Problem 2: Credit Card Data for GODIGT Bank

Problem Statement: A bank can generate revenue in a variety of ways, such as charging interest, transaction fees and financial advice. Interest charged on the capital that the bank lends out to customers has historically been the most significant method of revenue generation. The bank earns profits from the difference between the interest rates it pays on deposits and other sources of funds, and the interest rates it charges on the loans it gives out.

GODIGT Bank is a mid-sized private bank that deals in all kinds of banking products, such as savings accounts, current accounts, investment products, etc. among other offerings. The bank also cross-sells asset products to its existing customers through personal loans, auto loans, business loans, etc., and to do so they use various communication methods including cold calling, e-mails, recommendations on the net banking, mobile banking, etc.

GODIGT Bank also has a set of customers who were given credit cards based on risk policy and customer category class but due to huge competition in the credit card market, the bank is observing high attrition in credit card spending. The bank makes money only if customers spend more on credit cards. Given the attrition, the Bank wants to revisit its credit card policy and make sure that the card given to the customer is the right credit card. The bank will make a profit only through the customers that show higher intent towards a recommended credit card. (Higher intent means consumers would want to use the card and hence not be attrite.)

Solution:

Bank Main Objective: **The bank wants customers that show higher intent towards a recommended credit card (Higher intent means consumers would want to use the card and hence not be attrite).**

1.

We checked the shape/size of the dataset:

- **There are 8448 rows and 29 columns.**
- **The dataset has 17 integer, 1 decimal and 11 object type data type columns.**

2.

We did Preliminary analysis & Removed Discrepancies in Dataset

- We deleted "Unnmaed: 28" column which has unnecessary null values.

```
Occupation_at_source          0
cc_limit                      0
Unnamed: 28                8448
dtype: int64
```

```
[7]: bank.drop(['Unnamed: 28'],axis=1,inplace=True) #deleted "Unnmaed: 28" column which has unnnecessary null values.
```

- We checked again null values in columns and found Column Name "Transactor_revolver" has 38 null values.
- By looking at data in Column "Transactor_revolver" we got to know that highest occurrence is value T.
- We now know that value R is less than value T.
- So we filled 38 null values with highest occurring values that is T.

```
T+2_month_activity            0
T+3_month_activity            0
T+6_month_activity            0
T+12_month_activity           0
Transactor_revolver          38
 avg_spends_13m               0
Occupation_at_source          0
cc_limit                      0
dtype: int64
```

```
[10]: bank['Transactor_revolver'].unique()

      # Checking unique values in "Transactor_revolver" Column.
```

```
[10]: array(['T', 'R', nan], dtype=object)
```

```
[11]: bank.groupby('Transactor_revolver')['Transactor_revolver'].count()

      # Finding which Value comes in higher numbers.
      # We now know that value R is less than value T.
      # So we will fill 38 null values with highest occuring values that is T.
```

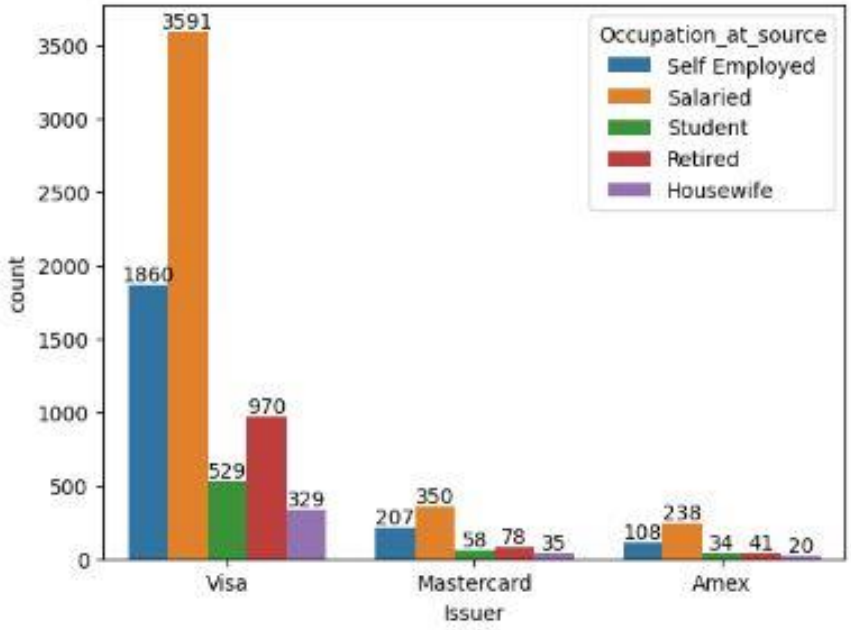
```
[11]: Transactor_revolver
      R    1295
      T    7115
      Name: Transactor_revolver, dtype: int64
```

```
[12]: bank['Transactor_revolver'].fillna("T", inplace=True)

      # Filled 38 null values with highest occuring values that is T.
```

Now that our dataset is clean and without discrepancies, we can now create graphs to share insights:

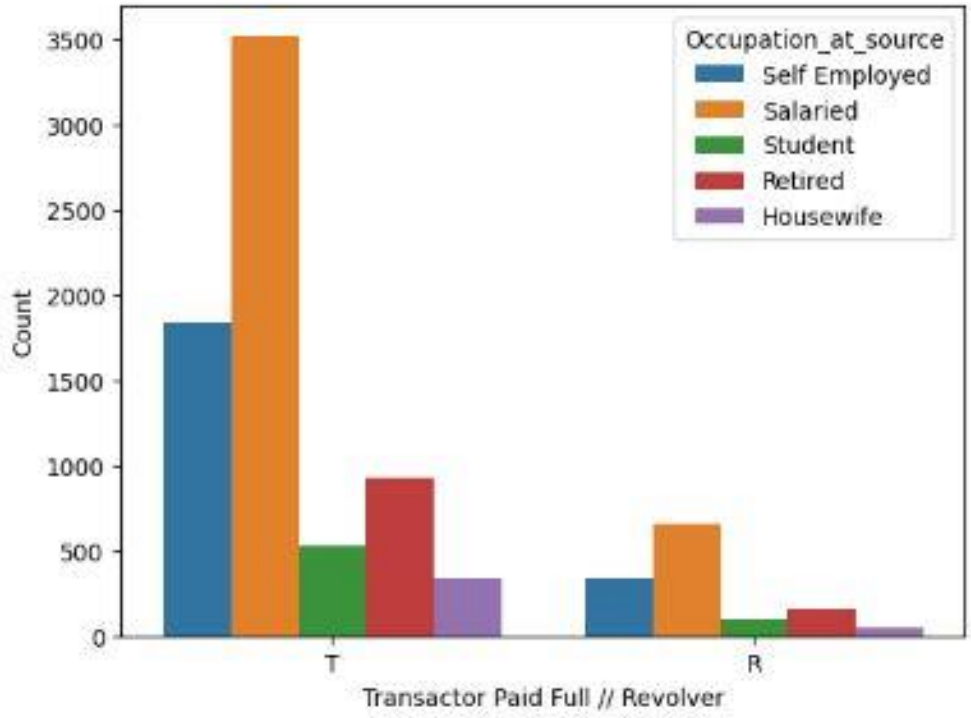
1. **Below chart shows Customers likely to buy Visa card because it is best among Salaried and Self-employed (Also most occupation likes to have VISA card)**

```
ax=sns.countplot(data=bank,x="Issuer",hue="Occupation_at_source")

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


# Visa card is best among Salaried and Self employed ( I mean most occupation likes VISA card)
```



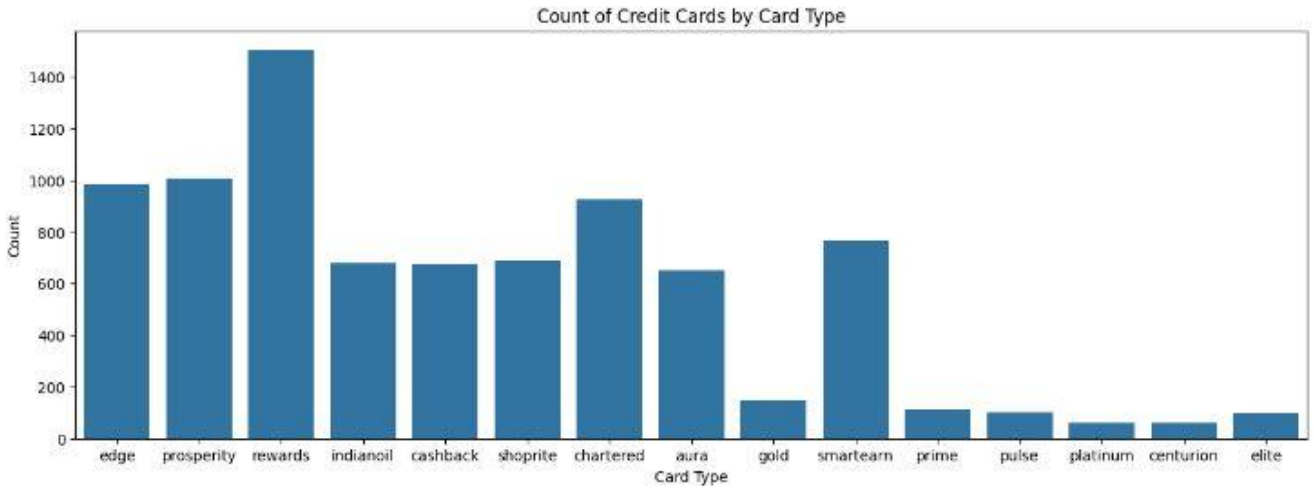
2. **Below chart shows Salaried Customer pays full amount every month on credit card spendings.**

```
ax=sns.countplot(data=bank,x="Transactor_revolver",hue="Occupation_at_source")
plt.xlabel("Transactor Paid Full // Revolver")
plt.ylabel("Count")
plt.show()
for bars in ax.containers:
    ax.bar_label(bars)


# Salaried Customer pays full amount every month on credit card spendings.
```



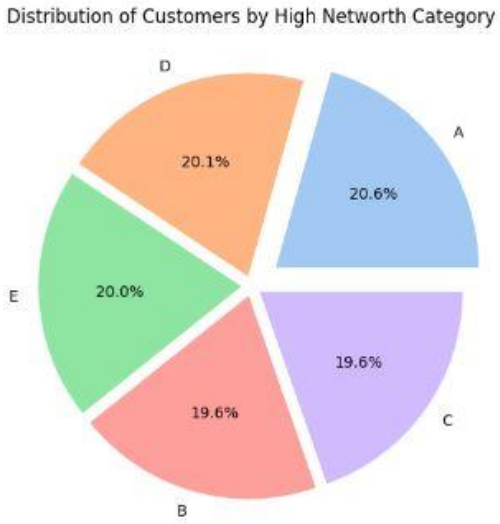
3. **Below chart shows that Best credit card types are rewards, prosperity and edge.**

```
# Example: Bar plot for count of credit cards by card type
plt.figure(figsize=(15, 5))
sns.countplot(data=bank, x='card_type')
plt.xlabel('Card Type')
plt.ylabel('Count')
plt.title('Count of Credit Cards by Card Type')
plt.show()

# Best credit card types are rewards, prosperity and edge.
```

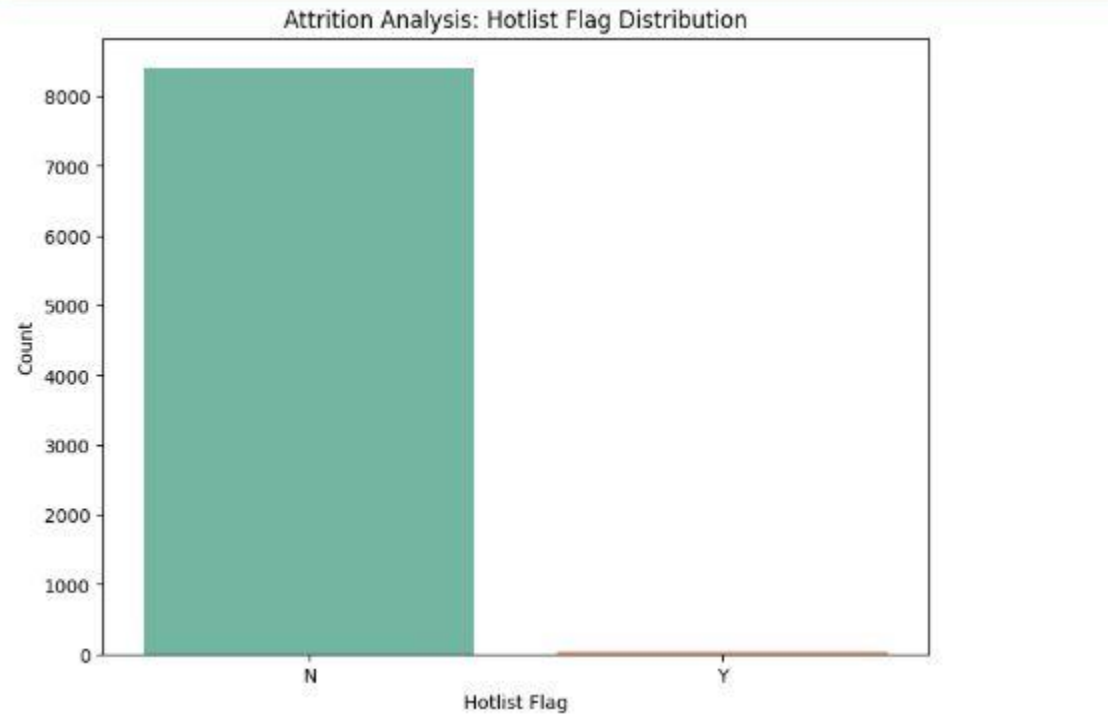


4. **Customer with high net worth are "A" category are more than others**

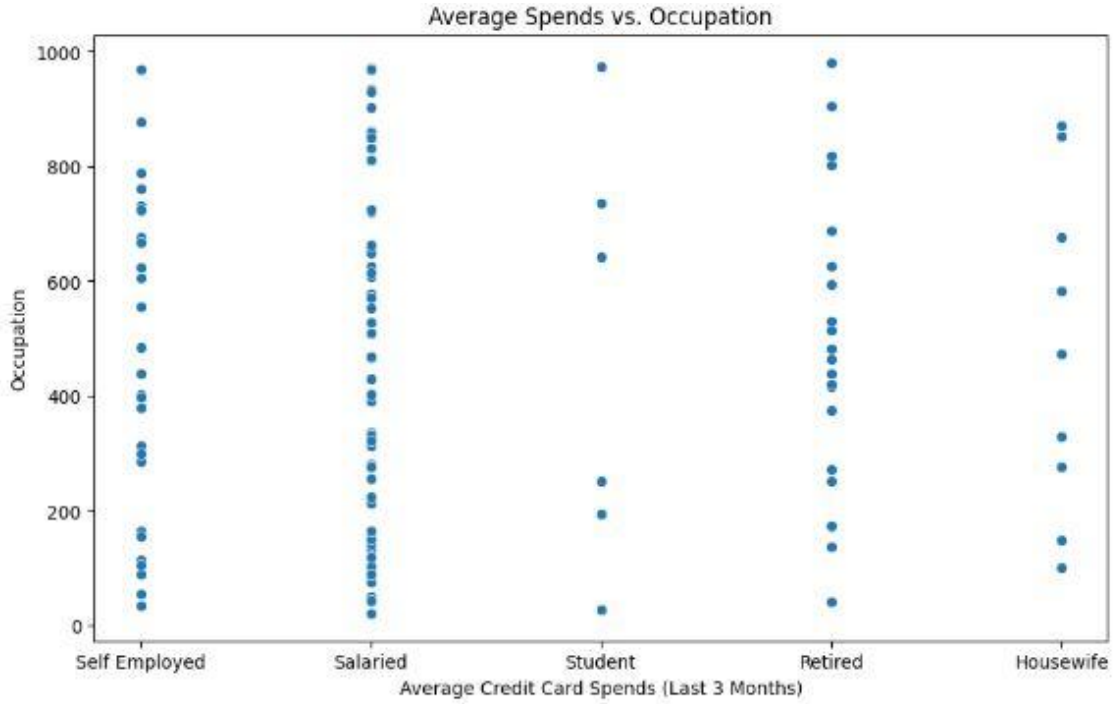
```
# Example: Pie chart for high_networth distribution
plt.figure(figsize=(8, 6))
bank['high_networth'].value_counts().plot(kind='pie', autopct='%1.1f%%', colors=sns.color_palette('pastel'), explode=(0.15, 0.05, 0.05, 0.05, 0.05))
plt.title('Distribution of Customers by High Networth Category')
plt.ylabel('')
plt.show()

# Customer with high networth A are more than others
```



5. **Below chart shows Very less Customers that has their card hotlisted.**



6. **Below chart we can see Salaried customers have spend more in last 3 months**



Conclusion: Salaried and Self-Employed Customers are more likely to use VISA Credit Card and they will not attrite. The best card type works here is Rewards, Prosperity, and Edge.

----------*----------*----------

Project By: Amit Bahukhandi (MSCIT)