In [75]:
```
import pandas as pd
```

In [76]:
```
df=pd.read_csv("C:/Users/NITISH/OneDrive/Documents/Desktop/Bank Data/TOP BANKS.csv"
```

In [77]:
```
print(df.head())
```

```
   Gender       City Age group          Occupation                BANK1  \
0    Male     Mumbai     31-40        Own Business            HDFC Bank
1  Female  Bangalore     41-50  Salaried □ Private           ICICI Bank
2    Male  Ahmedabad     41-50        Own Business  Kotak Mahindra Bank
3    Male  Bangalore     26-30  Salaried □ Private            HDFC Bank
4    Male  Ahmedabad     51-60        Own Business  Kotak Mahindra Bank

        BANK2                BANK3  Kotak Mahindra Bank  HDFC Bank ICICI Bank  \
0  ICICI Bank             Yes Bank  Kotak Mahindra Bank          -          -
1   HDFC Bank  Kotak Mahindra Bank  Kotak Mahindra Bank          -          -
2   HDFC Bank            Axis Bank                    -          -          -
3  ICICI Bank  Kotak Mahindra Bank                    -  HDFC Bank          -
4   HDFC Bank  State Bank Of India  Kotak Mahindra Bank          -          -

  Axis Bank  IndusInd Bank RBL Bank            Other 1 Other 2  \
0         -              -        -                  -       -
1         -              -        -                  -       -
2         -  IndusInd Bank        -  South Indian Bank       -
3         -              -        -      Bank Of India       -
4         -              -        -                  -       -

           Salary Acc          Primary Acc
0            Yes Bank             Yes Bank
1       Bank Of Baroda            HDFC Bank
2            Axis Bank  Kotak Mahindra Bank
3    None of the above           ICICI Bank
4  Kotak Mahindra Bank  Kotak Mahindra Bank
```

In [78]:
```
Age_group = ['18-25', '26-30', '31-40', '41-50', '51-60']
```

In [79]:
```
df['Age group'] = df['Age group'].replace({
    '31-40': '31-40',
    '41-50': '41-50',
    '51-60': '51-60',
    '26-35': '26-30',
    '18-25': '18-25'
})
```

In [80]:
```
gender_distribution = df['Gender'].value_counts()
```

In [81]:
```
city_distribution = df['City'].value_counts()
age_group_distribution = df['Age group'].value_counts()
```

In [82]:
```
all_banks = df.iloc[:, 4:15]  # Selecting columns corresponding to banks
bank_counts = all_banks.apply(pd.Series.value_counts).sum(axis=1)
most_preferred_banks = bank_counts.sort_values(ascending=False).head()
```

In [83]:
```
occupation_banks = df.groupby('Occupation').agg({
    'BANK1': lambda x: x.value_counts().index[0],
    'BANK2': lambda x: x.value_counts().index[0],
    'BANK3': lambda x: x.value_counts().index[0]
})
```

In [84]:
```
primary_salary_accounts = df['Primary Acc'].value_counts()
```

In [85]:
```
print("Distribution of respondents by gender:\n", gender_distribution)
print("\nDistribution of respondents by city:\n", city_distribution)
print("\nDistribution of respondents by age group:\n", age_group_distribution)
print("\nMost preferred banks among respondents:\n", most_preferred_banks)
print("\nOccupation-wise preferred banks:\n", occupation_banks)
print("\nDistribution of primary salary accounts among banks:\n", primary_salary_ac
```

```
Distribution of respondents by gender:
 Gender
Male      255
Female    200
Name: count, dtype: int64

Distribution of respondents by city:
 City
Mumbai         77
Delhi          61
Kolkata        57
Ahmedabad      54
Lucknow        51
Bangalore      46
Chennai        42
Indore         27
Kochi          24
Bhubaneswar    16
Name: count, dtype: int64

Distribution of respondents by age group:
 Age group
18-25    134
31-40    121
26-30    116
41-50     65
51-60     19
Name: count, dtype: int64

Most preferred banks among respondents:
 -                      2895.0
HDFC Bank               447.0
Kotak Mahindra Bank     370.0
ICICI Bank              334.0
State Bank Of India     229.0
dtype: float64

Occupation-wise preferred banks:
                                      BANK1                BANK2  \
Occupation
Housewife                         HDFC Bank  State Bank Of India
Own Business                      HDFC Bank            HDFC Bank
Salaried - Government  Kotak Mahindra Bank       Bank Of Baroda
Salaried □ Private                HDFC Bank            HDFC Bank
Student                          ICICI Bank  State Bank Of India

                                      BANK3
Occupation
Housewife                        ICICI Bank
Own Business           State Bank Of India
Salaried - Government           ICICI Bank
Salaried □ Private              ICICI Bank
Student                          HDFC Bank

Distribution of primary salary accounts among banks:
 Primary Acc
HDFC Bank                    121
Kotak Mahindra Bank          120
ICICI Bank                    90
Axis Bank                     48
State Bank Of India           17
IndusInd Bank                 16
Punjab National Bank           5
RBL Bank                       5
```

```
Union Bank of India            4
-                              4
Canara Bank                    4
Yes Bank                       4
Bank Of Baroda                 3
Bank Of India                  1
Karnataka Bank                 1
Indian Bank                    1
South Indian Bank              1
Equitas Small Finance Bank     1
IDBI Bank                      1
DBS Bank                       1
Paytm Payments Bank            1
IDFC Bank                      1
HSBC India                     1
Kerala Bank                    1
Tamilnad Mercantile Bank       1
Federal Bank                   1
Bank Of Maharashtra            1
Name: count, dtype: int64
```

In [86]:
```
categorical_vars = ['Gender', 'City', 'Age group', 'Occupation', 'Primary Acc','Sal
freq_distribution = {}
for col in categorical_vars:
    freq_distribution[col] = df[col].value_counts()
```

In [87]:
```
numerical_vars = df.select_dtypes(include=['int64', 'float64']).columns
correlation_matrix = df[numerical_vars].corr()
```

In [88]:
```
from scipy.stats import chi2_contingency
```

In [89]:
```
summary_stats = df.describe()


observed = pd.crosstab(df['Gender'], df['Age group'])
chi2, p, _, _ = chi2_contingency(observed)
if p < 0.05:
    hypothesis_result = "Reject null hypothesis: There is a significant difference.
else:
    hypothesis_result = "Fail to reject null hypothesis: There is no significant di
```

In [90]:
```
print("Summary statistics:\n", summary_stats)
print("\nFrequency distribution for categorical variables:\n")
for col, freq_dist in freq_distribution.items():
    print(col + ":\n", freq_dist)
print("\nCorrelation matrix:\n", correlation_matrix)
print("\nHypothesis testing result:\n", hypothesis_result)
```

```
Summary statistics:
        Gender    City Age group          Occupation      BANK1      BANK2  \
count      455     455       455                 455        455        453
unique       2      10         5                   5         23         32
top       Male  Mumbai     18-25  Salaried □ Private  HDFC Bank  HDFC Bank
freq       255      77       134                 300        139         77

                      BANK3 Kotak Mahindra Bank HDFC Bank ICICI Bank  \
count                   449                 455       455        455
unique                   35                   2         2          2
top     State Bank Of India                   -         -          -
freq                     62                 277       279        331

       Axis Bank IndusInd Bank RBL Bank Other 1 Other 2         Salary Acc  \
count        455           455      455     455     455                455
unique         2             2        2      24       7                 22
top            -             -        -       -       -  None of the above
freq         375           424      446     320     443                134

       Primary Acc
count          455
unique          27
top      HDFC Bank
freq           121

Frequency distribution for categorical variables:

Gender:
 Gender
Male      255
Female    200
Name: count, dtype: int64
City:
 City
Mumbai         77
Delhi          61
Kolkata        57
Ahmedabad      54
Lucknow        51
Bangalore      46
Chennai        42
Indore         27
Kochi          24
Bhubaneswar    16
Name: count, dtype: int64
Age group:
 Age group
18-25    134
31-40    121
26-30    116
41-50     65
51-60     19
Name: count, dtype: int64
Occupation:
 Occupation
Salaried □ Private       300
Own Business             73
Housewife                67
Student                  14
Salaried - Government     1
Name: count, dtype: int64
Primary Acc:
 Primary Acc
HDFC Bank                    121
```

```
Kotak Mahindra Bank          120
ICICI Bank                    90
Axis Bank                     48
State Bank Of India           17
IndusInd Bank                 16
Punjab National Bank           5
RBL Bank                       5
Union Bank of India            4
-                              4
Canara Bank                    4
Yes Bank                       4
Bank Of Baroda                 3
Bank Of India                  1
Karnataka Bank                 1
Indian Bank                    1
South Indian Bank              1
Equitas Small Finance Bank     1
IDBI Bank                      1
DBS Bank                       1
Paytm Payments Bank            1
IDFC Bank                      1
HSBC India                     1
Kerala Bank                    1
Tamilnad Mercantile Bank       1
Federal Bank                   1
Bank Of Maharashtra            1
Name: count, dtype: int64
Salary Acc:
 Salary Acc
None of the above            134
HDFC Bank                     91
Kotak Mahindra Bank           73
ICICI Bank                    71
Axis Bank                     38
IndusInd Bank                 13
State Bank Of India            8
Bank Of Baroda                 6
Punjab National Bank           4
Canara Bank                    2
Union Bank of India            2
Yes Bank                       2
Bank Of India                  2
HSBC India                     1
Bank Of Maharashtra            1
Federal Bank                   1
DBS Bank                       1
IDFC Bank                      1
Equitas Small Finance Bank     1
IDBI Bank                      1
South Indian Bank              1
Indian Bank                    1
Name: count, dtype: int64

Correlation matrix:
 Empty DataFrame
Columns: []
Index: []

Hypothesis testing result:
 Fail to reject null hypothesis: There is no significant difference.
```

In [91]:
```
primary_salary_bank1 = pd.crosstab(df['Primary Acc'], df['Salary Acc'], margins=Tru
primary_salary_bank2 = pd.crosstab(df['Primary Acc'], df['Salary Acc'], margins=Tru
primary_salary_bank3 = pd.crosstab(df['Primary Acc'], df['Salary Acc'], margins=Tru
```

In [92]:
```
print("Cross-tabulation of Primary Acc and Salary Acc with Bank1:\n", primary_salar
print("\nCross-tabulation of Primary Acc and Salary Acc with Bank2:\n", primary_sal
print("\nCross-tabulation of Primary Acc and Salary Acc with Bank3:\n", primary_sal
```

Cross-tabulation of Primary Acc and Salary Acc with Bank1:

Primary Acc \ Salary Acc	Axis Bank	Bank Of Baroda	Bank Of India
-	0	0	0
Axis Bank	29	0	0
Bank Of Baroda	0	2	0
Bank Of India	0	0	1
Bank Of Maharashtra	0	0	0
Canara Bank	0	0	0
DBS Bank	0	0	0
Equitas Small Finance Bank	0	0	0
Federal Bank	0	0	0
HDFC Bank	3	2	1
HSBC India	0	0	0
ICICI Bank	0	0	0
IDBI Bank	1	0	0
IDFC Bank	0	0	0
Indian Bank	0	0	0
IndusInd Bank	0	0	0
Karnataka Bank	0	0	0
Kerala Bank	1	0	0
Kotak Mahindra Bank	1	2	0
Paytm Payments Bank	0	0	0
Punjab National Bank	0	0	0
RBL Bank	0	0	0
South Indian Bank	0	0	0
State Bank Of India	3	0	0
Tamilnad Mercantile Bank	0	0	0
Union Bank of India	0	0	0
Yes Bank	0	0	0
All	38	6	2

Primary Acc \ Salary Acc	Bank Of Maharashtra	Canara Bank	DBS Bank
-	0	0	0
Axis Bank	0	0	0
Bank Of Baroda	0	0	0
Bank Of India	0	0	0
Bank Of Maharashtra	1	0	0
Canara Bank	0	2	0
DBS Bank	0	0	0
Equitas Small Finance Bank	0	0	0
Federal Bank	0	0	0
HDFC Bank	0	0	0
HSBC India	0	0	0
ICICI Bank	0	0	0
IDBI Bank	0	0	0
IDFC Bank	0	0	0
Indian Bank	0	0	0
IndusInd Bank	0	0	0
Karnataka Bank	0	0	0
Kerala Bank	0	0	0
Kotak Mahindra Bank	0	0	1
Paytm Payments Bank	0	0	0
Punjab National Bank	0	0	0
RBL Bank	0	0	0
South Indian Bank	0	0	0
State Bank Of India	0	0	0
Tamilnad Mercantile Bank	0	0	0
Union Bank of India	0	0	0
Yes Bank	0	0	0
All	1	2	1

Primary Acc \ Salary Acc	Equitas Small Finance Bank	Federal Bank

Primary Acc		
-	0	0
Axis Bank	0	0
Bank Of Baroda	0	0
Bank Of India	0	0
Bank Of Maharashtra	0	0
Canara Bank	0	0
DBS Bank	0	0
Equitas Small Finance Bank	1	0
Federal Bank	0	1
HDFC Bank	0	0
HSBC India	0	0
ICICI Bank	0	0
IDBI Bank	0	0
IDFC Bank	0	0
Indian Bank	0	0
IndusInd Bank	0	0
Karnataka Bank	0	0
Kerala Bank	0	0
Kotak Mahindra Bank	0	0
Paytm Payments Bank	0	0
Punjab National Bank	0	0
RBL Bank	0	0
South Indian Bank	0	0
State Bank Of India	0	0
Tamilnad Mercantile Bank	0	0
Union Bank of India	0	0
Yes Bank	0	0
All	1	1

Salary Acc	HDFC Bank	HSBC India	...	Indian Bank	\
Primary Acc			...		
-	1	0	...	0	
Axis Bank	0	0	...	0	
Bank Of Baroda	0	0	...	0	
Bank Of India	0	0	...	0	
Bank Of Maharashtra	0	0	...	0	
Canara Bank	1	0	...	0	
DBS Bank	0	0	...	0	
Equitas Small Finance Bank	0	0	...	0	
Federal Bank	0	0	...	0	
HDFC Bank	78	0	...	0	
HSBC India	0	1	...	0	
ICICI Bank	3	0	...	0	
IDBI Bank	0	0	...	0	
IDFC Bank	1	0	...	0	
Indian Bank	0	0	...	0	
IndusInd Bank	0	0	...	1	
Karnataka Bank	0	0	...	0	
Kerala Bank	0	0	...	0	
Kotak Mahindra Bank	1	0	...	0	
Paytm Payments Bank	0	0	...	0	
Punjab National Bank	0	0	...	0	
RBL Bank	2	0	...	0	
South Indian Bank	0	0	...	0	
State Bank Of India	3	0	...	0	
Tamilnad Mercantile Bank	0	0	...	0	
Union Bank of India	1	0	...	0	
Yes Bank	0	0	...	0	
All	91	1	...	1	

Salary Acc	IndusInd Bank	Kotak Mahindra Bank	\
Primary Acc			
-	0	0	

Axis Bank	0
Bank Of Baroda	0
Bank Of India	0
Bank Of Maharashtra	0
Canara Bank	0
DBS Bank	0
Equitas Small Finance Bank	0
Federal Bank	0
HDFC Bank	1
HSBC India	0
ICICI Bank	1
IDBI Bank	0
IDFC Bank	0
Indian Bank	0
IndusInd Bank	11
Karnataka Bank	0
Kerala Bank	0
Kotak Mahindra Bank	0
Paytm Payments Bank	0
Punjab National Bank	0
RBL Bank	0
South Indian Bank	0
State Bank Of India	0
Tamilnad Mercantile Bank	0
Union Bank of India	0
Yes Bank	0
All	13

Salary Acc	None of the above	Punjab National Bank
Primary Acc		
-	2	0
Axis Bank	18	0
Bank Of Baroda	1	0
Bank Of India	0	0
Bank Of Maharashtra	0	0
Canara Bank	1	0
DBS Bank	1	0
Equitas Small Finance Bank	0	0
Federal Bank	0	0
HDFC Bank	31	0
HSBC India	0	0
ICICI Bank	19	0
IDBI Bank	0	0
IDFC Bank	0	0
Indian Bank	1	0
IndusInd Bank	3	0
Karnataka Bank	1	0
Kerala Bank	0	0
Kotak Mahindra Bank	43	1
Paytm Payments Bank	1	0
Punjab National Bank	0	3
RBL Bank	3	0
South Indian Bank	0	0
State Bank Of India	5	0
Tamilnad Mercantile Bank	1	0
Union Bank of India	2	0
Yes Bank	1	0
All	134	4

Salary Acc	South Indian Bank	State Bank Of India
Primary Acc		
-	0	0
Axis Bank	0	0
Bank Of Baroda	0	0

```
Bank Of India                          0                   0
Bank Of Maharashtra                    0                   0
Canara Bank                            0                   0
DBS Bank                               0                   0
Equitas Small Finance Bank             0                   0
Federal Bank                           0                   0
HDFC Bank                              0                   1
HSBC India                             0                   0
ICICI Bank                             0                   1
IDBI Bank                              0                   0
IDFC Bank                              0                   0
Indian Bank                            0                   0
IndusInd Bank                          0                   0
Karnataka Bank                         0                   0
Kerala Bank                            0                   0
Kotak Mahindra Bank                    0                   1
Paytm Payments Bank                    0                   0
Punjab National Bank                   0                   0
RBL Bank                               0                   0
South Indian Bank                      1                   0
State Bank Of India                    0                   5
Tamilnad Mercantile Bank               0                   0
Union Bank of India                    0                   0
Yes Bank                               0                   0
All                                    1                   8

Salary Acc                  Union Bank of India  Yes Bank  All
Primary Acc
-                                         0         0    4
Axis Bank                                 0         0   48
Bank Of Baroda                            0         0    3
Bank Of India                             0         0    1
Bank Of Maharashtra                       0         0    1
Canara Bank                               0         0    4
DBS Bank                                  0         0    1
Equitas Small Finance Bank                0         0    1
Federal Bank                              0         0    1
HDFC Bank                                 0         0  121
HSBC India                                0         0    1
ICICI Bank                                1         0   90
IDBI Bank                                 0         0    1
IDFC Bank                                 0         0    1
Indian Bank                               0         0    1
IndusInd Bank                             0         0   16
Karnataka Bank                            0         0    1
Kerala Bank                               0         0    1
Kotak Mahindra Bank                       0         0  120
Paytm Payments Bank                       0         0    1
Punjab National Bank                      0         0    5
RBL Bank                                  0         0    5
South Indian Bank                         0         0    1
State Bank Of India                       0         0   17
Tamilnad Mercantile Bank                  0         0    1
Union Bank of India                       1         0    4
Yes Bank                                  0         2    4
All                                       2         2  455

[28 rows x 23 columns]

Cross-tabulation of Primary Acc and Salary Acc with Bank2:
 Salary Acc                  Axis Bank  Bank Of Baroda  Bank Of India  \
Primary Acc
-                                   0               0              0
Axis Bank                          29               0              0
```

Bank Of Baroda	0	2	0
Bank Of India	0	0	1
Bank Of Maharashtra	0	0	0
Canara Bank	0	0	0
DBS Bank	0	0	0
Equitas Small Finance Bank	0	0	0
Federal Bank	0	0	0
HDFC Bank	3	2	1
HSBC India	0	0	0
ICICI Bank	0	0	0
IDBI Bank	1	0	0
IDFC Bank	0	0	0
Indian Bank	0	0	0
IndusInd Bank	0	0	0
Karnataka Bank	0	0	0
Kerala Bank	1	0	0
Kotak Mahindra Bank	1	2	0
Paytm Payments Bank	0	0	0
Punjab National Bank	0	0	0
RBL Bank	0	0	0
South Indian Bank	0	0	0
State Bank Of India	3	0	0
Tamilnad Mercantile Bank	0	0	0
Union Bank of India	0	0	0
Yes Bank	0	0	0
All	38	6	2

Salary Acc Primary Acc	Bank Of Maharashtra	Canara Bank	DBS Bank	\
-	0	0	0	
Axis Bank	0	0	0	
Bank Of Baroda	0	0	0	
Bank Of India	0	0	0	
Bank Of Maharashtra	1	0	0	
Canara Bank	0	2	0	
DBS Bank	0	0	0	
Equitas Small Finance Bank	0	0	0	
Federal Bank	0	0	0	
HDFC Bank	0	0	0	
HSBC India	0	0	0	
ICICI Bank	0	0	0	
IDBI Bank	0	0	0	
IDFC Bank	0	0	0	
Indian Bank	0	0	0	
IndusInd Bank	0	0	0	
Karnataka Bank	0	0	0	
Kerala Bank	0	0	0	
Kotak Mahindra Bank	0	0	1	
Paytm Payments Bank	0	0	0	
Punjab National Bank	0	0	0	
RBL Bank	0	0	0	
South Indian Bank	0	0	0	
State Bank Of India	0	0	0	
Tamilnad Mercantile Bank	0	0	0	
Union Bank of India	0	0	0	
Yes Bank	0	0	0	
All	1	2	1	

Salary Acc Primary Acc	Equitas Small Finance Bank	Federal Bank	\
-	0	0	
Axis Bank	0	0	
Bank Of Baroda	0	0	
Bank Of India	0	0	

Bank Of Maharashtra	0	0
Canara Bank	0	0
DBS Bank	0	0
Equitas Small Finance Bank	1	0
Federal Bank	0	1
HDFC Bank	0	0
HSBC India	0	0
ICICI Bank	0	0
IDBI Bank	0	0
IDFC Bank	0	0
Indian Bank	0	0
IndusInd Bank	0	0
Karnataka Bank	0	0
Kerala Bank	0	0
Kotak Mahindra Bank	0	0
Paytm Payments Bank	0	0
Punjab National Bank	0	0
RBL Bank	0	0
South Indian Bank	0	0
State Bank Of India	0	0
Tamilnad Mercantile Bank	0	0
Union Bank of India	0	0
Yes Bank	0	0
All	1	1

Salary Acc	HDFC Bank	HSBC India	...	Indian Bank	\
Primary Acc			...		
-	1	0	...	0	
Axis Bank	0	0	...	0	
Bank Of Baroda	0	0	...	0	
Bank Of India	0	0	...	0	
Bank Of Maharashtra	0	0	...	0	
Canara Bank	1	0	...	0	
DBS Bank	0	0	...	0	
Equitas Small Finance Bank	0	0	...	0	
Federal Bank	0	0	...	0	
HDFC Bank	78	0	...	0	
HSBC India	0	1	...	0	
ICICI Bank	3	0	...	0	
IDBI Bank	0	0	...	0	
IDFC Bank	1	0	...	0	
Indian Bank	0	0	...	0	
IndusInd Bank	0	0	...	1	
Karnataka Bank	0	0	...	0	
Kerala Bank	0	0	...	0	
Kotak Mahindra Bank	1	0	...	0	
Paytm Payments Bank	0	0	...	0	
Punjab National Bank	0	0	...	0	
RBL Bank	2	0	...	0	
South Indian Bank	0	0	...	0	
State Bank Of India	3	0	...	0	
Tamilnad Mercantile Bank	0	0	...	0	
Union Bank of India	1	0	...	0	
Yes Bank	0	0	...	0	
All	91	1	...	1	

Salary Acc	IndusInd Bank	Kotak Mahindra Bank	\
Primary Acc			
-	0	0	
Axis Bank	0	0	
Bank Of Baroda	0	0	
Bank Of India	0	0	
Bank Of Maharashtra	0	0	
Canara Bank	0	0	

DBS Bank	0	0
Equitas Small Finance Bank	0	0
Federal Bank	0	0
HDFC Bank	1	2
HSBC India	0	0
ICICI Bank	1	0
IDBI Bank	0	0
IDFC Bank	0	0
Indian Bank	0	0
IndusInd Bank	11	0
Karnataka Bank	0	0
Kerala Bank	0	0
Kotak Mahindra Bank	0	68
Paytm Payments Bank	0	0
Punjab National Bank	0	2
RBL Bank	0	0
South Indian Bank	0	0
State Bank Of India	0	0
Tamilnad Mercantile Bank	0	0
Union Bank of India	0	0
Yes Bank	0	1
All	13	73

Salary Acc	None of the above	Punjab National Bank
Primary Acc		
-	2	0
Axis Bank	18	0
Bank Of Baroda	1	0
Bank Of India	0	0
Bank Of Maharashtra	0	0
Canara Bank	1	0
DBS Bank	1	0
Equitas Small Finance Bank	0	0
Federal Bank	0	0
HDFC Bank	31	0
HSBC India	0	0
ICICI Bank	19	0
IDBI Bank	0	0
IDFC Bank	0	0
Indian Bank	1	0
IndusInd Bank	3	0
Karnataka Bank	1	0
Kerala Bank	0	0
Kotak Mahindra Bank	43	1
Paytm Payments Bank	1	0
Punjab National Bank	0	3
RBL Bank	3	0
South Indian Bank	0	0
State Bank Of India	5	0
Tamilnad Mercantile Bank	1	0
Union Bank of India	2	0
Yes Bank	1	0
All	134	4

Salary Acc	South Indian Bank	State Bank Of India
Primary Acc		
-	0	0
Axis Bank	0	0
Bank Of Baroda	0	0
Bank Of India	0	0
Bank Of Maharashtra	0	0
Canara Bank	0	0
DBS Bank	0	0
Equitas Small Finance Bank	0	0

```
Federal Bank                              0                   0
HDFC Bank                                 0                   1
HSBC India                                0                   0
ICICI Bank                                0                   1
IDBI Bank                                 0                   0
IDFC Bank                                 0                   0
Indian Bank                               0                   0
IndusInd Bank                             0                   0
Karnataka Bank                            0                   0
Kerala Bank                               0                   0
Kotak Mahindra Bank                       0                   1
Paytm Payments Bank                       0                   0
Punjab National Bank                      0                   0
RBL Bank                                  0                   0
South Indian Bank                         1                   0
State Bank Of India                       0                   5
Tamilnad Mercantile Bank                  0                   0
Union Bank of India                       0                   0
Yes Bank                                  0                   0
All                                       1                   8

Salary Acc                  Union Bank of India  Yes Bank  All
Primary Acc
-                                             0         0    4
Axis Bank                                     0         0   48
Bank Of Baroda                                0         0    3
Bank Of India                                 0         0    1
Bank Of Maharashtra                           0         0    1
Canara Bank                                   0         0    4
DBS Bank                                      0         0    1
Equitas Small Finance Bank                    0         0    1
Federal Bank                                  0         0    1
HDFC Bank                                     0         0  121
HSBC India                                    0         0    1
ICICI Bank                                    1         0   90
IDBI Bank                                     0         0    1
IDFC Bank                                     0         0    1
Indian Bank                                   0         0    1
IndusInd Bank                                 0         0   16
Karnataka Bank                                0         0    1
Kerala Bank                                   0         0    1
Kotak Mahindra Bank                           0         0  120
Paytm Payments Bank                           0         0    1
Punjab National Bank                          0         0    5
RBL Bank                                      0         0    5
South Indian Bank                             0         0    1
State Bank Of India                           0         0   17
Tamilnad Mercantile Bank                      0         0    1
Union Bank of India                           1         0    4
Yes Bank                                      0         2    4
All                                           2         2  455

[28 rows x 23 columns]

Cross-tabulation of Primary Acc and Salary Acc with Bank3:
 Salary Acc                 Axis Bank  Bank Of Baroda  Bank Of India  \
Primary Acc
-                                  0               0              0
Axis Bank                         29               0              0
Bank Of Baroda                     0               2              0
Bank Of India                      0               0              1
Bank Of Maharashtra                0               0              0
Canara Bank                        0               0              0
DBS Bank                           0               0              0
```

Equitas Small Finance Bank	0	0	0
Federal Bank	0	0	0
HDFC Bank	3	2	1
HSBC India	0	0	0
ICICI Bank	0	0	0
IDBI Bank	1	0	0
IDFC Bank	0	0	0
Indian Bank	0	0	0
IndusInd Bank	0	0	0
Karnataka Bank	0	0	0
Kerala Bank	1	0	0
Kotak Mahindra Bank	1	2	0
Paytm Payments Bank	0	0	0
Punjab National Bank	0	0	0
RBL Bank	0	0	0
South Indian Bank	0	0	0
State Bank Of India	3	0	0
Tamilnad Mercantile Bank	0	0	0
Union Bank of India	0	0	0
Yes Bank	0	0	0
All	38	6	2

Salary Acc	Bank Of Maharashtra	Canara Bank	DBS Bank
Primary Acc			
-	0	0	0
Axis Bank	0	0	0
Bank Of Baroda	0	0	0
Bank Of India	0	0	0
Bank Of Maharashtra	1	0	0
Canara Bank	0	2	0
DBS Bank	0	0	0
Equitas Small Finance Bank	0	0	0
Federal Bank	0	0	0
HDFC Bank	0	0	0
HSBC India	0	0	0
ICICI Bank	0	0	0
IDBI Bank	0	0	0
IDFC Bank	0	0	0
Indian Bank	0	0	0
IndusInd Bank	0	0	0
Karnataka Bank	0	0	0
Kerala Bank	0	0	0
Kotak Mahindra Bank	0	0	1
Paytm Payments Bank	0	0	0
Punjab National Bank	0	0	0
RBL Bank	0	0	0
South Indian Bank	0	0	0
State Bank Of India	0	0	0
Tamilnad Mercantile Bank	0	0	0
Union Bank of India	0	0	0
Yes Bank	0	0	0
All	1	2	1

Salary Acc	Equitas Small Finance Bank	Federal Bank
Primary Acc		
-	0	0
Axis Bank	0	0
Bank Of Baroda	0	0
Bank Of India	0	0
Bank Of Maharashtra	0	0
Canara Bank	0	0
DBS Bank	0	0
Equitas Small Finance Bank	1	0
Federal Bank	0	1

HDFC Bank	0	0
HSBC India	0	0
ICICI Bank	0	0
IDBI Bank	0	0
IDFC Bank	0	0
Indian Bank	0	0
IndusInd Bank	0	0
Karnataka Bank	0	0
Kerala Bank	0	0
Kotak Mahindra Bank	0	0
Paytm Payments Bank	0	0
Punjab National Bank	0	0
RBL Bank	0	0
South Indian Bank	0	0
State Bank Of India	0	0
Tamilnad Mercantile Bank	0	0
Union Bank of India	0	0
Yes Bank	0	0
All	1	1

Salary Acc Primary Acc	HDFC Bank	HSBC India	...	Indian Bank	\
-	1	0	...	0	
Axis Bank	0	0	...	0	
Bank Of Baroda	0	0	...	0	
Bank Of India	0	0	...	0	
Bank Of Maharashtra	0	0	...	0	
Canara Bank	1	0	...	0	
DBS Bank	0	0	...	0	
Equitas Small Finance Bank	0	0	...	0	
Federal Bank	0	0	...	0	
HDFC Bank	78	0	...	0	
HSBC India	0	1	...	0	
ICICI Bank	3	0	...	0	
IDBI Bank	0	0	...	0	
IDFC Bank	1	0	...	0	
Indian Bank	0	0	...	0	
IndusInd Bank	0	0	...	1	
Karnataka Bank	0	0	...	0	
Kerala Bank	0	0	...	0	
Kotak Mahindra Bank	1	0	...	0	
Paytm Payments Bank	0	0	...	0	
Punjab National Bank	0	0	...	0	
RBL Bank	2	0	...	0	
South Indian Bank	0	0	...	0	
State Bank Of India	3	0	...	0	
Tamilnad Mercantile Bank	0	0	...	0	
Union Bank of India	1	0	...	0	
Yes Bank	0	0	...	0	
All	91	1	...	1	

Salary Acc Primary Acc	IndusInd Bank	Kotak Mahindra Bank	\
-	0	0	
Axis Bank	0	0	
Bank Of Baroda	0	0	
Bank Of India	0	0	
Bank Of Maharashtra	0	0	
Canara Bank	0	0	
DBS Bank	0	0	
Equitas Small Finance Bank	0	0	
Federal Bank	0	0	
HDFC Bank	1	2	
HSBC India	0	0	

ICICI Bank	1	0
IDBI Bank	0	0
IDFC Bank	0	0
Indian Bank	0	0
IndusInd Bank	11	0
Karnataka Bank	0	0
Kerala Bank	0	0
Kotak Mahindra Bank	0	68
Paytm Payments Bank	0	0
Punjab National Bank	0	2
RBL Bank	0	0
South Indian Bank	0	0
State Bank Of India	0	0
Tamilnad Mercantile Bank	0	0
Union Bank of India	0	0
Yes Bank	0	1
All	13	73

Salary Acc	None of the above	Punjab National Bank	\
Primary Acc			
-	2	0	
Axis Bank	18	0	
Bank Of Baroda	1	0	
Bank Of India	0	0	
Bank Of Maharashtra	0	0	
Canara Bank	1	0	
DBS Bank	1	0	
Equitas Small Finance Bank	0	0	
Federal Bank	0	0	
HDFC Bank	31	0	
HSBC India	0	0	
ICICI Bank	19	0	
IDBI Bank	0	0	
IDFC Bank	0	0	
Indian Bank	1	0	
IndusInd Bank	3	0	
Karnataka Bank	1	0	
Kerala Bank	0	0	
Kotak Mahindra Bank	43	1	
Paytm Payments Bank	1	0	
Punjab National Bank	0	3	
RBL Bank	3	0	
South Indian Bank	0	0	
State Bank Of India	5	0	
Tamilnad Mercantile Bank	1	0	
Union Bank of India	2	0	
Yes Bank	1	0	
All	134	4	

Salary Acc	South Indian Bank	State Bank Of India	\
Primary Acc			
-	0	0	
Axis Bank	0	0	
Bank Of Baroda	0	0	
Bank Of India	0	0	
Bank Of Maharashtra	0	0	
Canara Bank	0	0	
DBS Bank	0	0	
Equitas Small Finance Bank	0	0	
Federal Bank	0	0	
HDFC Bank	0	1	
HSBC India	0	0	
ICICI Bank	0	1	
IDBI Bank	0	0	

```
IDFC Bank                          0                  0
Indian Bank                        0                  0
IndusInd Bank                      0                  0
Karnataka Bank                     0                  0
Kerala Bank                        0                  0
Kotak Mahindra Bank                0                  1
Paytm Payments Bank                0                  0
Punjab National Bank               0                  0
RBL Bank                           0                  0
South Indian Bank                  1                  0
State Bank Of India                0                  5
Tamilnad Mercantile Bank           0                  0
Union Bank of India                0                  0
Yes Bank                           0                  0
All                                1                  8

Salary Acc                 Union Bank of India  Yes Bank  All
Primary Acc
-                                            0         0    4
Axis Bank                                    0         0   48
Bank Of Baroda                               0         0    3
Bank Of India                                0         0    1
Bank Of Maharashtra                          0         0    1
Canara Bank                                  0         0    4
DBS Bank                                     0         0    1
Equitas Small Finance Bank                   0         0    1
Federal Bank                                 0         0    1
HDFC Bank                                    0         0  121
HSBC India                                   0         0    1
ICICI Bank                                   1         0   90
IDBI Bank                                    0         0    1
IDFC Bank                                    0         0    1
Indian Bank                                  0         0    1
IndusInd Bank                                0         0   16
Karnataka Bank                               0         0    1
Kerala Bank                                  0         0    1
Kotak Mahindra Bank                          0         0  120
Paytm Payments Bank                          0         0    1
Punjab National Bank                         0         0    5
RBL Bank                                     0         0    5
South Indian Bank                            0         0    1
State Bank Of India                          0         0   17
Tamilnad Mercantile Bank                     0         0    1
Union Bank of India                          1         0    4
Yes Bank                                     0         2    4
All                                          2         2  455

[28 rows x 23 columns]
```

In [93]:
```
import pandas as pd
import numpy as np
import statsmodels.api as sm
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
```

In [94]:
```
print("Summary statistics:")
print(df.describe())
```

```
Summary statistics:
       Gender    City Age group          Occupation      BANK1      BANK2  \
count     455     455       455                 455        455        453
unique      2      10         5                   5         23         32
top      Male  Mumbai     18-25  Salaried □ Private  HDFC Bank  HDFC Bank
freq      255      77       134                 300        139         77

                     BANK3 Kotak Mahindra Bank HDFC Bank ICICI Bank  \
count                  449                 455       455        455
unique                  35                   2         2          2
top     State Bank Of India                   -         -          -
freq                    62                 277       279        331

       Axis Bank IndusInd Bank RBL Bank Other 1 Other 2         Salary Acc  \
count        455           455      455     455     455                455
unique         2             2        2      24       7                 22
top            -             -        -       -       -  None of the above
freq         375           424      446     320     443                134

       Primary Acc
count          455
unique          27
top      HDFC Bank
freq           121
```

In [95]:
```
print("Missing values:")
print(df.isnull().sum())
```

```
Missing values:
Gender                0
City                  0
Age group             0
Occupation            0
BANK1                 0
BANK2                 2
BANK3                 6
Kotak Mahindra Bank   0
HDFC Bank             0
ICICI Bank            0
Axis Bank             0
IndusInd Bank         0
RBL Bank              0
Other 1               0
Other 2               0
Salary Acc            0
Primary Acc           0
dtype: int64
```

In [96]:
```
numeric_df = df.select_dtypes(include=np.number)
correlation_matrix = numeric_df.corr()
```

In [97]:
```
non_numeric_columns = df.select_dtypes(exclude=np.number).columns
print("Non-numeric columns:", non_numeric_columns)
```

```
Non-numeric columns: Index(['Gender', 'City', 'Age group', 'Occupation', 'BANK1',
'BANK2', 'BANK3',
       'Kotak Mahindra Bank', 'HDFC Bank', 'ICICI Bank', 'Axis Bank',
       'IndusInd Bank', 'RBL Bank', 'Other 1', 'Other 2', 'Salary Acc',
       'Primary Acc'],
      dtype='object')
```

In [98]:
```
encoded_df = pd.get_dummies(df)
```

In [99]:
```
corr_matrix = encoded_df.corr()
```

In [100…
```
print("Correlation matrix:")
print(corr_matrix)
```

```
Correlation matrix:
                                    Gender_Female  Gender_Male  \
Gender_Female                            1.000000    -1.000000
Gender_Male                             -1.000000     1.000000
City_Ahmedabad                           0.003611    -0.003611
City_Bangalore                          -0.003228     0.003228
City_Bhubaneswar                        -0.000793     0.000793
...                                           ...          ...
Primary Acc_South Indian Bank           -0.041564     0.041564
Primary Acc_State Bank Of India          0.012316    -0.012316
Primary Acc_Tamilnad Mercantile Bank    -0.041564     0.041564
Primary Acc_Union Bank of India         -0.035968     0.035968
Primary Acc_Yes Bank                     0.058904    -0.058904

                                    City_Ahmedabad  City_Bangalore  \
Gender_Female                             0.003611       -0.003228
Gender_Male                              -0.003611        0.003228
City_Ahmedabad                            1.000000       -0.123067
City_Bangalore                           -0.123067        1.000000
City_Bhubaneswar                         -0.070057       -0.064024
...                                            ...             ...
Primary Acc_South Indian Bank            -0.017223       -0.015739
Primary Acc_State Bank Of India           0.071036       -0.027628
Primary Acc_Tamilnad Mercantile Bank     -0.017223        0.139944
Primary Acc_Union Bank of India          -0.034559        0.046517
Primary Acc_Yes Bank                      0.038239       -0.031583

                                    City_Bhubaneswar  City_Chennai  \
Gender_Female                              -0.000793      0.008237
Gender_Male                                 0.000793     -0.008237
City_Ahmedabad                             -0.070057     -0.117024
City_Bangalore                             -0.064024     -0.106947
City_Bhubaneswar                            1.000000     -0.060880
...                                              ...           ...
Primary Acc_South Indian Bank              -0.008960     -0.014967
Primary Acc_State Bank Of India            -0.037611     -0.062826
Primary Acc_Tamilnad Mercantile Bank       -0.008960     -0.014967
Primary Acc_Union Bank of India            -0.017979      0.132643
Primary Acc_Yes Bank                       -0.017979     -0.030032

                                    City_Delhi  City_Indore  City_Kochi  \
Gender_Female                         0.015424     0.002472   -0.090124
Gender_Male                          -0.015424    -0.002472    0.090124
City_Ahmedabad                       -0.144391    -0.092169   -0.086595
City_Bangalore                       -0.131957    -0.084232   -0.079138
City_Bhubaneswar                     -0.075118    -0.047950   -0.045050
...                                        ...          ...         ...
Primary Acc_South Indian Bank        -0.018467    -0.011788    0.198886
Primary Acc_State Bank Of India      -0.043506     0.048620    0.057200
Primary Acc_Tamilnad Mercantile Bank -0.018467    -0.011788   -0.011075
Primary Acc_Union Bank of India      -0.037056    -0.023654   -0.022223
Primary Acc_Yes Bank                 -0.037056    -0.023654   -0.022223

                                    City_Kolkata  ...  \
Gender_Female                           0.119655  ...
Gender_Male                            -0.119655  ...
City_Ahmedabad                         -0.138874  ...
City_Bangalore                         -0.126915  ...
City_Bhubaneswar                       -0.072248  ...
...                                          ...  ...
Primary Acc_South Indian Bank          -0.017761  ...
Primary Acc_State Bank Of India         0.030469  ...
Primary Acc_Tamilnad Mercantile Bank   -0.017761  ...
Primary Acc_Union Bank of India        -0.035640  ...
```

```
Primary Acc_Yes Bank                   -0.035640  ...

                                    Primary Acc_Kerala Bank  \
Gender_Female                                     -0.041564
Gender_Male                                        0.041564
City_Ahmedabad                                    -0.017223
City_Bangalore                                    -0.015739
City_Bhubaneswar                                  -0.008960
...                                                     ...
Primary Acc_South Indian Bank                     -0.002203
Primary Acc_State Bank Of India                   -0.009246
Primary Acc_Tamilnad Mercantile Bank              -0.002203
Primary Acc_Union Bank of India                   -0.004420
Primary Acc_Yes Bank                              -0.004420

                                    Primary Acc_Kotak Mahindra Bank  \
Gender_Female                                          -0.007509
Gender_Male                                             0.007509
City_Ahmedabad                                         -0.019150
City_Bangalore                                         -0.068360
City_Bhubaneswar                                        0.075281
...                                                          ...
Primary Acc_South Indian Bank                          -0.028089
Primary Acc_State Bank Of India                        -0.117911
Primary Acc_Tamilnad Mercantile Bank                   -0.028089
Primary Acc_Union Bank of India                        -0.056365
Primary Acc_Yes Bank                                   -0.056365

                                    Primary Acc_Paytm Payments Bank  \
Gender_Female                                           0.052994
Gender_Male                                            -0.052994
City_Ahmedabad                                         -0.017223
City_Bangalore                                         -0.015739
City_Bhubaneswar                                       -0.008960
...                                                          ...
Primary Acc_South Indian Bank                          -0.002203
Primary Acc_State Bank Of India                        -0.009246
Primary Acc_Tamilnad Mercantile Bank                   -0.002203
Primary Acc_Union Bank of India                        -0.004420
Primary Acc_Yes Bank                                   -0.004420

                                    Primary Acc_Punjab National Bank  \
Gender_Female                                           -0.050877
Gender_Male                                              0.050877
City_Ahmedabad                                          -0.038682
City_Bangalore                                           0.034582
City_Bhubaneswar                                        -0.020124
...                                                           ...
Primary Acc_South Indian Bank                           -0.004947
Primary Acc_State Bank Of India                         -0.020767
Primary Acc_Tamilnad Mercantile Bank                    -0.004947
Primary Acc_Union Bank of India                         -0.009927
Primary Acc_Yes Bank                                    -0.009927

                                    Primary Acc_RBL Bank  \
Gender_Female                                   0.034073
Gender_Male                                    -0.034073
City_Ahmedabad                                  0.026504
City_Bangalore                                  0.034582
City_Bhubaneswar                               -0.020124
...                                                  ...
Primary Acc_South Indian Bank                  -0.004947
Primary Acc_State Bank Of India                -0.020767
Primary Acc_Tamilnad Mercantile Bank           -0.004947
```

```
Primary Acc_Union Bank of India                -0.009927
Primary Acc_Yes Bank                           -0.009927

                                           Primary Acc_South Indian Bank  \
Gender_Female                                                  -0.041564
Gender_Male                                                     0.041564
City_Ahmedabad                                                 -0.017223
City_Bangalore                                                 -0.015739
City_Bhubaneswar                                               -0.008960
...                                                                  ...
Primary Acc_South Indian Bank                                   1.000000
Primary Acc_State Bank Of India                                -0.009246
Primary Acc_Tamilnad Mercantile Bank                           -0.002203
Primary Acc_Union Bank of India                                -0.004420
Primary Acc_Yes Bank                                           -0.004420

                                           Primary Acc_State Bank Of India  \
Gender_Female                                                     0.012316
Gender_Male                                                      -0.012316
City_Ahmedabad                                                    0.071036
City_Bangalore                                                   -0.027628
City_Bhubaneswar                                                 -0.037611
...                                                                    ...
Primary Acc_South Indian Bank                                    -0.009246
Primary Acc_State Bank Of India                                   1.000000
Primary Acc_Tamilnad Mercantile Bank                             -0.009246
Primary Acc_Union Bank of India                                  -0.018554
Primary Acc_Yes Bank                                             -0.018554

                                           Primary Acc_Tamilnad Mercantile Bank  \
Gender_Female                                                          -0.041564
Gender_Male                                                             0.041564
City_Ahmedabad                                                         -0.017223
City_Bangalore                                                          0.139944
City_Bhubaneswar                                                       -0.008960
...                                                                          ...
Primary Acc_South Indian Bank                                          -0.002203
Primary Acc_State Bank Of India                                        -0.009246
Primary Acc_Tamilnad Mercantile Bank                                    1.000000
Primary Acc_Union Bank of India                                        -0.004420
Primary Acc_Yes Bank                                                   -0.004420

                                           Primary Acc_Union Bank of India  \
Gender_Female                                                     -0.035968
Gender_Male                                                        0.035968
City_Ahmedabad                                                    -0.034559
City_Bangalore                                                     0.046517
City_Bhubaneswar                                                  -0.017979
...                                                                     ...
Primary Acc_South Indian Bank                                     -0.004420
Primary Acc_State Bank Of India                                   -0.018554
Primary Acc_Tamilnad Mercantile Bank                              -0.004420
Primary Acc_Union Bank of India                                    1.000000
Primary Acc_Yes Bank                                              -0.008869

                                           Primary Acc_Yes Bank
Gender_Female                                          0.058904
Gender_Male                                           -0.058904
City_Ahmedabad                                         0.038239
City_Bangalore                                        -0.031583
City_Bhubaneswar                                      -0.017979
...                                                         ...
Primary Acc_South Indian Bank                         -0.004420
Primary Acc_State Bank Of India                       -0.018554
```

```
Primary Acc_Tamilnad Mercantile Bank            -0.004420
Primary Acc_Union Bank of India                 -0.008869
Primary Acc_Yes Bank                             1.000000

[204 rows x 204 columns]
```

In [101…
```
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report
from sklearn.preprocessing import OneHotEncoder
```

In [102…
```
X = df.drop(columns=['Primary Acc', 'Salary Acc'])
y = df['Primary Acc']
```

In [103…
```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_sta
```

In [104…
```
encoder = OneHotEncoder()
```

In [105…
```
encoder = OneHotEncoder(drop='first', sparse=False)
```

In [106…
```
combined_data = pd.concat([X_train, X_test])
```

In [107…
```
combined_encoded = encoder.fit_transform(combined_data)
```

```
C:\Users\NITISH\anaconda3\Lib\site-packages\sklearn\preprocessing\_encoders.py:97
2: FutureWarning: `sparse` was renamed to `sparse_output` in version 1.2 and will
be removed in 1.4. `sparse_output` is ignored unless you leave `sparse` to its def
ault value.
  warnings.warn(
```

In [108…
```
X_train_encoded = combined_encoded[:len(X_train)]
X_test_encoded = combined_encoded[len(X_train):]
```

In [109…
```
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report
```

In [110…
```
from sklearn.compose import ColumnTransformer
from sklearn.preprocessing import OneHotEncoder
```

In [111…
```
categorical_cols = ['Gender', 'City', 'Age group', 'Occupation', 'BANK1', 'BANK2',
encoder = ColumnTransformer(transformers=[('onehot', OneHotEncoder(), categorical_c
```

In [112…
```
X_encoded = encoder.fit_transform(df)
```

In [113…
```
X_train, X_test, y_train, y_test = train_test_split(X_encoded, y, test_size=0.2, ra
```

In [114…
```
rf_classifier = RandomForestClassifier(n_estimators=100, random_state=42)
rf_classifier.fit(X_train, y_train)
```

Out[114]:
```
▾ RandomForestClassifier

RandomForestClassifier(random_state=42)
```

In [115…
```
y_pred = rf_classifier.predict(X_test)
```

In [116…
```
print(classification_report(y_test, y_pred))
```

```
                      precision    recall  f1-score   support

                    -       0.00      0.00      0.00         1
            Axis Bank       0.88      1.00      0.93         7
  Bank Of Maharashtra       0.00      0.00      0.00         1
          Canara Bank       1.00      1.00      1.00         1
             DBS Bank       0.00      0.00      0.00         1
         Federal Bank       0.00      0.00      0.00         1
            HDFC Bank       0.84      1.00      0.91        26
           ICICI Bank       1.00      1.00      1.00        19
        IndusInd Bank       0.50      1.00      0.67         2
          Kerala Bank       0.00      0.00      0.00         1
  Kotak Mahindra Bank       0.96      1.00      0.98        25
 Punjab National Bank       0.00      0.00      0.00         1
             RBL Bank       0.00      0.00      0.00         1
  State Bank Of India       1.00      0.67      0.80         3
             Yes Bank       0.00      0.00      0.00         1

             accuracy                           0.90        91
            macro avg       0.41      0.44      0.42        91
         weighted avg       0.83      0.90      0.86        91
```

```
C:\Users\NITISH\anaconda3\Lib\site-packages\sklearn\metrics\_classification.py:146
9: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to
0.0 in labels with no predicted samples. Use `zero_division` parameter to control
this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
C:\Users\NITISH\anaconda3\Lib\site-packages\sklearn\metrics\_classification.py:146
9: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to
0.0 in labels with no predicted samples. Use `zero_division` parameter to control
this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
C:\Users\NITISH\anaconda3\Lib\site-packages\sklearn\metrics\_classification.py:146
9: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to
0.0 in labels with no predicted samples. Use `zero_division` parameter to control
this behavior.
  _warn_prf(average, modifier, msg_start, len(result))
```

In [ ]:

In [117…

```
subset = df[['Age group', 'Salary Acc', 'Primary Acc']]
```

In [118…

```
count_df = subset.groupby(['Age group', 'Salary Acc', 'Primary Acc']).size().reset_
```

In [119…

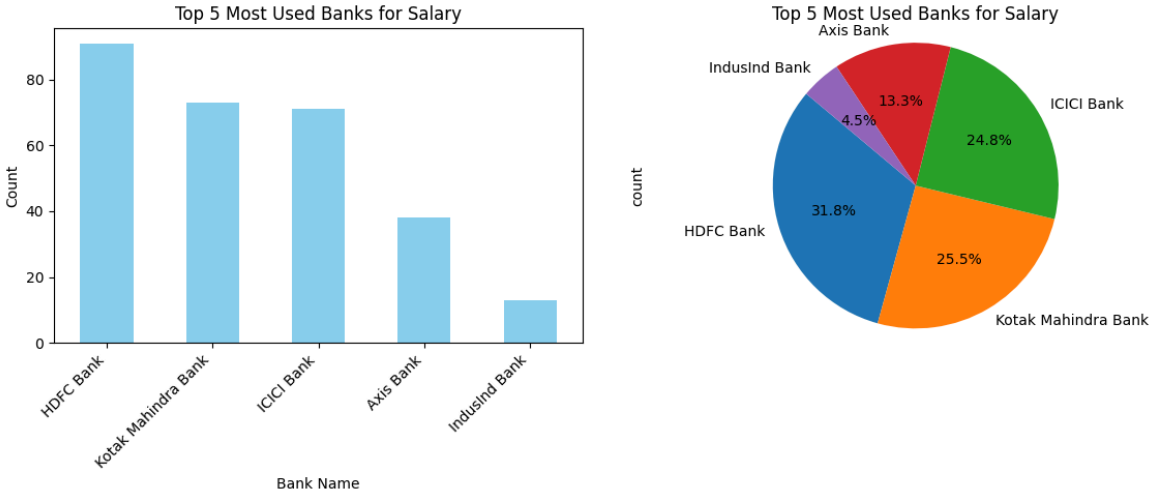
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

MOST USED AS A SALARY ACC

In [120…

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

# Sample data
data = {'Salary Acc': ['Yes Bank', 'Bank Of Baroda', 'Axis Bank', 'None of the abov

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

# Filter out "None of the above" entries
```

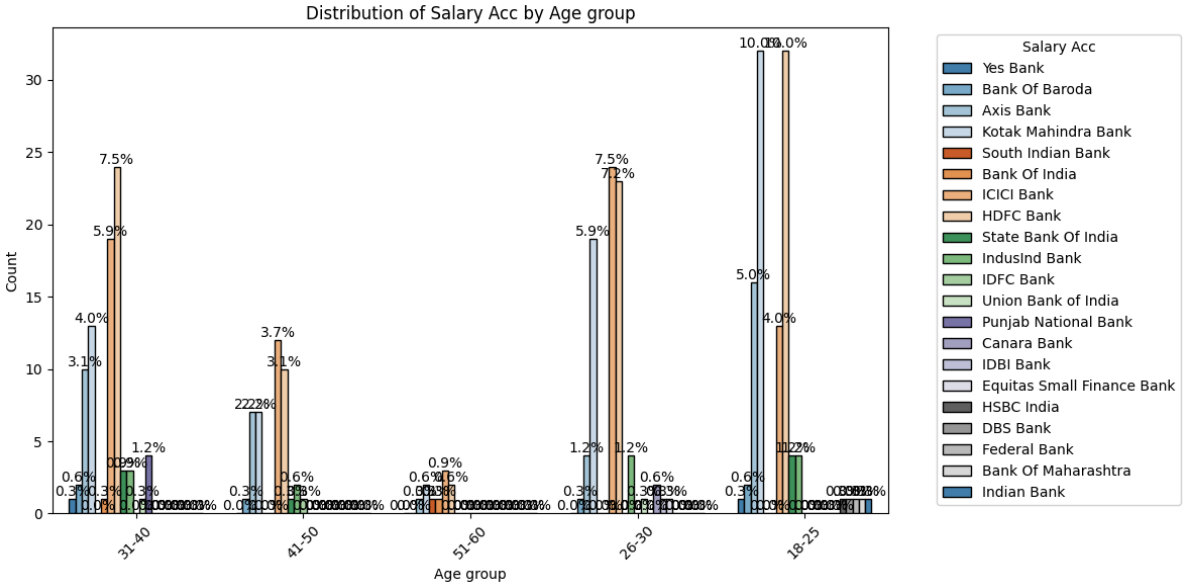
```
df_filtered = df[df['Salary Acc'] != 'None of the above']

# Get top 5 banks
top_5_banks = df_filtered['Salary Acc'].value_counts().nlargest(5)

# Plot
plt.figure(figsize=(12, 5))

plt.subplot(1, 2, 1)
top_5_banks.plot(kind='bar', color='skyblue')
plt.title('Top 5 Most Used Banks for Salary')
plt.xlabel('Bank Name')
plt.ylabel('Count')
plt.xticks(rotation=45, ha='right')

plt.subplot(1, 2, 2)
top_5_banks.plot(kind='pie', autopct='%1.1f%%', startangle=140)
plt.axis('equal')  # Equal aspect ratio ensures that pie is drawn as a circle.
plt.title('Top 5 Most Used Banks for Salary')

plt.tight_layout()
plt.show()
```



In [122…

```
import seaborn as sns
import matplotlib.pyplot as plt
from collections import Counter

# Remove 'None of the above' options and combine bank accounts
subset = subset[(subset[['Salary Acc', 'Primary Acc']] != 'None of the above').all(
combined_accounts = subset['Salary Acc'].tolist() + subset['Primary Acc'].tolist()

# Count the frequency of each bank
bank_counts = Counter(combined_accounts)

# Sort banks and counts based on counts
banks_sorted, counts_sorted = zip(*sorted(bank_counts.items(), key=lambda x: x[1],

# Define custom colors for the bars
custom_colors = plt.cm.tab20c.colors[:len(banks_sorted)]

# Distribution of Salary Acc by Age group
plt.figure(figsize=(12, 6))
ax = sns.countplot(x='Age group', hue='Salary Acc', data=subset, palette=custom_col

# Add percentage labels above each bar
for p in ax.patches:
    percentage = '{:.1f}%'.format(100 * p.get_height() / len(subset))
```

```
    ax.annotate(percentage, (p.get_x() + p.get_width() / 2, p.get_height()), ha='ce

plt.title('Distribution of Salary Acc by Age group')
plt.xlabel('Age group')
plt.ylabel('Count')
plt.legend(title='Salary Acc', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



In [123…

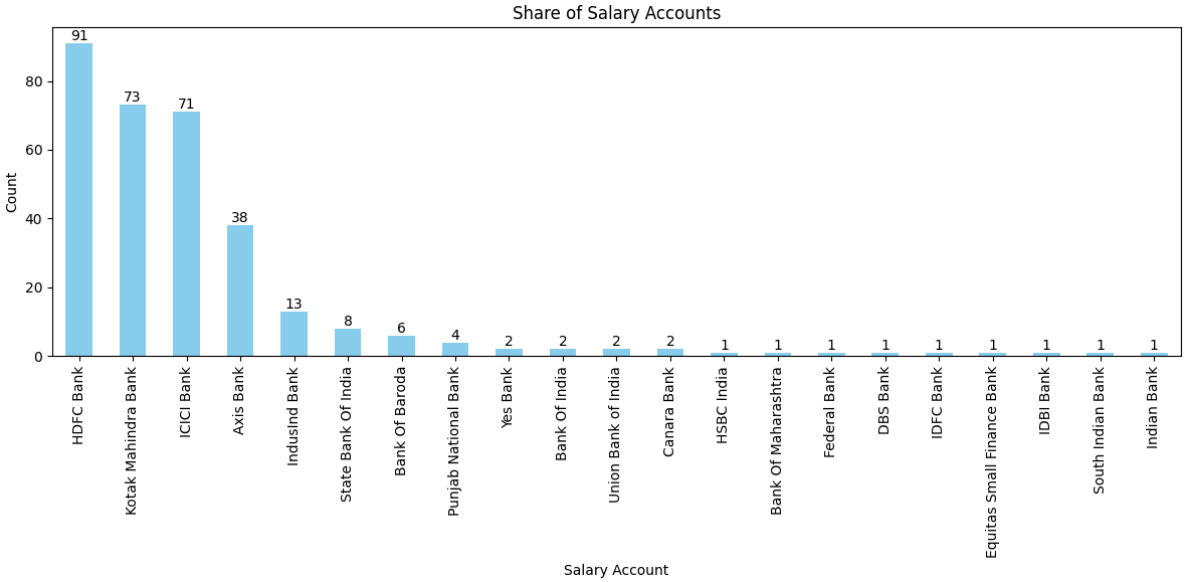
```
import pandas as pd

# Sample data
data = {'Salary Acc': ['Yes Bank', 'Bank Of Baroda', 'Axis Bank', 'None of the abov

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

print(len(df))
```

454

In [124…

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

# Sample data
data = {'Salary Acc': ['Yes Bank', 'Bank Of Baroda', 'Axis Bank', 'None of the abov

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

# Filter out "None of the above"
df_filtered = df[df['Salary Acc'] != 'None of the above']

# Plot
plt.figure(figsize=(12, 6))
df_filtered['Salary Acc'].value_counts().plot(kind='bar', color='skyblue')
plt.title('Share of Salary Accounts')
plt.xlabel('Salary Account')
plt.ylabel('Count')
plt.xticks(rotation=90)

# Display counts on each bar
counts = df_filtered['Salary Acc'].value_counts()
for i, count in enumerate(counts):
```

```
    plt.text(i, count, str(count), ha='center', va='bottom')

plt.tight_layout()
plt.show()
```



Primary Acc

In [125…

```
plt.figure(figsize=(8, 8))
df['Primary Acc'].value_counts().plot(kind='pie', autopct='%1.1f%%', startangle=140
plt.title('Total Number of Primary Acc')
plt.ylabel('')
plt.show()
```

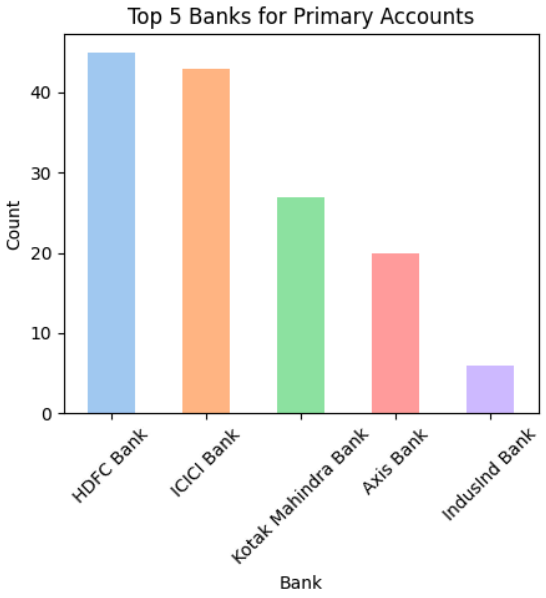
```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3805, in Index.get_
loc(self, key)
   3804 try:
-> 3805     return self._engine.get_loc(casted_key)
   3806 except KeyError as err:

File index.pyx:167, in pandas._libs.index.IndexEngine.get_loc()

File index.pyx:196, in pandas._libs.index.IndexEngine.get_loc()

File pandas\\_libs\\hashtable_class_helper.pxi:7081, in pandas._libs.hashtable.PyO
bjectHashTable.get_item()

File pandas\\_libs\\hashtable_class_helper.pxi:7089, in pandas._libs.hashtable.PyO
bjectHashTable.get_item()

KeyError: 'Primary Acc'

The above exception was the direct cause of the following exception:

KeyError                                  Traceback (most recent call last)
Cell In[125], line 2
      1 plt.figure(figsize=(8, 8))
----> 2 df['Primary Acc'].value_counts().plot(kind='pie', autopct='%1.1f%%', start
angle=140, colors=sns.color_palette('pastel'))
      3 plt.title('Total Number of Primary Acc')
      4 plt.ylabel('')

File ~\anaconda3\Lib\site-packages\pandas\core\frame.py:4090, in DataFrame.__getit
em__(self, key)
   4088 if self.columns.nlevels > 1:
   4089     return self._getitem_multilevel(key)
-> 4090 indexer = self.columns.get_loc(key)
   4091 if is_integer(indexer):
   4092     indexer = [indexer]

File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3812, in Index.get_
loc(self, key)
   3807     if isinstance(casted_key, slice) or (
   3808         isinstance(casted_key, abc.Iterable)
   3809         and any(isinstance(x, slice) for x in casted_key)
   3810     ):
   3811         raise InvalidIndexError(key)
-> 3812     raise KeyError(key) from err
   3813 except TypeError:
   3814     # If we have a listlike key, _check_indexing_error will raise
   3815     #  InvalidIndexError. Otherwise we fall through and re-raise
   3816     #  the TypeError.
   3817     self._check_indexing_error(key)

KeyError: 'Primary Acc'
```

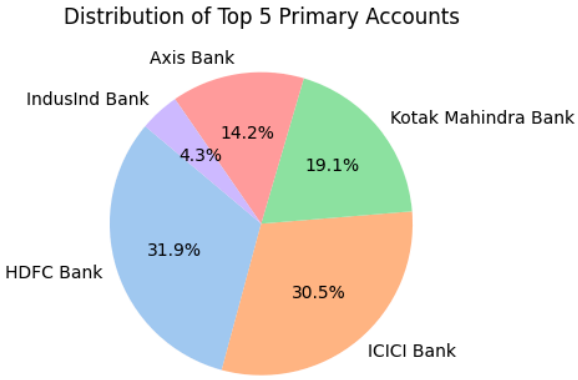
<Figure size 800x800 with 0 Axes>

In [126…

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Sample data
data = {'Primary Acc': ['Yes Bank', 'Bank Of Baroda', 'Axis Bank', 'None of the abo

# Convert to DataFrame
df = pd.DataFrame(data)
```

```
# Remove "None of the above" entries
df_filtered = df[df['Primary Acc'] != 'None of the above']

# Get top 5 banks
top_5_banks = df_filtered['Primary Acc'].value_counts().head(5)

# Plot bar chart for top 5 banks
plt.figure(figsize=(10, 5))
plt.subplot(1, 2, 1)
top_5_banks.plot(kind='bar', color=sns.color_palette('pastel'))
plt.title('Top 5 Banks for Primary Accounts')
plt.xlabel('Bank')
plt.ylabel('Count')
plt.xticks(rotation=45)

# Plot pie chart for top 5 banks
plt.subplot(1, 2, 2)
top_5_banks.plot(kind='pie', autopct='%1.1f%%', startangle=140, colors=sns.color_pa
plt.title('Distribution of Top 5 Primary Accounts')
plt.ylabel('')

plt.tight_layout()
plt.show()
```

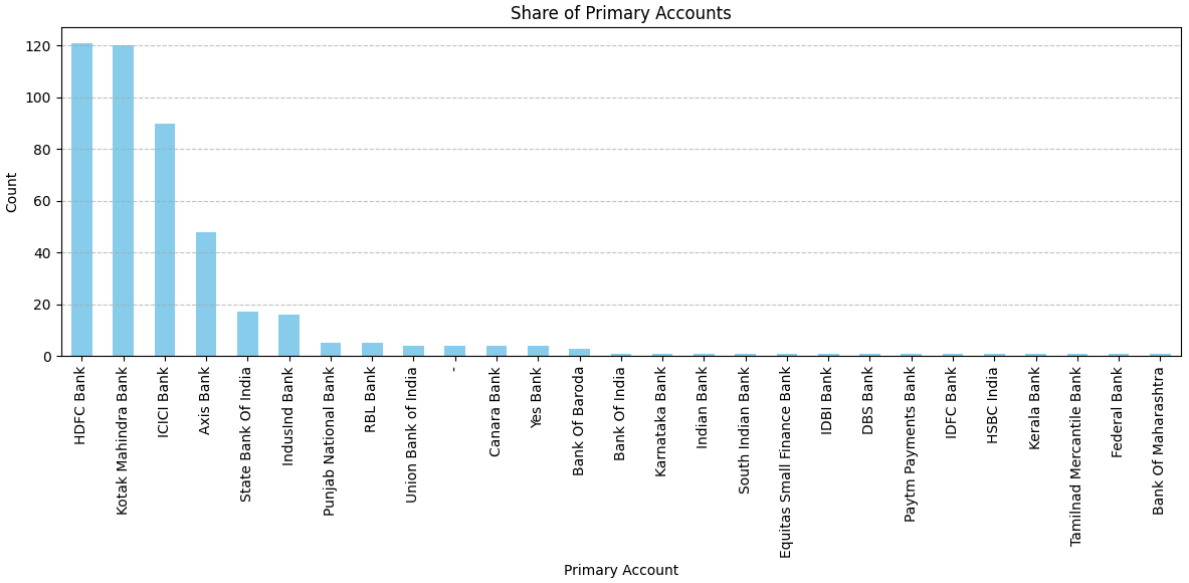




In [127…

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

# Sample data
data = {'Primary Acc': ['Yes Bank', 'HDFC Bank', 'Kotak Mahindra Bank', 'ICICI Bank

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

# Count occurrences of each primary account
primary_acc_counts = df['Primary Acc'].value_counts()

# Plot
plt.figure(figsize=(12, 6))
primary_acc_counts.plot(kind='bar', color='skyblue')
plt.title('Share of Primary Accounts')
plt.xlabel('Primary Account')
plt.ylabel('Count')
plt.xticks(rotation=90)
```

```
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```



In [128…
```
df=pd.read_csv("C:/Users/NITISH/OneDrive/Documents/Desktop/Bank Data/TOP BANKS.csv"
```

In [129…
```
print(df)
```

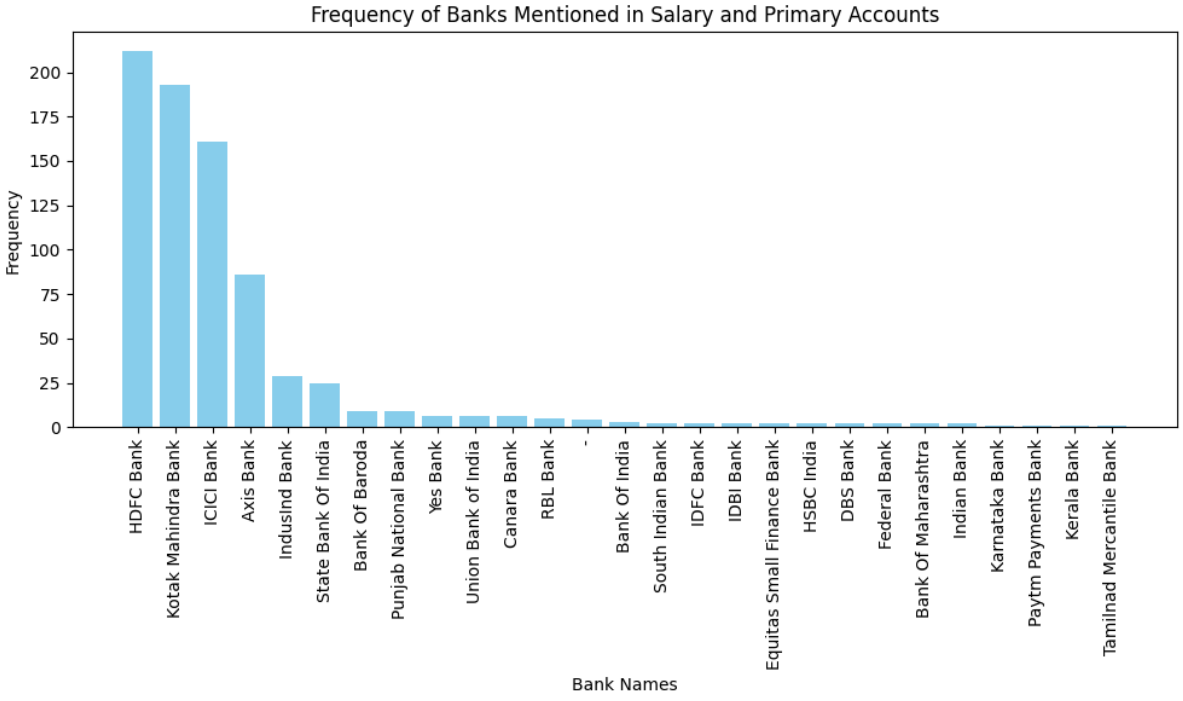
```
     Gender       City Age group          Occupation               BANK1  \
0      Male     Mumbai     31-40        Own Business           HDFC Bank
1    Female  Bangalore     41-50  Salaried  Private          ICICI Bank
2      Male  Ahmedabad     41-50        Own Business  Kotak Mahindra Bank
3      Male  Bangalore     26-30  Salaried  Private           HDFC Bank
4      Male  Ahmedabad     51-60        Own Business  Kotak Mahindra Bank
..      ...        ...       ...                 ...                  ...
450    Male  Bangalore     18-25  Salaried  Private  Kotak Mahindra Bank
451    Male  Bangalore     18-25  Salaried  Private       IndusInd Bank
452    Male      Delhi     18-25  Salaried  Private           HDFC Bank
453    Male  Ahmedabad     18-25  Salaried  Private  Kotak Mahindra Bank
454    Male  Bangalore     18-25  Salaried  Private  Kotak Mahindra Bank

                   BANK2                BANK3  Kotak Mahindra Bank  \
0             ICICI Bank             Yes Bank  Kotak Mahindra Bank
1              HDFC Bank  Kotak Mahindra Bank  Kotak Mahindra Bank
2              HDFC Bank           Axis Bank                    -
3             ICICI Bank  Kotak Mahindra Bank                    -
4              HDFC Bank  State Bank Of India  Kotak Mahindra Bank
..                   ...                  ...                  ...
450        IndusInd Bank           HDFC Bank  Kotak Mahindra Bank
451          Indian Bank         Federal Bank                    -
452  Kotak Mahindra Bank           ICICI Bank                    -
453          ICICI Bank  Central Bank Of India  Kotak Mahindra Bank
454  State Bank Of India                  NaN  Kotak Mahindra Bank

     HDFC Bank ICICI Bank Axis Bank  IndusInd Bank RBL Bank  \
0            -          -         -              -        -
1            -          -         -              -        -
2            -          -         -  IndusInd Bank        -
3    HDFC Bank          -         -              -        -
4            -          -         -              -        -
..         ...        ...       ...            ...      ...
450          -          -         -              -        -
451          -          -         -  IndusInd Bank        -
452  HDFC Bank          -         -              -        -
453          -          -         -              -        -
454          -          -         -              -        -

                Other 1 Other 2           Salary Acc          Primary Acc
0                     -       -             Yes Bank             Yes Bank
1                     -       -       Bank Of Baroda            HDFC Bank
2      South Indian Bank       -            Axis Bank  Kotak Mahindra Bank
3          Bank Of India       -    None of the above           ICICI Bank
4                     -       -  Kotak Mahindra Bank  Kotak Mahindra Bank
..                  ...     ...                  ...                  ...
450                   -       -  Kotak Mahindra Bank  Kotak Mahindra Bank
451          Indian Bank       -          Indian Bank        IndusInd Bank
452                   -       -            HDFC Bank            HDFC Bank
453                   -       -           ICICI Bank  Kotak Mahindra Bank
454  State Bank Of India       -    None of the above  Kotak Mahindra Bank

[455 rows x 17 columns]
```

Total Number of (Primary+Salary) Bank Users

In [130…

```
import matplotlib.pyplot as plt
from collections import Counter

# Extract salary accounts and primary accounts
```

```
salary_accounts = df.loc[df['Salary Acc'] != 'None of the above', 'Salary Acc'].tol
primary_accounts = df.loc[df['Primary Acc'] != 'None of the above', 'Primary Acc'].

# Combine the lists
combined_accounts = salary_accounts + primary_accounts

# Count the frequency of each bank
bank_counts = Counter(combined_accounts)

# Extract bank names and corresponding counts
banks = list(bank_counts.keys())
counts = list(bank_counts.values())

# Sort banks and counts based on counts
sorted_indices = sorted(range(len(counts)), key=lambda k: counts[k], reverse=True)
banks_sorted = [banks[i] for i in sorted_indices]
counts_sorted = [counts[i] for i in sorted_indices]

# Plotting
plt.figure(figsize=(10, 6))
plt.bar(banks_sorted, counts_sorted, color='skyblue')
plt.xlabel('Bank Names')
plt.ylabel('Frequency')
plt.title('Frequency of Banks Mentioned in Salary and Primary Accounts')
plt.xticks(rotation=90)
plt.tight_layout()
plt.show()
```



In [ ]: