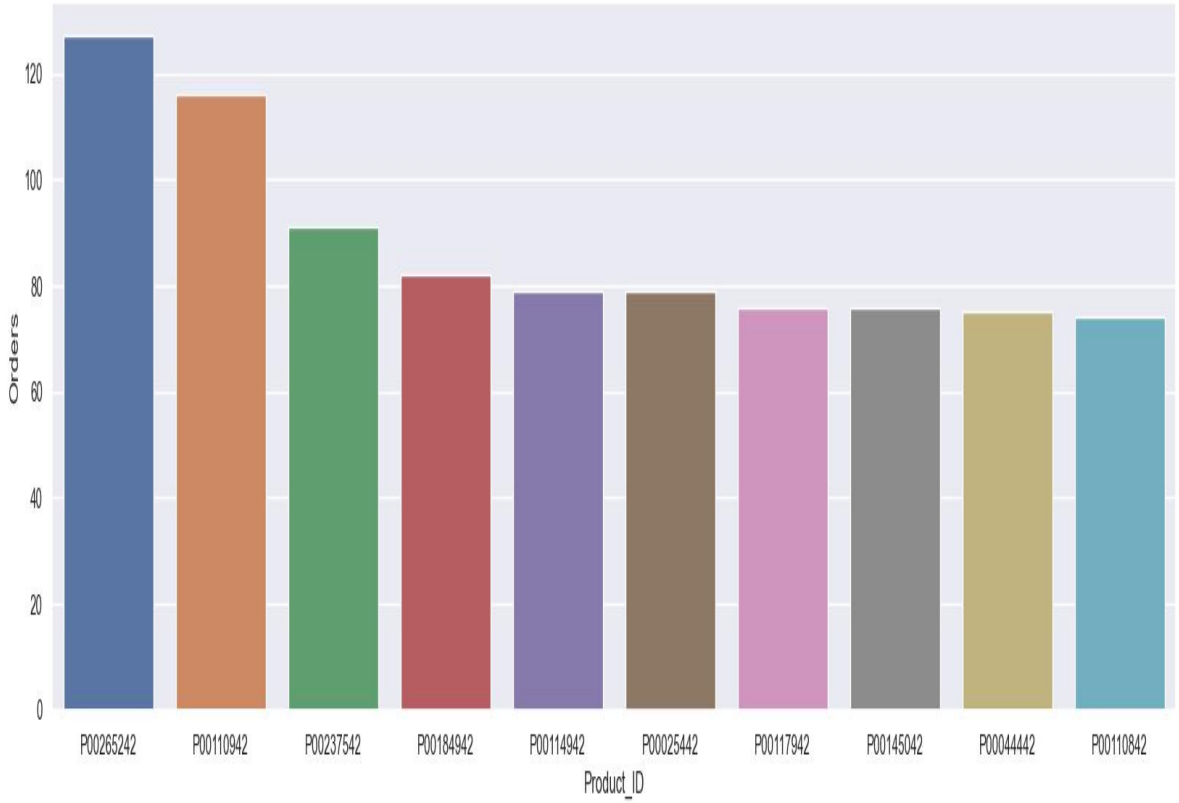
```
sns.barplot(data = sales_state, x = 'Product_Category',y= 'Amount')
```



Q6. Which product category has the most sales?

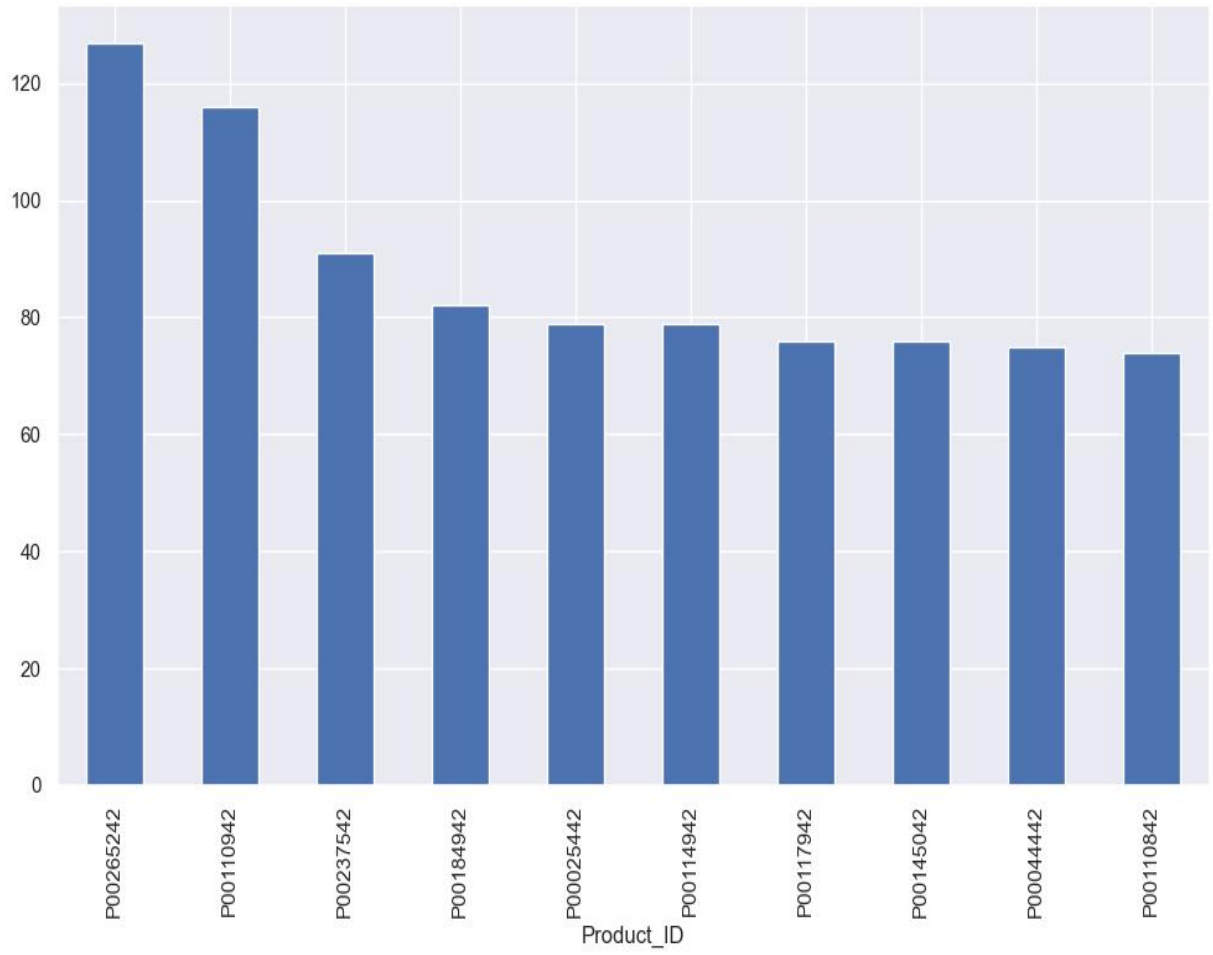
Ans: From the above graphs, we can see that most of the sold products are from Food, Clothings and Electronic category.

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


sns.set(rc={'figure.figsize':(20,5)})

sns.barplot(data = sales_state, x = 'Product_ID',y= 'Orders')
```



top 10 most sold products (same thing as above)

```
fig1, ax1 = plt.subplots(figsize=(12,7))
df.groupby('Product_ID')['Orders'].sum().nlargest(10).sort_values(ascending=False).plot(kind='bar')
```



Q7. How can we figure out which product to buy in large amount so that its's sell can bring us profit?

Ans: From the above graphs, we can get the top 10 most sold products to identify it.

#CONCLUSION

Married women of age group 26-35 yrs from UP, Maharastra and Karnataka working in IT, Healthcare and Aviation are more likely to buy products from Food, Clothing and Electronics category...

Our Diwali Sales Analysis project aimed to uncover insights into the sales performance during the festive season, specifically focusing on Diwali. Through a combination of data analysis and visualization techniques, we sought to understand customer behavior, identify popular products, and provide actionable recommendations for future marketing strategies.

By leveraging the insights gained from this analysis, businesses can position themselves strategically to

maximize sales and customer satisfaction during future Diwali seasons.