ANP-D0449

DATA ANALYTICS USING PYTHON

CREDIT CARD USAGE TRENDS ACROSS DIFFERENT DEMOGRAPHICS

SUBMITTED BY:

SANDHIYA N

Abstract:

Credit card transactions provide valuable insights into consumer spending behavior across different demographic groups. This study analyzes credit card usage trends based on factors such as income level, occupation, and geographical location. By leveraging data analytics techniques such as descriptive statistics, trend analysis, and correlation analysis, we examine variations in spending patterns, repayment behavior, and credit utilization rates.

For example, high-income professionals tend to use credit cards for travel and luxury purchases, whereas lower-income groups rely on credit for daily necessities, often carrying higher balances. Additionally, urban consumers show a preference for digital payments, while rural users may still rely more on traditional banking methods.

The findings from this study can help financial institutions, policymakers, and businesses better understand consumer financial behavior, enabling them to design more effective credit policies, financial literacy programs, and targeted marketing strategies. By identifying key trends and spending habits, this research contributes to a data-driven approach for improving financial decision-making at both individual and institutional levels.

Problem Statement:

"Analyzing Credit Card Usage Trends Across Different Demographics to Understand Spending Behavior and Financial Patterns"

Context:

Credit card transactions serve as a crucial indicator of consumer spending behavior. However, spending habits, credit utilization, and repayment patterns vary significantly across different demographic groups, including age, income levels, occupation, and geographic location. Understanding these variations can help financial institutions, policymakers, and businesses tailor financial products, optimize risk assessment models, and promote responsible credit usage.

Objective:

This study aims to analyze credit card usage trends among different demographic groups to identify key spending patterns, credit utilization behaviors, and financial habits. By leveraging data analytics, we seek to answer:

- How do spending habits differ across age groups and income levels?
- What are the common categories of expenditure for different demographics?
- How does credit utilization vary across different geographical locations ?

Significance:

The insights derived from this analysis will assist:

- Banks & Financial Institutions in designing better credit policies and personalized financial products.
- Businesses & Marketers in targeting specific customer segments with relevant offers.
- Policymakers in developing financial literacy programs to promote responsible credit usage

Libraries used:

Pandas (import pandas as pd)

Pandas is a powerful data analysis library designed for handling and manipulating structured data, such as tables and spreadsheets. It provides efficient tools for data cleaning, transformation, and aggregation.

- Key Features:

1.Works with DataFrames (tabular data) and Series (1D data).
2.Provides functions for data cleaning, handling missing values, and filtering.
3.Supports grouping and aggregation for summarizing data.
4.Allows merging and joining multiple datasets.
5.Reads and writes data in various formats like CSV, Excel, and JSON.

Matplotlib (import matplotlib.pyplot as plt)

Matplotlib is a basic visualization library used to create static, animated, and interactive plots. It provides extensive control over graph elements, making it highly customizable for various types of charts.

- Key Features:

1.Supports line charts, bar plots, scatter plots, histograms, and more.
2.Allows full customization of titles, labels, legends, and colors.
3.Enables the creation of subplots for multiple charts in one figure.
4.Can export charts in multiple formats, including PNG, JPG, and PDF.

Seaborn (import seaborn as sns)

Seaborn is a statistical data visualization library built on top of Matplotlib. It provides elegant and easy-to-use functions for creating visually appealing and informative charts.

- Key Features:

1.Offers built-in themes and color palettes for aesthetic charts.
2.Supports statistical plots like heatmaps, boxplots, violin plots, and regression plots.
3.Integrates seamlessly with Pandas DataFrames for efficient visualization.
4.Automatically handles data aggregation and categorical plotting.

Code:

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

#Reading the CSV file

```
data=pd.read_csv("/content/credit_card_trend.csv")
```

#Reading and viewing the dataset

```
data.head(5)
```

#Displaying columns

```
data.columns
```

#Finding null values

```
data.isnull().sum()
```

#Finding duplicate values

```
data.duplicated().sum()
```

#Deriving statistical information

```
data.describe()
```

#Deriving statistical information of object columns

```
data.describe(include='object')
```

#Merge function

```
data1=data[['Customer_ID','Age','Credit_Score']]
data2=data[['Customer_ID','Monthly_Income','Transaction_Amount']]
result=pd.merge(data1,data2,on='Customer_ID')
result
```

#Pivot tables to find the credit card usage based on age

```
table1=pd.pivot_table(data,values='Transaction_Amount',index='Age',aggfunc=np.sum)
table1
```

#Visualizing the data using plots

```
age=data['Age']
transamt=data['Transaction_Amount']

plt.scatter(age,transamt,color='m')
plt.xlabel('Age')
plt.ylabel('Transaction_Amount')
plt.title('Transactions based on Age')
plt.show()
```

#Pivot tables to find the credit card usage based on Gender

```
table2=pd.pivot_table(data,values='Transaction_Amount',index='Gender',aggfunc=np.sum)
table2
```

```
#Visualizing the data using plots
gender=data['Gender']
transamt=data['Transaction_Amount']

plt.bar(gender,transamt,color='c')
plt.xlabel('Gender')
plt.ylabel('Transaction_Amount')
plt.title('Credit card Usage Patterns based on Gender')
plt.show()

#Pivot tables to find the credit card usage based on Monthly Income
table3=pd.pivot_table(data,values='Transaction_Amount',index='Monthly_Income',aggfunc=np.sum)
table3

#Visualizing the data using plots
x=data['Monthly_Income']
y=data['Transaction_Amount']

plt.bar(x,y,color='r')
plt.xlabel('Monthly_Income (in $)')
plt.ylabel('Transaction_Amount (in $)')
plt.title('Credit card usage based on Monthly Income')
plt.show()

#Pivot tables to find the credit card usage based on Marital Status
table4=pd.pivot_table(data,values='Transaction_Amount',index='Marital_Status',aggfunc=np.sum)
table4

#Visualizing the data using plots
x=data['Marital_Status']
y=data['Transaction_Amount']

plt.bar(x,y,color='#C8A2C8')
plt.xlabel('Marital_Status')
plt.ylabel('Transaction_Amount')
plt.title('Credit card usage based on MaritalStatus')
plt.show()

#Pivot tables to find the credit card usage based on income level
table5=pd.pivot_table(data,values='Transaction_Amount',index='Income_Level',aggfunc=np.sum)
table5

#Visualizing the data using pie chart
x = data['Income_Level']
counts = x.value_counts()  # Get unique categories and their counts

plt.pie(counts, labels=counts.index, autopct='%1.2f%%')  # Use unique labels
plt.legend(title=" Income Levels of customers", loc="lower right")
plt.show()
```

#Pivot tables to find the credit card usage on differrnt categories of expenses
```
table6=pd.pivot_table(data,values='Transaction_Amount',index='Spending_Category',aggfunc=np.su
m)
table6
```

#Visualizing the data using plots
```
x=data['Spending_Category']
y=data['Transaction_Amount']

plt.bar(x,y,color='#AC9362')
plt.xticks(rotation=90)
plt.xlabel('Spending_Category')
plt.ylabel('Transaction_Amount')
plt.title('Spendings on different categories')
plt.show()
```

#Pivot tables to find the credit card usage in different geographical locations
```
table7=pd.pivot_table(data,values='Transaction_Amount',index='City_or_State',aggfunc=np.sum)
table7
```

#Visualizing the data using plots
```
x=data['City_or_State']
y=data['Transaction_Amount']

plt.bar(x,y,color='#87CEEB')
plt.xticks(rotation=90)
plt.xlabel('City_or_State')
plt.ylabel('Transaction_Amount')
plt.title('Spending Patterns in different locations')
plt.show()
```

#Pivot tables to find the monthly income and monthly expenditure
```
table8=pd.pivot_table(data,values='Monthly_Expenditure',index='Monthly_Income',aggfunc=np.sum)
table8

#Visualizing the data using plots
x=data['Monthly_Income']
y=data['Monthly_Expenditure']

plt.bar(x,y,color='g')
plt.xlabel('Monthly_Income')
plt.ylabel('Monthly_Expenditure')
plt.title('Monthly_Expenditure based on Monthly_Income')
plt.show()
```

#Pie chart to depict different types of card holders
```
x=data['Card_Type']

counts = x.value_counts()  # Get unique categories and their counts
```

```
plt.pie(counts, labels=counts.index, autopct='%1.2f%%')  # Use unique labels
plt.legend(title="Card Types of Customers", loc="lower right")
plt.show()
```

#Visualizing the data using plots

```
x=data['Transaction_Amount']
y=data['Card_Type']

plt.bar(y,x,color='y')
plt.xlabel('Transaction_Amount')
plt.ylabel('Card_Type')
plt.title('Transactions based on CardType of customers')
plt.show()
```

#Categorizing Spending Patterns by Age

```
datas = {
    'Age': data['Age'],
    'Spending Category': data['Spending_Category'],
    'Amount': data['Transaction_Amount']
}

# Convert to DataFrame
df = pd.DataFrame(datas)

# Define Age Groups
age_bins = [18, 25, 35, 45, 60, 100]  # Define age range bins
age_labels = ['18-25', '26-35', '36-45', '46-60', '60+']  # Age group labels

# Categorize ages into age groups
df['Age Group'] = pd.cut(df['Age'], bins=age_bins, labels=age_labels, right=False)

# Group data by Age Group and Spending Category
grouped_data = df.groupby(['Age Group', 'Spending Category'])['Amount'].sum().reset_index()

# Pivot table for visualization
pivot_data = grouped_data.pivot(index='Spending Category', columns='Age Group',
values='Amount')

# Display grouped spending patterns
print(grouped_data)

# Bar Plot to visualize spending trends
plt.figure(figsize=(12, 6))
sns.barplot(x='Age Group', y='Amount', hue='Spending Category', data=grouped_data)
plt.title("Spending Categories by Age Group")
plt.xlabel("Age Group")
plt.ylabel("Total Spending")
plt.legend(title="Spending Category")
plt.show()
```

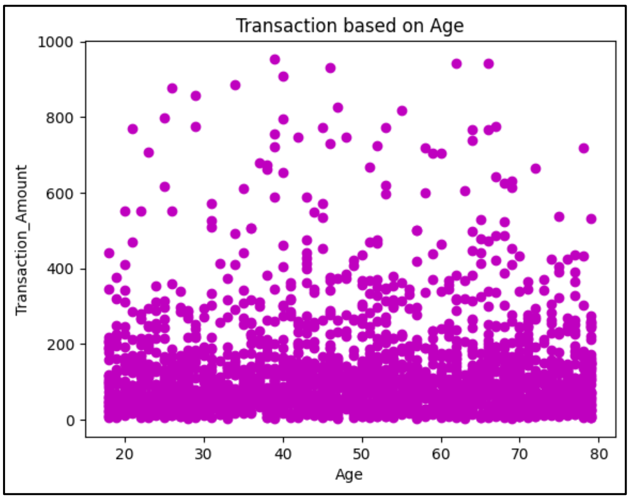
```
#Categorizing Spending Patterns by Marital Status
datas2 = {
    'Marital_Status': data['Marital_Status'],
    'Spending Category': data['Spending_Category'],
    'Amount': data['Transaction_Amount']
}

# Convert to DataFrame
df = pd.DataFrame(datas)

# Group data by Age Group and Spending Category
grouped_data2 = df.groupby(['Marital_Status', 'Spending Category'])['Amount'].sum().reset_index()

# Pivot table for visualization
pivot_data2 = grouped_data2.pivot(index='Spending Category', columns='Marital_Status',
values='Amount')

# Display grouped spending patterns
print(grouped_data2)

# Bar Plot to visualize spending trends
plt.figure(figsize=(12, 6))
sns.barplot(x='Marital_Status', y='Amount', hue='Spending Category', data=grouped_data)
plt.title("Spending Categories by Marital Status")
plt.xlabel("Marital Status ")
plt.ylabel("Total Spending")
plt.legend(title="Spending Category")
plt.show()


#Spendings of card holders by Income Level

datas3 = {
    'Income': data['Income_Level'],
    'Type': data['Card_Type'],
    'Amount': data['Transaction_Amount']
}

# Convert to DataFrame
df = pd.DataFrame(datas3)

# Group data by Age Group and Spending Category
grouped_data3 = df.groupby(['Income', 'Type'])['Amount'].sum().reset_index()

# Pivot table for visualization
pivot_data3 = grouped_data3.pivot(index='Type', columns='Income', values='Amount')

# Display grouped spending patterns
print(grouped_data3)
```

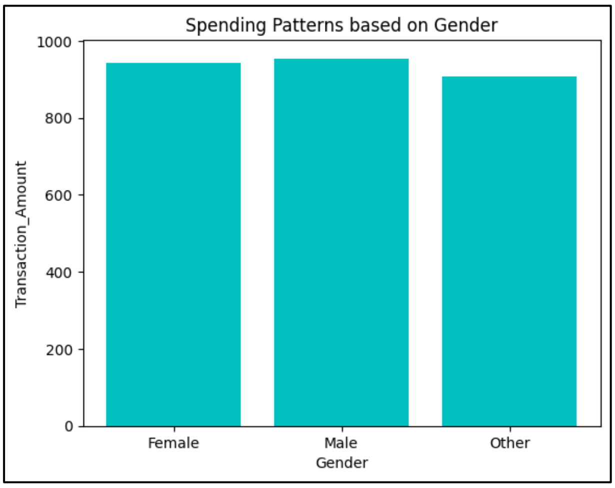
```
# Bar Plot to visualize spending trends
plt.figure(figsize=(12, 6))
plt.figure(figsize=(12, 6))
sns.barplot(x='Income', y='Amount', hue='Type', data=grouped_data3)
plt.title("Spendings of card holders by Income Level")
plt.xlabel("Income level")
plt.ylabel("Amount")
plt.legend(title="Type of Card")
plt.show()
```

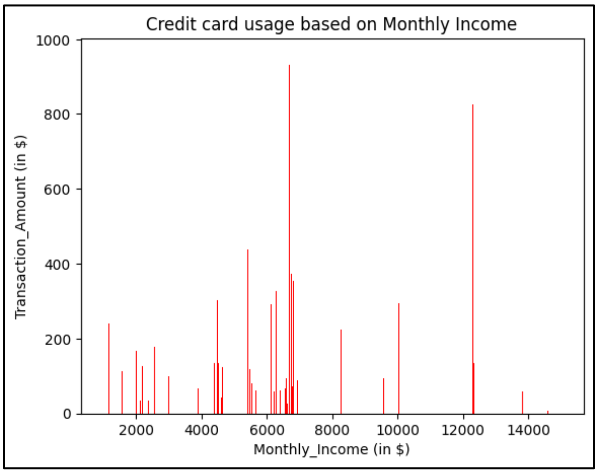
#Categorizing Spending Patterns by Income Level

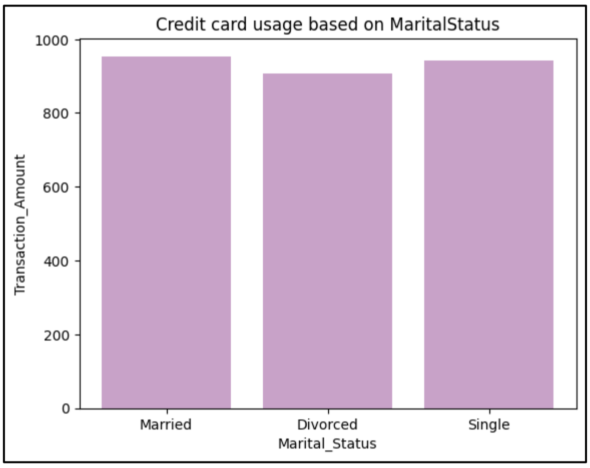
```
datas4 = {
    'Income': data['Income_Level'],
    'Spending Category': data['Spending_Category'],
    'Amount': data['Transaction_Amount']
}

# Convert to DataFrame
df = pd.DataFrame(datas4)

# Group data by Age Group and Spending Category
grouped_data4 = df.groupby(['Income', 'Spending Category'])['Amount'].sum().reset_index()

# Pivot table for visualization
pivot_data4 = grouped_data4.pivot(index='Spending Category', columns='Income', values='Amount')

# Display grouped spending patterns
print(grouped_data4)

# Bar Plot to visualize spending trends
plt.figure(figsize=(12, 6))
sns.barplot(x='Income', y='Amount', hue='Spending Category', data=grouped_data4)
plt.title("Spending Categories by Income Level")
plt.xlabel("Income Level")
plt.ylabel("Total Spending")
plt.legend(title="Spending Category")
plt.show()
```

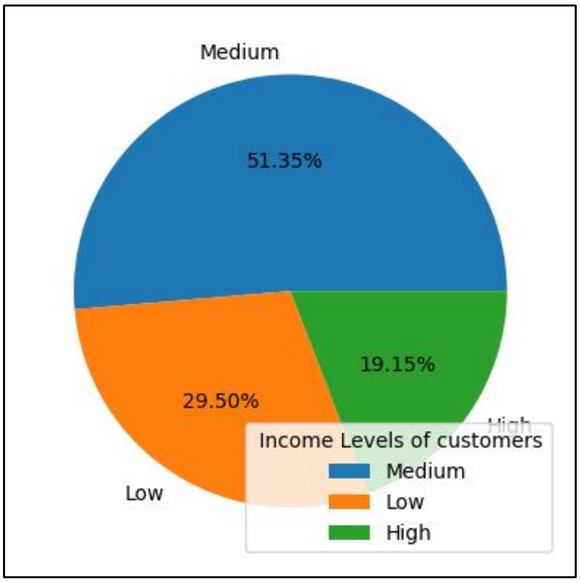
Visualizations:

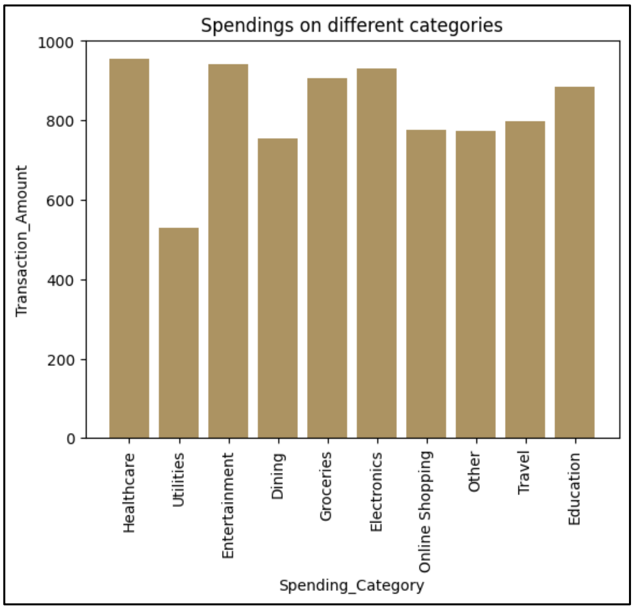


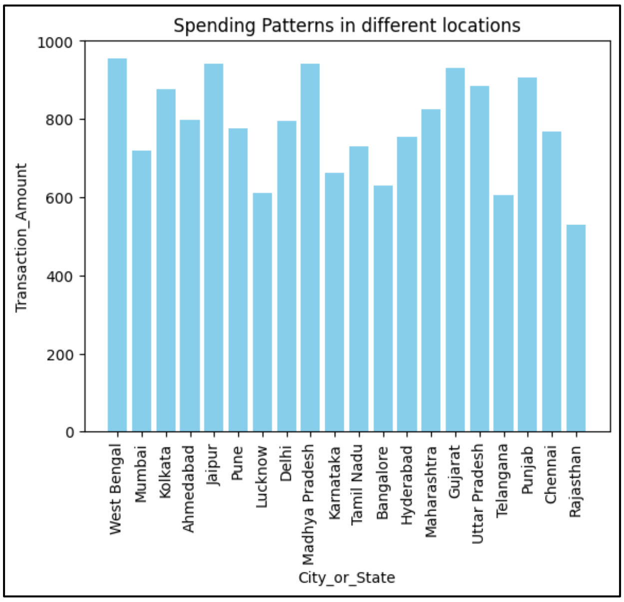


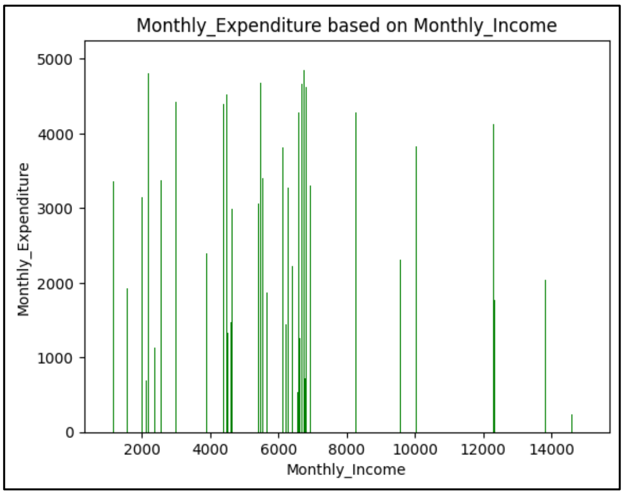


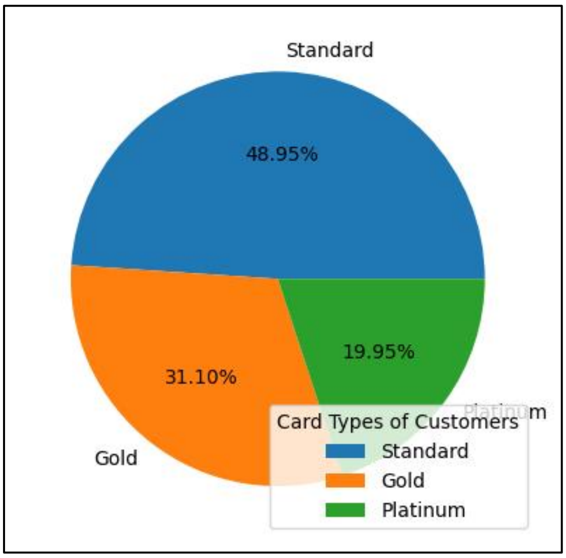


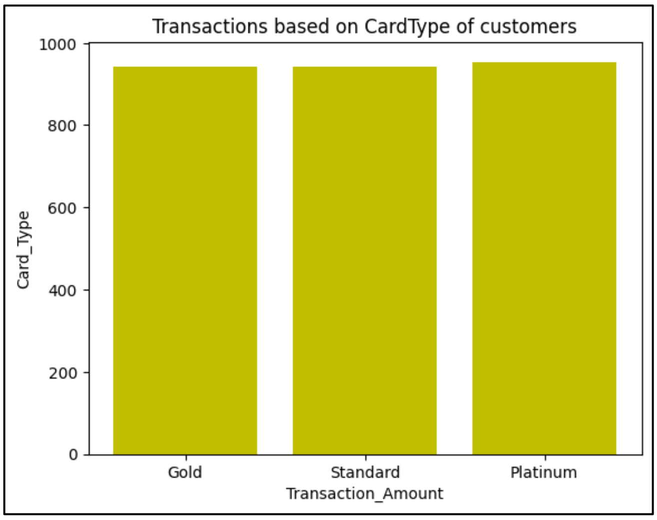


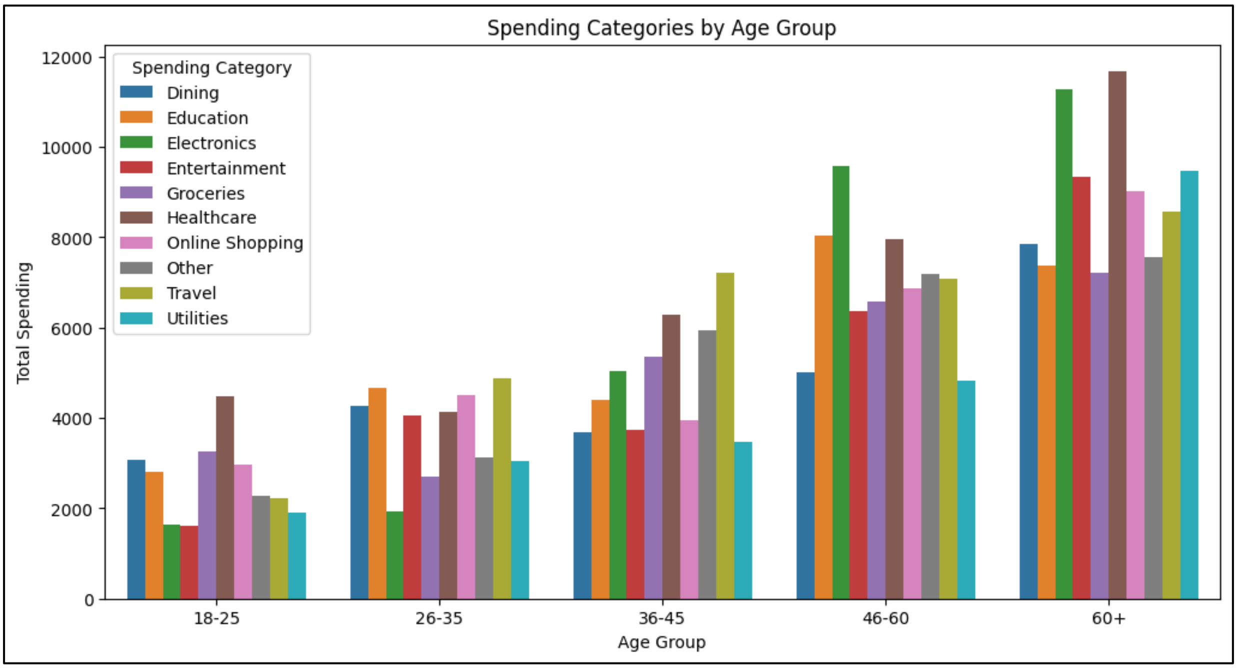


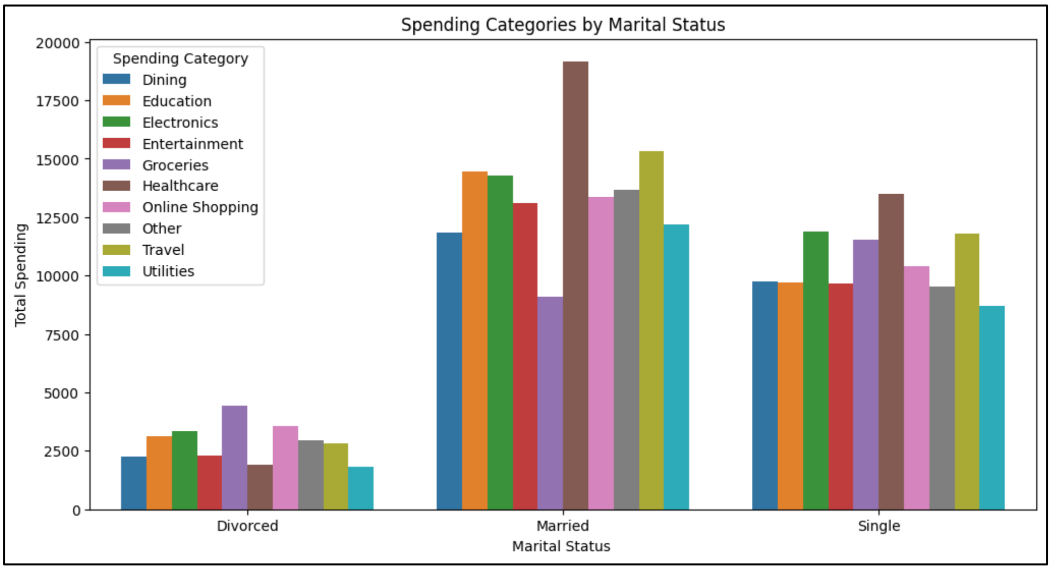

Spending Patterns in different locations
Transaction_Amount
City_or_State
West Bengal
Mumbai
Kolkata
Ahmedabad
Jaipur
Pune
Lucknow
Delhi
Madhya Pradesh
Karnataka
Tamil Nadu
Bangalore
Hyderabad
Maharashtra
Gujarat
Uttar Pradesh
Telangana
Punjab
Chennai
Rajasthan

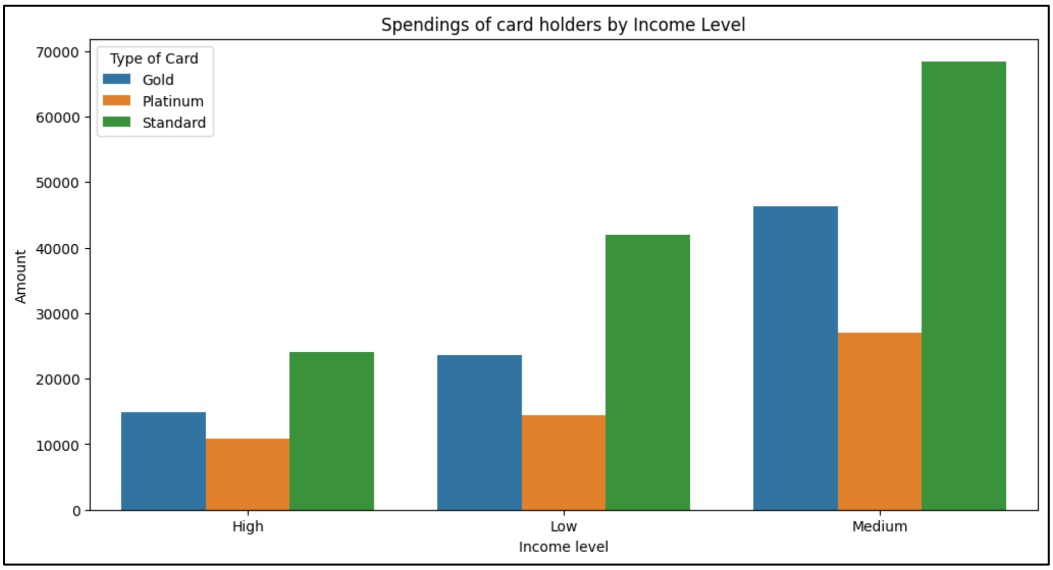

Monthly_Expenditure based on Monthly_Income
Monthly_Expenditure
Monthly_Income

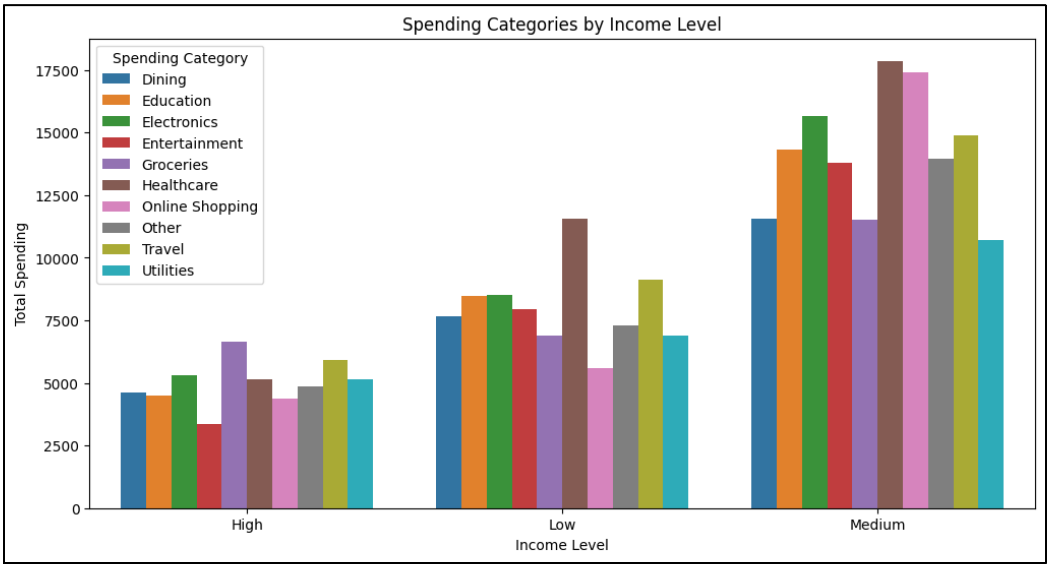

Standard
48.95%
31.10%
19.95%
Gold
Card Types of Customers
Standard
Gold
Platinum


Transactions based on CardType of customers
Card_Type
Gold
Standard
Platinum
Transaction_Amount


Spending Categories by Age Group
Spending Category
Dining
Education
Electronics
Entertainment
Groceries
Healthcare
Online Shopping
Other
Travel
Utilities
Total Spending
18-25
26-35
36-45
46-60
60+
Age Group


Spending Categories by Marital Status
Spending Category
Dining
Education
Electronics
Entertainment
Groceries
Healthcare
Online Shopping
Other
Travel
Utilities
Total Spending
20000
17500
15000
12500
10000
7500
5000
2500
0
Divorced
Married
Single
Marital Status


Spendings of card holders by Income Level
Type of Card
Gold
Platinum
Standard
Amount
70000
60000
50000
40000
30000
20000
10000
0
High
Low
Medium
Income level


Spending Categories by Income Level
Spending Category
Dining
Education
Electronics
Entertainment
Groceries
Healthcare
Online Shopping
Other
Travel
Utilities
Total Spending
17500
15000
12500
10000
7500
5000
2500
0
High
Low
Medium
Income Level

Results:

Based on the analysis of credit card usage trends across different demographics, the project is likely to reveal:

1. **Spending Patterns Vary by Age Group**
   - **Young Adults (18-30):** High spending on entertainment, online shopping, and travel.
   - **Middle-Aged (31-50):** Increased spending on necessities, loans, and family-related expenses.
   - **Older Adults (51+):** Focused on healthcare, utilities based transactions.
2. **Income Level Influences Credit Utilization**
   - **High-Income Earners:** Utilize premium credit cards
   - **Middle-Income Groups:** Balance between discretionary spending and essential expenses.
   - **Low-Income Groups:** Higher dependence on credit for daily needs.
3. **Category-Wise Spending Trends**
   - **Groceries, Utilities, Healthcare:** Universal across demographics.
   - **Luxury Goods & Travel:** More common in higher income.
   - **Education Loans & EMIs:** Higher in younger age group

Conclusion:

The study highlights significant variations in credit card usage trends based on demographic factors, offering key insights for financial institutions, businesses, and policymakers:

- **Banks & Financial Institutions** can use these insights to design personalized credit card offers, adjust credit limits, and optimize risk assessment models.
- **Businesses & Marketers** can target specific consumer segments more effectively based on spending behavior.
- **Policymakers** can implement financial literacy programs to encourage responsible credit usage, especially among high-risk groups.

Overall, this project demonstrates how **data-driven insights can enhance financial decision-making**, improve customer segmentation, and promote **responsible credit usage**.