Diwali Sales Analysis

March 14, 2025

[4]:
```
import pandas as pd
import numpy as np
from sqlalchemy import create_engine
from urllib.parse import quote_plus
import matplotlib.pyplot as plt
import seaborn as sns
from matplotlib import gridspec

import warnings
warnings.filterwarnings("ignore")
```

[5]:
```
df = pd.read_csv('diwali_sales_data.csv',encoding='latin1')
```

[6]:
```
df.shape
```

[6]: (11251, 13)

[7]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 11251 entries, 0 to 11250
Data columns (total 13 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   User_ID           11251 non-null  int64
 1   Cust_name         11251 non-null  object
 2   Product_ID        11251 non-null  object
 3   Gender            11251 non-null  object
 4   Age Group         11251 non-null  object
 5   Age               11251 non-null  int64
 6   Marital_Status    11251 non-null  int64
 7   State             11251 non-null  object
 8   Zone              11251 non-null  object
 9   Occupation        11251 non-null  object
 10  Product_Category  11251 non-null  object
 11  Orders            11251 non-null  int64
 12  Amount            11239 non-null  float64
dtypes: float64(1), int64(4), object(8)
```

```
memory usage: 1.1+ MB
```

[8]:
```
df.describe()
```

[8]:

	User_ID	Age	Marital_Status	Orders	Amount
count	1.125100e+04	11251.000000	11251.000000	11251.000000	11239.000000
mean	1.003004e+06	35.421207	0.420318	2.489290	9453.610858
std	1.716125e+03	12.754122	0.493632	1.115047	5222.355869
min	1.000001e+06	12.000000	0.000000	1.000000	188.000000
25%	1.001492e+06	27.000000	0.000000	1.500000	5443.000000
50%	1.003065e+06	33.000000	0.000000	2.000000	8109.000000
75%	1.004430e+06	43.000000	1.000000	3.000000	12675.000000
max	1.006040e+06	92.000000	1.000000	4.000000	23952.000000

[9]:
```
df.columns
```

[9]:
```
Index(['User_ID', 'Cust_name', 'Product_ID', 'Gender', 'Age Group', 'Age',
       'Marital_Status', 'State', 'Zone', 'Occupation', 'Product_Category',
       'Orders', 'Amount'],
      dtype='object')
```

[10]:
```
df.head()
```

[10]:

	User_ID	Cust_name	Product_ID	Gender	Age Group	Age	Marital_Status	State	Zone	Occupation	Product_Category	Orders	Amount
0	1002903	Sanskriti	P00125942	F	26-35	28	0	Maharashtra	Western	Healthcare	Auto	1	23952.0
1	1000732	Kartik	P00110942	F	26-35	35	1	Andhra Pradesh	Southern	Govt	Auto	3	23934.0
2	1001990	Bindu	P00118542	F	26-35	35	1	Uttar Pradesh	Central	Automobile	Auto	3	23924.0
3	1001425	Sudevi	P00237842	M	0-17	16	0	Karnataka	Southern	Construction	Auto	2	23912.0
4	1000588	Joni	P00057942	M	26-35	28	1	Gujarat	Western	Food Processing	Auto	2	23877.0

[11]:
```
def plotTwoCharts(df, chartParams):
    # print(df.columns)
    """
    Function to plot two charts side by side with different chart types (line, scatter, bar, pie, histogram).
    Parameters:
    df (DataFrame): The dataframe containing the data
    chartParams (dict): Dictionary containing chart details
    """
    totalCharts = len(chartParams['chartData'])
```

```
    rows = (totalCharts + 1) // 2  # Calculate rows for the fixed 2-column layout

    # Create subplots
    fig, axes = plt.subplots(rows, 2, figsize=(19, 5 * rows))
    axes = axes.flatten()  # Flatten to simplify indexing

    for chart in range(totalCharts):
        chartDetails = chartParams['chartData'][chart]
        chartType = chartDetails['type']
        xvalue = chartDetails['xCol']
        yvalues = chartDetails.get('yCol', [])
        lvalue = chartDetails.get('legend', None)  # Use .get to handle optional keys

        sns.set_style("darkgrid")
        if chartType == 'line':
            if lvalue:  # If 'legend' is specified, restructure the data for grouped plotting
                plot_df = pd.melt(
                    df,
                    id_vars=[xvalue],
                    value_vars=yvalues,
                    var_name='Group',
                    value_name='Value'
                )
                plot_df['Group'] = plot_df['Group'].replace(dict(zip(yvalues, lvalue)))
                sns.lineplot(
                    data=plot_df,
                    x=xvalue,
                    y='Value',
                    hue='Group',
                    marker='o',
                    ax=axes[chart]
                )
            else:  # Simple line plot
                for col in yvalues:
                    sns.lineplot(
                        data=df,
                        x=xvalue,
                        y=col,
                        marker='o',
                        ax=axes[chart]
                    )
        elif chartType == 'scatter':
```

```
            sns.scatterplot(data=df, x=xvalue, y=yvalues[0], hue=lvalue,␣
↪ax=axes[chart])
        elif chartType == 'bar':
            if len(yvalues) > 1 and lvalue:
                    melted_df = pd.melt(
                        df,
                        id_vars=[xvalue],
                        value_vars=yvalues,
                        var_name='Group',
                        value_name='Value'
                    )
                    melted_df['Group'] = melted_df['Group'].replace(
                        dict(zip(yvalues, lvalue))
                    )
                    sns.barplot(
                        data=melted_df,
                        x=xvalue,
                        y='Value',
                        hue='Group',
                        palette="Set2",
                        ax=axes[chart]
                    )
            else:
                bars = sns.barplot(
                    data=df,
                    x=xvalue,
                    y=yvalues[0],
                    color="#66b3ff",
                    ax=axes[chart]
                )
                for bar in bars.patches:
                    height = bar.get_height()
                    bars.annotate(
                        f'{height:.1f}',
                        (bar.get_x() + bar.get_width() / 2, height),
                        ha='center',
                        va='bottom',
                        fontsize=9,
                        color='black'
                    )
        elif chartType == 'pie':
            # For pie chart: Use the first column in xCol as categories
            # Enhanced Pie Chart Code
            pie_data = df[xvalue].value_counts()

            # Create the pie chart
            wedges, texts, autotexts = axes[chart].pie(
```

```
                pie_data,
                autopct='%1.1f%%',
                startangle=90,
                labels=pie_data.index,
                colors=['#66b3ff', '#ff9999', '#99ff99', '#ffcc99'],  # Custom␣
↪color palette
                textprops={'fontsize': 10, 'color': 'black'}  # Text properties␣
↪for better readability
            )

            # Style the percentage labels
            for autotext in autotexts:
                autotext.set_fontsize(12)
                autotext.set_fontweight('bold')

            # Set title with better styling
            axes[chart].set_title(
                "Loan Default Distribution".upper(),
                fontsize=14,
                fontweight='bold',
                pad=20
            )

            # Remove y-axis label
            axes[chart].set_ylabel("")

            # Add a legend outside the chart
            axes[chart].legend(
                pie_data.index,
                title="Categories",
                loc="upper right",
                bbox_to_anchor=(1.2, 0.9),  # Position outside the chart
                fontsize=10
            )

        elif chartType == 'histogram':
            sns.histplot(data=df, x=xvalue, bins=20, kde=True, ax=axes[chart])

        axes[chart].set_title(chartDetails['chartTitle'].upper(), fontsize=12)
        axes[chart].set_xlabel(chartDetails.get('xlabel', xvalue.upper()),␣
↪fontsize=10)
        axes[chart].set_ylabel(chartDetails.get('ylabel', ', '.join(yvalues)),␣
↪fontsize=10)
        axes[chart].tick_params(axis='both', which='major', labelsize=10)

    # Hide any unused axes
    for ax in axes[totalCharts:]:
```

```
        ax.set_visible(False)

    # Adjust layout for better spacing
    plt.tight_layout()
    plt.show()
```

[12]:
```
df['User_ID'].nunique()
```

[12]:
```
3755
```

[13]:
```
df.shape[0] / df['User_ID'].nunique()
```

[13]:
```
2.996271637816245
```

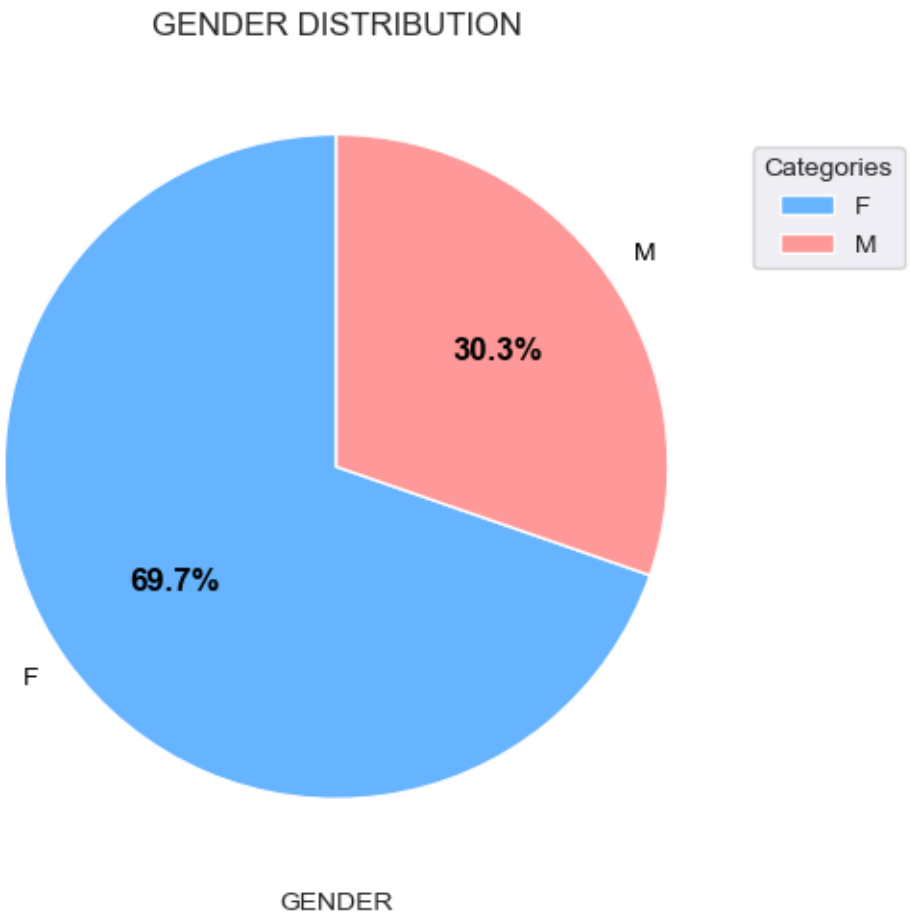
[14]:
```
df['Product_ID'].nunique()
```

[14]:
```
2351
```

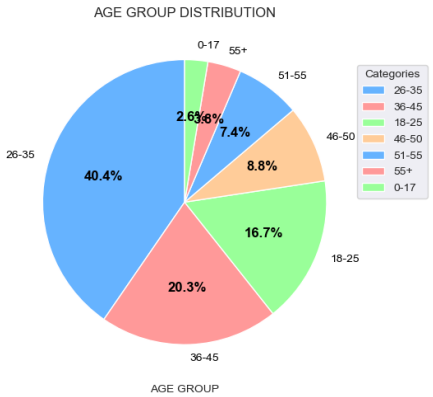
[15]:
```
df.shape[0] / df['Product_ID'].nunique()
```

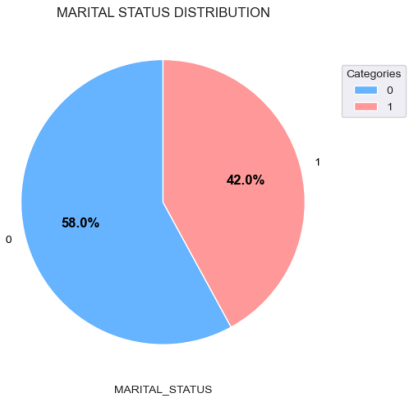
[15]:
```
4.785623139089749
```

[18]:
```
chartParams = {
    "chartData": [
       {
            "type": "pie",  # Simple bar chart
            "xCol": "Gender",  # States as the x-axis
            "chartTitle": "Gender DIstribution",
            "legend": None  # Simple bar chart without grouping
        },
    ]
}

plotTwoCharts(df, chartParams)
```



[24]:
```
chartParams = {
    "chartData": [
       {
            "type": "pie",  # Simple bar chart
            "xCol": "Age Group",  # States as the x-axis
            "chartTitle": "Age Group Distribution",
            "legend": None  # Simple bar chart without grouping
        },
        {
            "type": "pie",  # Simple bar chart
            "xCol": "Marital_Status",  # States as the x-axis
            "chartTitle": "Marital Status Distribution",
            "legend": None  # Simple bar chart without grouping
        },
    ]
}

plotTwoCharts(df, chartParams)
```





[24]:
```
print(df['Age'].mean())
print(df['Age'].max())
print(df['Age'].min())
```

```
35.421207003821884
92
12
```

[25]:
```
df['Marital_Status'].value_counts()
```

[25]:
```
Marital_Status
0    6522
1    4729
Name: count, dtype: int64
```

[26]:
```
df['State'].value_counts()
```

[26]:
```
State
Uttar Pradesh       1946
Maharashtra         1526
Karnataka           1305
Delhi               1107
Madhya Pradesh       923
Andhra Pradesh       812
Himachal Pradesh     608
Kerala               453
Haryana              452
Bihar                434
Gujarat              429
Jharkhand            380
Uttarakhand          320
Rajasthan            231
Punjab               200
```

```
Telangana            125
Name: count, dtype: int64
```

[27]:
```
df['Zone'].value_counts()
```

[27]:
```
Zone
Central     4296
Southern    2695
Western     1955
Northern    1491
Eastern      814
Name: count, dtype: int64
```

[28]:
```
df['Occupation'].value_counts()
```

[28]:
```
Occupation
IT Sector          1588
Healthcare         1408
Aviation           1310
Banking            1139
Govt                854
Hospitality         705
Media               637
Automobile          566
Chemical            542
Lawyer              531
Retail              501
Food Processing     423
Construction        414
Textile             350
Agriculture         283
Name: count, dtype: int64
```

[29]:
```
df['Product_Category'].value_counts()
```

[29]:
```
Product_Category
Clothing & Apparel       2655
Food                     2493
Electronics & Gadgets    2087
Footwear & Shoes         1064
Household items           520
Beauty                    422
Games & Toys              386
Sports Products           356
Furniture                 353
Pet Care                  212
Office                    113
```

```
Stationery                112
Books                     103
Auto                      100
Decor                      96
Veterinary                 81
Tupperware                 72
Hand & Power Tools         26
Name: count, dtype: int64
```

[31]:
```
print(df['Orders'].mean())
print(df['Orders'].max())
print(df['Orders'].min())
```

```
2.4892898409030306
4
1
```

[32]:
```
print(df['Amount'].mean())
print(df['Amount'].max())
print(df['Amount'].min())
```

```
9453.610857727557
23952.0
188.0
```

[33]:
```
df.columns
```

[33]:
```
Index(['User_ID', 'Cust_name', 'Product_ID', 'Gender', 'Age Group', 'Age',
       'Marital_Status', 'State', 'Zone', 'Occupation', 'Product_Category',
       'Orders', 'Amount'],
      dtype='object')
```

dividing the dataset into two groups

The customer features - 1. User_ID 2. Gender 3. Age Group 4. Age 5. Marital Status 6. State 7. Zone 8. Occupation

The product features 1. Product_ID 2. Product_Category 3. Orders 4. Amount

[39]:
```
df.groupby('Age Group').agg(
    TotalUsers=('User_ID', 'count'),
    Females=('Gender', lambda x: (x == 'F').sum()),
    Males=('Gender', lambda x: (x == 'M').sum()),
    Married=('Marital_Status', lambda x: (x == 1).sum()),
    UnMarried=('Marital_Status', lambda x: (x == 0).sum()),
)
```

[39]:

Age Group	TotalUsers	Females	Males	Married	UnMarried
0-17	296	162	134	112	184

```
18-25            1879     1305    574      801       1078
26-35            4543     3271   1272     1932       2611
36-45            2286     1581    705      968       1318
46-50             987      696    291      410        577
51-55             832      554    278      337        495
55+               428      273    155      169        259
```

[40]:
```
df.groupby('Occupation').agg(
    TotalUsers=('User_ID', 'count'),
    Females=('Gender', lambda x: (x == 'F').sum()),
    Males=('Gender', lambda x: (x == 'M').sum()),
    Married=('Marital_Status', lambda x: (x == 1).sum()),
    UnMarried=('Marital_Status', lambda x: (x == 0).sum()),
)
```

[40]:

Occupation	TotalUsers	Females	Males	Married	UnMarried
Agriculture	283	203	80	131	152
Automobile	566	392	174	257	309
Aviation	1310	932	378	583	727
Banking	1139	825	314	569	570
Chemical	542	379	163	241	301
Construction	414	300	114	124	290
Food Processing	423	295	128	194	229
Govt	854	605	249	366	488
Healthcare	1408	968	440	394	1014
Hospitality	705	469	236	346	359
IT Sector	1588	1075	513	719	869
Lawyer	531	364	167	259	272
Media	637	428	209	292	345
Retail	501	370	131	120	381
Textile	350	237	113	134	216

[ ]:
```
```

[42]:
```
df.columns
```

[42]:
```
Index(['User_ID', 'Cust_name', 'Product_ID', 'Gender', 'Age Group', 'Age',
       'Marital_Status', 'State', 'Zone', 'Occupation', 'Product_Category',
       'Orders', 'Amount'],
      dtype='object')
```

[62]:
```
df.groupby('Product_Category').agg(
    TotalProducts=('Product_ID', 'nunique'),
    TotalOrders=('Orders', 'count'),
    MinOrders=('Orders', 'min'),
    MaxOrders=('Orders', 'max'),
```

```
    MeanOrders=('Orders', 'mean')
).sort_values(by='TotalOrders',ascending=False)
```

[62]:

Product_Category	TotalProducts	TotalOrders	MinOrders	MaxOrders	MeanOrders
Clothing & Apparel	1344	2655	1	4	2.498682
Food	1298	2493	1	4	2.454874
Electronics & Gadgets	1120	2087	1	4	2.504073
Footwear & Shoes	730	1064	1	4	2.494361
Household items	434	520	1	4	2.559615
Beauty	370	422	1	4	2.573460
Games & Toys	327	386	1	4	2.435233
Sports Products	315	356	1	4	2.443820
Furniture	313	353	1	4	2.521246
Pet Care	195	212	1	4	2.528302
Office	107	113	1	4	2.309735
Stationery	107	112	1	4	2.508929
Books	99	103	1	4	2.378641
Auto	93	100	1	4	2.460000
Decor	94	96	1	4	2.447917
Veterinary	75	81	1	4	2.543210
Tupperware	72	72	1	4	2.305556
Hand & Power Tools	25	26	1	4	3.076923

[61]:
```
df.groupby('Product_Category').agg(
    TotalOrders=('Amount', 'sum'),
```

```
    MinOrders=('Amount', 'min'),
    MaxOrders=('Amount', 'max'),
    MeanOrders=('Amount', 'mean')
).sort_values(by='TotalOrders',ascending=False)
```

[61]:

Product_Category	TotalOrders	MinOrders	MaxOrders	MeanOrders
Food	33933883.50	3791.0	19708.0	13628.065663
Clothing & Apparel	16495019.00	1715.0	8907.0	6212.813183
Electronics & Gadgets	15643846.00	1939.0	10082.0	7495.853378
Footwear & Shoes	15575209.45	4054.0	20965.0	14707.468791
Furniture	5440051.99	3992.0	20689.0	15454.693153
Games & Toys	4331694.00	3199.0	16499.0	11222.005181
Sports Products	3635933.00	2700.0	13702.0	10213.294944
Beauty	1959484.00	1474.0	7637.0	4643.327014
Auto	1958609.99	9386.0	23952.0	20191.855567
Stationery	1676051.50	4192.0	21563.0	14964.745536
Household items	1569337.00	774.0	3900.0	3017.955769
Tupperware	1155642.00	4642.0	21079.0	16050.583333
Books	1061478.00	5167.0	13264.0	10305.611650
Decor	730360.00	2088.0	10072.0	7607.916667
Pet Care	482277.00	686.0	3555.0	2274.891509
Hand & Power Tools	405618.00	4250.0	23434.0	15600.692308
Veterinary	112702.00	367.0	1777.0	1391.382716
Office	81936.00	188.0	960.0	725.097345

[49]:
```
df.groupby('Product_ID').agg(
    TotalOrders=('Orders', 'count'),
    TotalOrderAmount=('Amount', 'sum'),
    MinOrdersAmount=('Amount', 'min'),
    MaxOrdersAmount=('Amount', 'max'),
    MeanOrdersAmount=('Amount', 'mean'),
    MinOrders=('Orders', 'min'),
    MaxOrders=('Orders', 'max'),
    MeanOrders=('Orders', 'mean')
).sort_values(by='TotalOrders',ascending=False).head()
```

[49]:

Product_ID	TotalOrders	TotalOrderAmount	MinOrdersAmount	MaxOrdersAmount	MeanOrdersAmount	MinOrders	MaxOrders	MeanOrders
P00265242	53	540136.0	1846.0	21325.0				
P00110942	44	424833.0	720.0	23934.0				
P00184942	37	401816.0	744.0	20883.0				
P00237542	35	322363.0	1073.0	19575.0				
P00112142	34	341020.0	3738.0	19280.0				

```
P00265242       10191.245283          1          4    2.396226
P00110942        9655.295455          1          4    2.636364
P00184942       10859.891892          1          4    2.216216
P00237542        9210.371429          1          4    2.600000
P00112142       10030.000000          1          4    1.882353
```

[53]:
```
df.groupby(['User_ID','Cust_name','Gender']).agg(
    TotalOrders=('Orders', 'count'),
    TotalOrderAmount=('Amount', 'sum'),
).sort_values(by='TotalOrders',ascending=False).head(10)
```

[53]:

User_ID	Cust_name	Gender	TotalOrders	TotalOrderAmount
1001680	Vasudev	M	24	281034.0
1003808	Vishakha	F	23	197660.0
1001941	Gopal	M	22	239147.0
1004425	Indulekha	F	20	194343.0
1003476	Lalita	F	19	220435.0
1006036	Halladay	M	19	158407.0
1004682	Vishakha	F	19	185122.0
1000424	Sudevi	F	19	187679.0
1002665	Champaklata	F	19	201104.0
1003410	Kamberova	F	17	128249.0

[66]:
```
df.groupby(['Gender', 'Marital_Status', 'Product_Category']).agg(
    TotalOrders=('Orders', 'count'),
    TotalOrderAmount=('Amount', 'sum')
).reset_index()\
.sort_values(by=['Gender', 'Marital_Status', 'TotalOrders'], ascending=[True,␣
 ↪True, False])\
.groupby(['Gender', 'Marital_Status']).first()\
.reset_index()\
.sort_values(by='TotalOrders', ascending=False)
```

[66]:

	Gender	Marital_Status	Product_Category	TotalOrders	TotalOrderAmount
0	F	0	Food	1079	14718268.5
1	F	1	Clothing & Apparel	780	4847131.0
2	M	0	Clothing & Apparel	461	2898890.0
3	M	1	Clothing & Apparel	340	2117026.0

[57]:
```
df.groupby(['Age_Group']).agg(
    TotalOrders=('Orders', 'count'),
    TotalOrderAmount=('Amount', 'sum'),
).sort_values(by='TotalOrders',ascending=False).head(10)
```

[57]:

Age Group	TotalOrders	TotalOrderAmount

```
26-35              4543       42613443.94
36-45              2286       22144995.49
18-25              1879       17240732.00
46-50               987        9207844.00
51-55               832        8261477.00
55+                 428        4080987.00
0-17                296        2699653.00
```

[68]:
```
df.groupby(['Age Group', 'Product_Category']).agg(
    TotalOrders=('Orders', 'count'),
    TotalOrderAmount=('Amount', 'sum'),
).reset_index().sort_values(['Age Group', 'TotalOrders'], ascending=[True,␣
 ↪False]).groupby('Age Group').first().reset_index().
 ↪sort_values(by='TotalOrders',ascending=False)
```

[68]:

	Age Group	Product_Category	TotalOrders	TotalOrderAmount
2	26-35	Clothing & Apparel	1057	6568722.0
3	36-45	Clothing & Apparel	532	3254060.0
1	18-25	Food	478	6578809.0
4	46-50	Clothing & Apparel	234	1450180.0
5	51-55	Clothing & Apparel	188	1196558.0
6	55+	Clothing & Apparel	103	626051.0
0	0-17	Food	83	1079989.0

[69]:
```
df.groupby(['Occupation', 'Product_Category']).agg(
    TotalOrders=('Orders', 'count'),
    TotalOrderAmount=('Amount', 'sum'),
).reset_index().sort_values(['Occupation', 'TotalOrders'], ascending=[True,␣
 ↪False]).groupby('Occupation').first().reset_index().
 ↪sort_values(by='TotalOrders',ascending=False)
```

[69]:

	Occupation	Product_Category	TotalOrders	TotalOrderAmount
8	Healthcare	Clothing & Apparel	361	2249961.0
10	IT Sector	Clothing & Apparel	352	2226012.0
2	Aviation	Food	312	4306343.0
3	Banking	Clothing & Apparel	274	1738400.0
7	Govt	Food	234	3242839.0
9	Hospitality	Clothing & Apparel	188	1131531.0
12	Media	Food	171	2351254.0
1	Automobile	Clothing & Apparel	136	870150.0
4	Chemical	Food	133	1821161.0
11	Lawyer	Food	130	1780666.0
13	Retail	Clothing & Apparel	123	744577.0
5	Construction	Clothing & Apparel	108	678794.0
6	Food Processing	Electronics & Gadgets	102	796316.0
14	Textile	Clothing & Apparel	91	545373.0
0	Agriculture	Clothing & Apparel	70	437296.0

[27]:
```
import pandas as pd

# Example data loading (comment if not required)
# df = pd.read_csv('your_dataset.csv')

def generate_state_insights(df):
    final_report = []  # Collect all insights here
    states = df['State'].unique()

    for state in states:
        state_insights = []  # Collect state-specific insights here
        state_df = df[df['State'] == state]

        print(f"\n{'='*80}")
        print(f"  STATE REPORT: {state}")
        print(f"{'='*80}")

        # --------- Users by Age Group ---------
        age_group_data = state_df.groupby('Age Group').agg(
            TotalUsers=('User_ID', 'count'),
            Females=('Gender', lambda x: (x == 'F').sum()),
            Males=('Gender', lambda x: (x == 'M').sum()),
            Married=('Marital_Status', lambda x: (x == 1).sum()),
            UnMarried=('Marital_Status', lambda x: (x == 0).sum())
        ).reset_index()

        top_age_group = age_group_data.sort_values(by='TotalUsers',␣
↪ascending=False).iloc[0]
        insight_age = (
            f"In {state}, the {top_age_group['Age Group']} age group makes up␣
↪the largest segment "
            f"with {top_age_group['TotalUsers']} users. Gender split:␣
↪{top_age_group['Males']} males and "
            f"{top_age_group['Females']} females. Married users are␣
↪{top_age_group['Married']}."
        )
        state_insights.append(insight_age)

        # --------- Users by Occupation ---------
        occupation_data = state_df.groupby('Occupation').agg(
            TotalUsers=('User_ID', 'count'),
            Females=('Gender', lambda x: (x == 'F').sum()),
            Males=('Gender', lambda x: (x == 'M').sum()),
            Married=('Marital_Status', lambda x: (x == 1).sum()),
            UnMarried=('Marital_Status', lambda x: (x == 0).sum())
        ).reset_index()
```

```
        top_occupation = occupation_data.sort_values(by='TotalUsers',␣
↪ascending=False).iloc[0]
        insight_occ = (
            f"In {state}, users from the '{top_occupation['Occupation']}'␣
↪occupation dominate with "
            f"{top_occupation['TotalUsers']} users, where␣
↪{top_occupation['Males']} are male and "
            f"{top_occupation['Females']} are female."
        )
        state_insights.append(insight_occ)

        # --------- Product Category by Orders ---------
        prod_cat_orders = state_df.groupby('Product_Category').agg(
            TotalOrders=('Orders', 'count')
        ).reset_index()

        top_prod_cat_orders = prod_cat_orders.sort_values(by='TotalOrders',␣
↪ascending=False).iloc[0]
        insight_prod_orders = (
            f"In {state}, the most ordered category is␣
↪'{top_prod_cat_orders['Product_Category']}' "
            f"with {top_prod_cat_orders['TotalOrders']} orders."
        )
        state_insights.append(insight_prod_orders)

        # --------- Product Category by Amount ---------
        prod_cat_amount = state_df.groupby('Product_Category').agg(
            TotalAmount=('Amount', 'sum')
        ).reset_index()

        top_prod_cat_amount = prod_cat_amount.sort_values(by='TotalAmount',␣
↪ascending=False).iloc[0]
        insight_prod_amount = (
            f"In {state}, the category with the highest sales amount is␣
↪'{top_prod_cat_amount['Product_Category']}' "
            f"with  {top_prod_cat_amount['TotalAmount']:.2f} in sales."
        )
        state_insights.append(insight_prod_amount)

        # --------- Product ID Analysis ---------
        prod_id_analysis = state_df.groupby('Product_ID').agg(
            TotalOrders=('Orders', 'count'),
            TotalOrderAmount=('Amount', 'sum')
        ).reset_index()
```

```
        top_prod_id = prod_id_analysis.sort_values(by='TotalOrders',␣
↪ascending=False).iloc[0]
        insight_prod_id = (
            f"In {state}, Product ID {top_prod_id['Product_ID']} had the␣
↪highest orders "
            f"({top_prod_id['TotalOrders']}) generating␣
↪ {top_prod_id['TotalOrderAmount']:.2f} in revenue."
        )
        state_insights.append(insight_prod_id)

        # --------- Top Users ---------
        top_users = state_df.groupby(['User_ID', 'Gender']).agg(
            TotalOrders=('Orders', 'count'),
            TotalOrderAmount=('Amount', 'sum')
        ).reset_index()

        top_user = top_users.sort_values(by='TotalOrders', ascending=False).
↪iloc[0]
        insight_top_user = (
            f"In {state}, the top user (User ID: {top_user['User_ID']}, Gender:␣
↪{top_user['Gender']}) "
            f"placed {top_user['TotalOrders']} orders totaling␣
↪ {top_user['TotalOrderAmount']:.2f}."
        )
        state_insights.append(insight_top_user)

        # --------- Gender & Marital Status Product Category ---------
        gm_prod_cat = state_df.groupby(['Gender', 'Marital_Status',␣
↪'Product_Category']).agg(
            TotalOrders=('Orders', 'count')
        ).reset_index()

        top_gm_prod = gm_prod_cat.sort_values(by='TotalOrders',␣
↪ascending=False).iloc[0]
        marital_status_str = 'Married' if top_gm_prod['Marital_Status'] == 1␣
↪else 'Unmarried'
        insight_gm_prod = (
            f"In {state}, {marital_status_str} {top_gm_prod['Gender']}s␣
↪primarily ordered "
            f"'{top_gm_prod['Product_Category']}' with␣
↪{top_gm_prod['TotalOrders']} orders."
        )
        state_insights.append(insight_gm_prod)

        # --------- Age Group & Product Category ---------
        age_prod = state_df.groupby(['Age Group', 'Product_Category']).agg(
```

```
            TotalOrders=('Orders', 'count')
        ).reset_index()

        top_age_prod = age_prod.sort_values(by='TotalOrders', ascending=False).
 ↪iloc[0]
        insight_age_prod = (
            f"In {state}, users aged {top_age_prod['Age Group']} favored␣
 ↪'{top_age_prod['Product_Category']}' "
            f"with {top_age_prod['TotalOrders']} orders."
        )
        state_insights.append(insight_age_prod)

        # --------- Occupation & Product Category ---------
        occ_prod = state_df.groupby(['Occupation', 'Product_Category']).agg(
            TotalOrders=('Orders', 'count')
        ).reset_index()

        top_occ_prod = occ_prod.sort_values(by='TotalOrders', ascending=False).
 ↪iloc[0]
        insight_occ_prod = (
            f"In {state}, people in '{top_occ_prod['Occupation']}' preferred␣
 ↪'{top_occ_prod['Product_Category']}' "
            f"with {top_occ_prod['TotalOrders']} orders."
        )
        state_insights.append(insight_occ_prod)

        # --------- Print and Collect State Insights ---------
        print("\n".join(state_insights))
        final_report.append({ 'State': state, 'Insights': state_insights })

    return final_report

# Run the report generator
state_insight_reports = generate_state_insights(df)

# If you want to display all reports together:
for report in state_insight_reports:
    print(f"\n\n{'#'*80}\nSTATE: {report['State']}\n{'#'*80}")
    for insight in report['Insights']:
        print(f"- {insight}")
```

```
================================================================================
 STATE REPORT: Maharashtra
================================================================================
In Maharashtra, the 26-35 age group makes up the largest segment with 645 users.
Gender split: 162 males and 483 females. Married users are 263.
```

```
In Maharashtra, users from the 'IT Sector' occupation dominate with 219 users,
where 73 are male and 146 are female.
In Maharashtra, the most ordered category is 'Food' with 480 orders.
In Maharashtra, the category with the highest sales amount is 'Food' with
 6421531.00 in sales.
In Maharashtra, Product ID P00110942 had the highest orders (11) generating
 100579.00 in revenue.
In Maharashtra, the top user (User ID: 1004425, Gender: F) placed 8 orders
totaling  87664.00.
In Maharashtra, Unmarried Fs primarily ordered 'Food' with 239 orders.
In Maharashtra, users aged 26-35 favored 'Food' with 192 orders.
In Maharashtra, people in 'Healthcare' preferred 'Food' with 77 orders.

================================================================================
  STATE REPORT: Andhra Pradesh
================================================================================
In Andhra Pradesh, the 26-35 age group makes up the largest segment with 319
users. Gender split: 97 males and 222 females. Married users are 129.
In Andhra Pradesh, users from the 'IT Sector' occupation dominate with 115
users, where 37 are male and 78 are female.
In Andhra Pradesh, the most ordered category is 'Electronics & Gadgets' with 277
orders.
In Andhra Pradesh, the category with the highest sales amount is 'Food' with
 2163209.00 in sales.
In Andhra Pradesh, Product ID P00080342 had the highest orders (5) generating
 58280.00 in revenue.
In Andhra Pradesh, the top user (User ID: 1000752, Gender: M) placed 4 orders
totaling  49114.00.
In Andhra Pradesh, Unmarried Fs primarily ordered 'Electronics & Gadgets' with
118 orders.
In Andhra Pradesh, users aged 26-35 favored 'Electronics & Gadgets' with 112
orders.
In Andhra Pradesh, people in 'IT Sector' preferred 'Electronics & Gadgets' with
40 orders.

================================================================================
  STATE REPORT: Uttar Pradesh
================================================================================
In Uttar Pradesh, the 26-35 age group makes up the largest segment with 784
users. Gender split: 214 males and 570 females. Married users are 346.
In Uttar Pradesh, users from the 'IT Sector' occupation dominate with 280 users,
where 98 are male and 182 are female.
In Uttar Pradesh, the most ordered category is 'Food' with 569 orders.
In Uttar Pradesh, the category with the highest sales amount is 'Food' with
 7983142.00 in sales.
In Uttar Pradesh, Product ID P00138542 had the highest orders (10) generating
 110191.00 in revenue.
In Uttar Pradesh, the top user (User ID: 1001680, Gender: M) placed 6 orders
```

```
totaling  71230.00.
In Uttar Pradesh, Unmarried Fs primarily ordered 'Food' with 261 orders.
In Uttar Pradesh, users aged 26-35 favored 'Food' with 243 orders.
In Uttar Pradesh, people in 'IT Sector' preferred 'Food' with 82 orders.

================================================================================
  STATE REPORT: Karnataka
================================================================================
In Karnataka, the 26-35 age group makes up the largest segment with 538 users.
Gender split: 128 males and 410 females. Married users are 237.
In Karnataka, users from the 'IT Sector' occupation dominate with 176 users,
where 54 are male and 122 are female.
In Karnataka, the most ordered category is 'Footwear & Shoes' with 339 orders.
In Karnataka, the category with the highest sales amount is 'Footwear & Shoes'
with  4963928.00 in sales.
In Karnataka, Product ID P00117942 had the highest orders (8) generating
 89275.00 in revenue.
In Karnataka, the top user (User ID: 1000424, Gender: F) placed 5 orders
totaling  49623.00.
In Karnataka, Unmarried Fs primarily ordered 'Footwear & Shoes' with 168 orders.
In Karnataka, users aged 26-35 favored 'Footwear & Shoes' with 162 orders.
In Karnataka, people in 'IT Sector' preferred 'Footwear & Shoes' with 58 orders.

================================================================================
  STATE REPORT: Gujarat
================================================================================
In Gujarat, the 26-35 age group makes up the largest segment with 165 users.
Gender split: 50 males and 115 females. Married users are 62.
In Gujarat, users from the 'Aviation' occupation dominate with 65 users, where
13 are male and 52 are female.
In Gujarat, the most ordered category is 'Clothing & Apparel' with 113 orders.
In Gujarat, the category with the highest sales amount is 'Food' with
 1342541.00 in sales.
In Gujarat, Product ID P00129642 had the highest orders (4) generating  56561.00
in revenue.
In Gujarat, the top user (User ID: 1001422, Gender: F) placed 4 orders totaling
 33527.00.
In Gujarat, Unmarried Fs primarily ordered 'Clothing & Apparel' with 61 orders.
In Gujarat, users aged 26-35 favored 'Clothing & Apparel' with 46 orders.
In Gujarat, people in 'Aviation' preferred 'Food' with 20 orders.

================================================================================
  STATE REPORT: Himachal Pradesh
================================================================================
In Himachal Pradesh, the 26-35 age group makes up the largest segment with 248
users. Gender split: 70 males and 178 females. Married users are 97.
In Himachal Pradesh, users from the 'Healthcare' occupation dominate with 80
users, where 30 are male and 50 are female.
```

```
In Himachal Pradesh, the most ordered category is 'Clothing & Apparel' with 240
orders.
In Himachal Pradesh, the category with the highest sales amount is 'Clothing &
Apparel' with  1445132.00 in sales.
In Himachal Pradesh, Product ID P00059442 had the highest orders (4) generating
 38010.00 in revenue.
In Himachal Pradesh, the top user (User ID: 1003808, Gender: F) placed 4 orders
totaling  28913.00.
In Himachal Pradesh, Unmarried Fs primarily ordered 'Clothing & Apparel' with
112 orders.
In Himachal Pradesh, users aged 26-35 favored 'Clothing & Apparel' with 88
orders.
In Himachal Pradesh, people in 'IT Sector' preferred 'Clothing & Apparel' with
32 orders.

================================================================================
  STATE REPORT: Delhi
================================================================================
In Delhi, the 26-35 age group makes up the largest segment with 472 users.
Gender split: 146 males and 326 females. Married users are 219.
In Delhi, users from the 'IT Sector' occupation dominate with 174 users, where
63 are male and 111 are female.
In Delhi, the most ordered category is 'Footwear & Shoes' with 338 orders.
In Delhi, the category with the highest sales amount is 'Footwear & Shoes' with
 5027449.45 in sales.
In Delhi, Product ID P00184942 had the highest orders (9) generating  104806.00
in revenue.
In Delhi, the top user (User ID: 1001899, Gender: M) placed 5 orders totaling
 37066.00.
In Delhi, Unmarried Fs primarily ordered 'Footwear & Shoes' with 175 orders.
In Delhi, users aged 26-35 favored 'Footwear & Shoes' with 148 orders.
In Delhi, people in 'IT Sector' preferred 'Footwear & Shoes' with 53 orders.

================================================================================
  STATE REPORT: Madhya Pradesh
================================================================================
In Madhya Pradesh, the 26-35 age group makes up the largest segment with 343
users. Gender split: 111 males and 232 females. Married users are 152.
In Madhya Pradesh, users from the 'IT Sector' occupation dominate with 142
users, where 40 are male and 102 are female.
In Madhya Pradesh, the most ordered category is 'Food' with 210 orders.
In Madhya Pradesh, the category with the highest sales amount is 'Food' with
 2821970.00 in sales.
In Madhya Pradesh, Product ID P00110942 had the highest orders (6) generating
 51831.00 in revenue.
In Madhya Pradesh, the top user (User ID: 1003618, Gender: M) placed 4 orders
totaling  46414.00.
In Madhya Pradesh, Unmarried Fs primarily ordered 'Beauty' with 103 orders.
```

```
In Madhya Pradesh, users aged 26-35 favored 'Food' with 78 orders.
In Madhya Pradesh, people in 'IT Sector' preferred 'Beauty' with 39 orders.

================================================================================
  STATE REPORT: Jharkhand
================================================================================
In Jharkhand, the 26-35 age group makes up the largest segment with 148 users.
Gender split: 37 males and 111 females. Married users are 58.
In Jharkhand, users from the 'Banking' occupation dominate with 50 users, where
15 are male and 35 are female.
In Jharkhand, the most ordered category is 'Clothing & Apparel' with 148 orders.
In Jharkhand, the category with the highest sales amount is 'Electronics &
Gadgets' with  913742.00 in sales.
In Jharkhand, Product ID P00265242 had the highest orders (4) generating
 23669.00 in revenue.
In Jharkhand, the top user (User ID: 1002038, Gender: M) placed 3 orders
totaling  17498.00.
In Jharkhand, Unmarried Fs primarily ordered 'Clothing & Apparel' with 71
orders.
In Jharkhand, users aged 26-35 favored 'Clothing & Apparel' with 60 orders.
In Jharkhand, people in 'Banking' preferred 'Electronics & Gadgets' with 23
orders.

================================================================================
  STATE REPORT: Kerala
================================================================================
In Kerala, the 26-35 age group makes up the largest segment with 192 users.
Gender split: 56 males and 136 females. Married users are 86.
In Kerala, users from the 'Healthcare' occupation dominate with 58 users, where
19 are male and 39 are female.
In Kerala, the most ordered category is 'Clothing & Apparel' with 182 orders.
In Kerala, the category with the highest sales amount is 'Clothing & Apparel'
with  1129045.00 in sales.
In Kerala, Product ID P00265242 had the highest orders (5) generating  29084.00
in revenue.
In Kerala, the top user (User ID: 1000329, Gender: F) placed 3 orders totaling
 27696.00.
In Kerala, Unmarried Fs primarily ordered 'Clothing & Apparel' with 82 orders.
In Kerala, users aged 26-35 favored 'Clothing & Apparel' with 78 orders.
In Kerala, people in 'Healthcare' preferred 'Clothing & Apparel' with 30 orders.

================================================================================
  STATE REPORT: Haryana
================================================================================
In Haryana, the 26-35 age group makes up the largest segment with 182 users.
Gender split: 49 males and 133 females. Married users are 67.
In Haryana, users from the 'IT Sector' occupation dominate with 67 users, where
20 are male and 47 are female.
```

```
In Haryana, the most ordered category is 'Clothing & Apparel' with 123 orders.
In Haryana, the category with the highest sales amount is 'Food' with
 1678205.00 in sales.
In Haryana, Product ID P00321042 had the highest orders (4) generating  23351.00
in revenue.
In Haryana, the top user (User ID: 1005643, Gender: M) placed 4 orders totaling
 21740.00.
In Haryana, Unmarried Fs primarily ordered 'Food' with 66 orders.
In Haryana, users aged 26-35 favored 'Food' with 53 orders.
In Haryana, people in 'IT Sector' preferred 'Food' with 23 orders.

================================================================================
  STATE REPORT: Bihar
================================================================================
In Bihar, the 26-35 age group makes up the largest segment with 173 users.
Gender split: 44 males and 129 females. Married users are 78.
In Bihar, users from the 'IT Sector' occupation dominate with 64 users, where 19
are male and 45 are female.
In Bihar, the most ordered category is 'Clothing & Apparel' with 140 orders.
In Bihar, the category with the highest sales amount is 'Food' with  1555848.00
in sales.
In Bihar, Product ID P00002242 had the highest orders (4) generating  31441.00
in revenue.
In Bihar, the top user (User ID: 1004448, Gender: F) placed 3 orders totaling
 17481.00.
In Bihar, Unmarried Fs primarily ordered 'Clothing & Apparel' with 63 orders.
In Bihar, users aged 26-35 favored 'Clothing & Apparel' with 54 orders.
In Bihar, people in 'Healthcare' preferred 'Clothing & Apparel' with 23 orders.

================================================================================
  STATE REPORT: Rajasthan
================================================================================
In Rajasthan, the 26-35 age group makes up the largest segment with 95 users.
Gender split: 32 males and 63 females. Married users are 42.
In Rajasthan, users from the 'IT Sector' occupation dominate with 34 users,
where 11 are male and 23 are female.
In Rajasthan, the most ordered category is 'Electronics & Gadgets' with 134
orders.
In Rajasthan, the category with the highest sales amount is 'Electronics &
Gadgets' with  999550.00 in sales.
In Rajasthan, Product ID P00057442 had the highest orders (2) generating
 19784.00 in revenue.
In Rajasthan, the top user (User ID: 1006036, Gender: M) placed 2 orders
totaling  13600.00.
In Rajasthan, Unmarried Fs primarily ordered 'Electronics & Gadgets' with 62
orders.
In Rajasthan, users aged 26-35 favored 'Electronics & Gadgets' with 58 orders.
In Rajasthan, people in 'IT Sector' preferred 'Electronics & Gadgets' with 22
```

orders.

```
================================================================================
  STATE REPORT: Uttarakhand
================================================================================
In Uttarakhand, the 26-35 age group makes up the largest segment with 111 users.
Gender split: 32 males and 79 females. Married users are 44.
In Uttarakhand, users from the 'IT Sector' occupation dominate with 42 users,
where 14 are male and 28 are female.
In Uttarakhand, the most ordered category is 'Clothing & Apparel' with 157
orders.
In Uttarakhand, the category with the highest sales amount is 'Clothing &
Apparel' with  972979.00 in sales.
In Uttarakhand, Product ID P00073642 had the highest orders (4) generating
 58518.00 in revenue.
In Uttarakhand, the top user (User ID: 1005604, Gender: F) placed 2 orders
totaling  22059.00.
In Uttarakhand, Unmarried Fs primarily ordered 'Clothing & Apparel' with 76
orders.
In Uttarakhand, users aged 26-35 favored 'Clothing & Apparel' with 52 orders.
In Uttarakhand, people in 'Media' preferred 'Clothing & Apparel' with 23 orders.

================================================================================
  STATE REPORT: Telangana
================================================================================
In Telangana, the 26-35 age group makes up the largest segment with 45 users.
Gender split: 20 males and 25 females. Married users are 17.
In Telangana, users from the 'Healthcare' occupation dominate with 17 users,
where 7 are male and 10 are female.
In Telangana, the most ordered category is 'Electronics & Gadgets' with 59
orders.
In Telangana, the category with the highest sales amount is 'Food' with
 496602.00 in sales.
In Telangana, Product ID P00238542 had the highest orders (2) generating
 17656.00 in revenue.
In Telangana, the top user (User ID: 1005759, Gender: M) placed 2 orders
totaling  12935.00.
In Telangana, Unmarried Fs primarily ordered 'Electronics & Gadgets' with 19
orders.
In Telangana, users aged 26-35 favored 'Electronics & Gadgets' with 26 orders.
In Telangana, people in 'Healthcare' preferred 'Electronics & Gadgets' with 11
orders.

================================================================================
  STATE REPORT: Punjab
================================================================================
In Punjab, the 26-35 age group makes up the largest segment with 83 users.
Gender split: 24 males and 59 females. Married users are 35.
```

In Punjab, users from the 'IT Sector' occupation dominate with 27 users, where 8 are male and 19 are female.
In Punjab, the most ordered category is 'Electronics & Gadgets' with 126 orders.
In Punjab, the category with the highest sales amount is 'Electronics & Gadgets' with 917585.00 in sales.
In Punjab, Product ID P00001142 had the highest orders (2) generating 14174.00 in revenue.
In Punjab, the top user (User ID: 1005472, Gender: F) placed 3 orders totaling 23813.00.
In Punjab, Unmarried Fs primarily ordered 'Electronics & Gadgets' with 67 orders.
In Punjab, users aged 26-35 favored 'Electronics & Gadgets' with 52 orders.
In Punjab, people in 'IT Sector' preferred 'Electronics & Gadgets' with 19 orders.

##
STATE: Maharashtra
##
- In Maharashtra, the 26-35 age group makes up the largest segment with 645 users. Gender split: 162 males and 483 females. Married users are 263.
- In Maharashtra, users from the 'IT Sector' occupation dominate with 219 users, where 73 are male and 146 are female.
- In Maharashtra, the most ordered category is 'Food' with 480 orders.
- In Maharashtra, the category with the highest sales amount is 'Food' with 6421531.00 in sales.
- In Maharashtra, Product ID P00110942 had the highest orders (11) generating 100579.00 in revenue.
- In Maharashtra, the top user (User ID: 1004425, Gender: F) placed 8 orders totaling 87664.00.
- In Maharashtra, Unmarried Fs primarily ordered 'Food' with 239 orders.
- In Maharashtra, users aged 26-35 favored 'Food' with 192 orders.
- In Maharashtra, people in 'Healthcare' preferred 'Food' with 77 orders.

##
STATE: Andhra Pradesh
##
- In Andhra Pradesh, the 26-35 age group makes up the largest segment with 319 users. Gender split: 97 males and 222 females. Married users are 129.
- In Andhra Pradesh, users from the 'IT Sector' occupation dominate with 115 users, where 37 are male and 78 are female.
- In Andhra Pradesh, the most ordered category is 'Electronics & Gadgets' with 277 orders.
- In Andhra Pradesh, the category with the highest sales amount is 'Food' with 2163209.00 in sales.
- In Andhra Pradesh, Product ID P00080342 had the highest orders (5) generating 58280.00 in revenue.

- In Andhra Pradesh, the top user (User ID: 1000752, Gender: M) placed 4 orders totaling 49114.00.
- In Andhra Pradesh, Unmarried Fs primarily ordered 'Electronics & Gadgets' with 118 orders.
- In Andhra Pradesh, users aged 26-35 favored 'Electronics & Gadgets' with 112 orders.
- In Andhra Pradesh, people in 'IT Sector' preferred 'Electronics & Gadgets' with 40 orders.

##
STATE: Uttar Pradesh
##
- In Uttar Pradesh, the 26-35 age group makes up the largest segment with 784 users. Gender split: 214 males and 570 females. Married users are 346.
- In Uttar Pradesh, users from the 'IT Sector' occupation dominate with 280 users, where 98 are male and 182 are female.
- In Uttar Pradesh, the most ordered category is 'Food' with 569 orders.
- In Uttar Pradesh, the category with the highest sales amount is 'Food' with 7983142.00 in sales.
- In Uttar Pradesh, Product ID P00138542 had the highest orders (10) generating 110191.00 in revenue.
- In Uttar Pradesh, the top user (User ID: 1001680, Gender: M) placed 6 orders totaling 71230.00.
- In Uttar Pradesh, Unmarried Fs primarily ordered 'Food' with 261 orders.
- In Uttar Pradesh, users aged 26-35 favored 'Food' with 243 orders.
- In Uttar Pradesh, people in 'IT Sector' preferred 'Food' with 82 orders.

##
STATE: Karnataka
##
- In Karnataka, the 26-35 age group makes up the largest segment with 538 users. Gender split: 128 males and 410 females. Married users are 237.
- In Karnataka, users from the 'IT Sector' occupation dominate with 176 users, where 54 are male and 122 are female.
- In Karnataka, the most ordered category is 'Footwear & Shoes' with 339 orders.
- In Karnataka, the category with the highest sales amount is 'Footwear & Shoes' with 4963928.00 in sales.
- In Karnataka, Product ID P00117942 had the highest orders (8) generating 89275.00 in revenue.
- In Karnataka, the top user (User ID: 1000424, Gender: F) placed 5 orders totaling 49623.00.
- In Karnataka, Unmarried Fs primarily ordered 'Footwear & Shoes' with 168 orders.
- In Karnataka, users aged 26-35 favored 'Footwear & Shoes' with 162 orders.
- In Karnataka, people in 'IT Sector' preferred 'Footwear & Shoes' with 58 orders.

```
################################################################################
STATE: Gujarat
################################################################################
- In Gujarat, the 26-35 age group makes up the largest segment with 165 users.
Gender split: 50 males and 115 females. Married users are 62.
- In Gujarat, users from the 'Aviation' occupation dominate with 65 users, where
13 are male and 52 are female.
- In Gujarat, the most ordered category is 'Clothing & Apparel' with 113 orders.
- In Gujarat, the category with the highest sales amount is 'Food' with
 1342541.00 in sales.
- In Gujarat, Product ID P00129642 had the highest orders (4) generating
 56561.00 in revenue.
- In Gujarat, the top user (User ID: 1001422, Gender: F) placed 4 orders
totaling  33527.00.
- In Gujarat, Unmarried Fs primarily ordered 'Clothing & Apparel' with 61
orders.
- In Gujarat, users aged 26-35 favored 'Clothing & Apparel' with 46 orders.
- In Gujarat, people in 'Aviation' preferred 'Food' with 20 orders.


################################################################################
STATE: Himachal Pradesh
################################################################################
- In Himachal Pradesh, the 26-35 age group makes up the largest segment with 248
users. Gender split: 70 males and 178 females. Married users are 97.
- In Himachal Pradesh, users from the 'Healthcare' occupation dominate with 80
users, where 30 are male and 50 are female.
- In Himachal Pradesh, the most ordered category is 'Clothing & Apparel' with
240 orders.
- In Himachal Pradesh, the category with the highest sales amount is 'Clothing &
Apparel' with  1445132.00 in sales.
- In Himachal Pradesh, Product ID P00059442 had the highest orders (4)
generating  38010.00 in revenue.
- In Himachal Pradesh, the top user (User ID: 1003808, Gender: F) placed 4
orders totaling  28913.00.
- In Himachal Pradesh, Unmarried Fs primarily ordered 'Clothing & Apparel' with
112 orders.
- In Himachal Pradesh, users aged 26-35 favored 'Clothing & Apparel' with 88
orders.
- In Himachal Pradesh, people in 'IT Sector' preferred 'Clothing & Apparel' with
32 orders.


################################################################################
STATE: Delhi
################################################################################
```

- In Delhi, the 26-35 age group makes up the largest segment with 472 users. Gender split: 146 males and 326 females. Married users are 219.
- In Delhi, users from the 'IT Sector' occupation dominate with 174 users, where 63 are male and 111 are female.
- In Delhi, the most ordered category is 'Footwear & Shoes' with 338 orders.
- In Delhi, the category with the highest sales amount is 'Footwear & Shoes' with 5027449.45 in sales.
- In Delhi, Product ID P00184942 had the highest orders (9) generating 104806.00 in revenue.
- In Delhi, the top user (User ID: 1001899, Gender: M) placed 5 orders totaling 37066.00.
- In Delhi, Unmarried Fs primarily ordered 'Footwear & Shoes' with 175 orders.
- In Delhi, users aged 26-35 favored 'Footwear & Shoes' with 148 orders.
- In Delhi, people in 'IT Sector' preferred 'Footwear & Shoes' with 53 orders.

##
STATE: Madhya Pradesh
##
- In Madhya Pradesh, the 26-35 age group makes up the largest segment with 343 users. Gender split: 111 males and 232 females. Married users are 152.
- In Madhya Pradesh, users from the 'IT Sector' occupation dominate with 142 users, where 40 are male and 102 are female.
- In Madhya Pradesh, the most ordered category is 'Food' with 210 orders.
- In Madhya Pradesh, the category with the highest sales amount is 'Food' with 2821970.00 in sales.
- In Madhya Pradesh, Product ID P00110942 had the highest orders (6) generating 51831.00 in revenue.
- In Madhya Pradesh, the top user (User ID: 1003618, Gender: M) placed 4 orders totaling 46414.00.
- In Madhya Pradesh, Unmarried Fs primarily ordered 'Beauty' with 103 orders.
- In Madhya Pradesh, users aged 26-35 favored 'Food' with 78 orders.
- In Madhya Pradesh, people in 'IT Sector' preferred 'Beauty' with 39 orders.

##
STATE: Jharkhand
##
- In Jharkhand, the 26-35 age group makes up the largest segment with 148 users. Gender split: 37 males and 111 females. Married users are 58.
- In Jharkhand, users from the 'Banking' occupation dominate with 50 users, where 15 are male and 35 are female.
- In Jharkhand, the most ordered category is 'Clothing & Apparel' with 148 orders.
- In Jharkhand, the category with the highest sales amount is 'Electronics & Gadgets' with 913742.00 in sales.
- In Jharkhand, Product ID P00265242 had the highest orders (4) generating 23669.00 in revenue.

- In Jharkhand, the top user (User ID: 1002038, Gender: M) placed 3 orders totaling 17498.00.
- In Jharkhand, Unmarried Fs primarily ordered 'Clothing & Apparel' with 71 orders.
- In Jharkhand, users aged 26-35 favored 'Clothing & Apparel' with 60 orders.
- In Jharkhand, people in 'Banking' preferred 'Electronics & Gadgets' with 23 orders.

##
STATE: Kerala
##
- In Kerala, the 26-35 age group makes up the largest segment with 192 users. Gender split: 56 males and 136 females. Married users are 86.
- In Kerala, users from the 'Healthcare' occupation dominate with 58 users, where 19 are male and 39 are female.
- In Kerala, the most ordered category is 'Clothing & Apparel' with 182 orders.
- In Kerala, the category with the highest sales amount is 'Clothing & Apparel' with 1129045.00 in sales.
- In Kerala, Product ID P00265242 had the highest orders (5) generating 29084.00 in revenue.
- In Kerala, the top user (User ID: 1000329, Gender: F) placed 3 orders totaling 27696.00.
- In Kerala, Unmarried Fs primarily ordered 'Clothing & Apparel' with 82 orders.
- In Kerala, users aged 26-35 favored 'Clothing & Apparel' with 78 orders.
- In Kerala, people in 'Healthcare' preferred 'Clothing & Apparel' with 30 orders.

##
STATE: Haryana
##
- In Haryana, the 26-35 age group makes up the largest segment with 182 users. Gender split: 49 males and 133 females. Married users are 67.
- In Haryana, users from the 'IT Sector' occupation dominate with 67 users, where 20 are male and 47 are female.
- In Haryana, the most ordered category is 'Clothing & Apparel' with 123 orders.
- In Haryana, the category with the highest sales amount is 'Food' with 1678205.00 in sales.
- In Haryana, Product ID P00321042 had the highest orders (4) generating 23351.00 in revenue.
- In Haryana, the top user (User ID: 1005643, Gender: M) placed 4 orders totaling 21740.00.
- In Haryana, Unmarried Fs primarily ordered 'Food' with 66 orders.
- In Haryana, users aged 26-35 favored 'Food' with 53 orders.
- In Haryana, people in 'IT Sector' preferred 'Food' with 23 orders.

```
################################################################################
STATE: Bihar
################################################################################
- In Bihar, the 26-35 age group makes up the largest segment with 173 users.
Gender split: 44 males and 129 females. Married users are 78.
- In Bihar, users from the 'IT Sector' occupation dominate with 64 users, where
19 are male and 45 are female.
- In Bihar, the most ordered category is 'Clothing & Apparel' with 140 orders.
- In Bihar, the category with the highest sales amount is 'Food' with
 1555848.00 in sales.
- In Bihar, Product ID P00002242 had the highest orders (4) generating  31441.00
in revenue.
- In Bihar, the top user (User ID: 1004448, Gender: F) placed 3 orders totaling
 17481.00.
- In Bihar, Unmarried Fs primarily ordered 'Clothing & Apparel' with 63 orders.
- In Bihar, users aged 26-35 favored 'Clothing & Apparel' with 54 orders.
- In Bihar, people in 'Healthcare' preferred 'Clothing & Apparel' with 23
orders.


################################################################################
STATE: Rajasthan
################################################################################
- In Rajasthan, the 26-35 age group makes up the largest segment with 95 users.
Gender split: 32 males and 63 females. Married users are 42.
- In Rajasthan, users from the 'IT Sector' occupation dominate with 34 users,
where 11 are male and 23 are female.
- In Rajasthan, the most ordered category is 'Electronics & Gadgets' with 134
orders.
- In Rajasthan, the category with the highest sales amount is 'Electronics &
Gadgets' with  999550.00 in sales.
- In Rajasthan, Product ID P00057442 had the highest orders (2) generating
 19784.00 in revenue.
- In Rajasthan, the top user (User ID: 1006036, Gender: M) placed 2 orders
totaling  13600.00.
- In Rajasthan, Unmarried Fs primarily ordered 'Electronics & Gadgets' with 62
orders.
- In Rajasthan, users aged 26-35 favored 'Electronics & Gadgets' with 58 orders.
- In Rajasthan, people in 'IT Sector' preferred 'Electronics & Gadgets' with 22
orders.


################################################################################
STATE: Uttarakhand
################################################################################
- In Uttarakhand, the 26-35 age group makes up the largest segment with 111
users. Gender split: 32 males and 79 females. Married users are 44.
- In Uttarakhand, users from the 'IT Sector' occupation dominate with 42 users,
```

where 14 are male and 28 are female.
- In Uttarakhand, the most ordered category is 'Clothing & Apparel' with 157 orders.
- In Uttarakhand, the category with the highest sales amount is 'Clothing & Apparel' with 972979.00 in sales.
- In Uttarakhand, Product ID P00073642 had the highest orders (4) generating 58518.00 in revenue.
- In Uttarakhand, the top user (User ID: 1005604, Gender: F) placed 2 orders totaling 22059.00.
- In Uttarakhand, Unmarried Fs primarily ordered 'Clothing & Apparel' with 76 orders.
- In Uttarakhand, users aged 26-35 favored 'Clothing & Apparel' with 52 orders.
- In Uttarakhand, people in 'Media' preferred 'Clothing & Apparel' with 23 orders.

##
STATE: Telangana
##
- In Telangana, the 26-35 age group makes up the largest segment with 45 users. Gender split: 20 males and 25 females. Married users are 17.
- In Telangana, users from the 'Healthcare' occupation dominate with 17 users, where 7 are male and 10 are female.
- In Telangana, the most ordered category is 'Electronics & Gadgets' with 59 orders.
- In Telangana, the category with the highest sales amount is 'Food' with 496602.00 in sales.
- In Telangana, Product ID P00238542 had the highest orders (2) generating 17656.00 in revenue.
- In Telangana, the top user (User ID: 1005759, Gender: M) placed 2 orders totaling 12935.00.
- In Telangana, Unmarried Fs primarily ordered 'Electronics & Gadgets' with 19 orders.
- In Telangana, users aged 26-35 favored 'Electronics & Gadgets' with 26 orders.
- In Telangana, people in 'Healthcare' preferred 'Electronics & Gadgets' with 11 orders.

##
STATE: Punjab
##
- In Punjab, the 26-35 age group makes up the largest segment with 83 users. Gender split: 24 males and 59 females. Married users are 35.
- In Punjab, users from the 'IT Sector' occupation dominate with 27 users, where 8 are male and 19 are female.
- In Punjab, the most ordered category is 'Electronics & Gadgets' with 126 orders.
- In Punjab, the category with the highest sales amount is 'Electronics &

```
Gadgets' with  917585.00 in sales.
- In Punjab, Product ID P00001142 had the highest orders (2) generating
 14174.00 in revenue.
- In Punjab, the top user (User ID: 1005472, Gender: F) placed 3 orders totaling
 23813.00.
- In Punjab, Unmarried Fs primarily ordered 'Electronics & Gadgets' with 67
orders.
- In Punjab, users aged 26-35 favored 'Electronics & Gadgets' with 52 orders.
- In Punjab, people in 'IT Sector' preferred 'Electronics & Gadgets' with 19
orders.
```

1 Closure Analysis

1.1 Demographic Highlights

- **Dominant Age Group**: The 26-35 age group is consistently the largest customer segment across all states
- **Gender Distribution**: Female users significantly outnumber male users in all states (roughly 3:1 ratio in most regions)
- **Marital Status**: Unmarried users form the majority of the customer base in most states

1.2 Industry Insights

- **Dominant Occupation**: IT Sector employees are the primary customer base in 12 out of 16 states
- **Regional Variations**:
  - Gujarat uniquely shows Aviation sector dominance
  - Healthcare professionals lead in Kerala, Himachal Pradesh, and Telangana
  - Banking sector professionals dominate in Jharkhand

1.3 Product Category Analysis

- **Top Categories by Orders**:
  - Food: Leading category in Maharashtra, Uttar Pradesh, Madhya Pradesh (4 states)
  - Clothing & Apparel: Leading category in Gujarat, Himachal Pradesh, Jharkhand, Kerala, Haryana, Bihar, Uttarakhand (7 states)
  - Electronics & Gadgets: Leading category in Andhra Pradesh, Rajasthan, Telangana, Punjab (4 states)
  - Footwear & Shoes: Leading category in Karnataka and Delhi (2 states)
- **Highest Revenue Generators**:
  - Food: Top revenue category in 7 states despite being the most ordered category in only 4 states
  - Footwear & Shoes: Generated highest revenue in Karnataka and Delhi
  - Clothing & Apparel: Highest revenue in 4 states
  - Electronics & Gadgets: Highest revenue in 3 states

1.4 Consumer Behavior Patterns

- **Unmarried Female Preferences**:

  - Clothing & Apparel: Preferred by unmarried females in 7 states
  - Electronics & Gadgets: Preferred in 5 states
  - Food: Preferred in 3 states
  - Footwear & Shoes: Preferred in 2 states
  - Beauty: Preferred in 1 state (Madhya Pradesh)
- **Regional Spending Power**:
  - Uttar Pradesh: Highest total food sales (7,983,142)
  - Maharashtra: Second highest in food sales (6,421,531)
  - Delhi: Highest footwear sales (5,027,449)

1.5 Industry-Specific Purchasing Patterns

- IT Sector professionals show varied preferences across states:
  - Food preferred in 5 states
  - Electronics & Gadgets preferred in 4 states
  - Clothing & Apparel preferred in 2 states
  - Footwear & Shoes preferred in 2 states
  - Beauty preferred in 1 state

1.6 High-Value Customers

- Top spenders are predominantly female in most states
- Highest individual spending: User ID 1004425 (Female) in Maharashtra with 87,664 across 8 orders

1.7 Product Performance

- Highest performing product: P00138542 in Uttar Pradesh with 10 orders generating 110,191
- Second highest: P00184942 in Delhi with 9 orders generating 104,806

[ ]: