Report on Call Center Financial Data Analysis

Introduction

This report provides an analysis of financial data extracted from the "National Call Center.xlsx" file. The focus of the analysis is on current account balances and past due amounts, segmented by caller and account holder gender. Various visualizations, including bar charts and scatter plots, were generated to offer deeper insights into the data.

Methodology

Data analysis

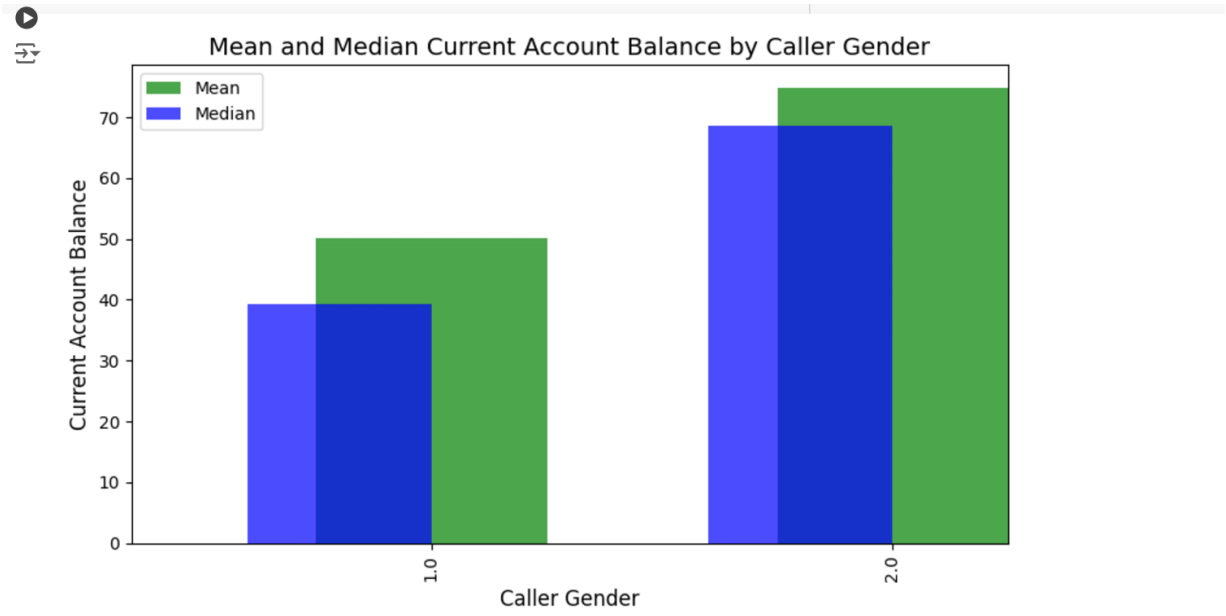
1.0. Loading of the data

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

# Load the dataset from Excel
file_path = "National Call Center.xlsx"
df = pd.read_excel(file_path)
```

1.1. Mean and Median Current Account Balance by Caller Gender

The analysis begins with the computation of the mean and median current account balances by caller gender.

```
# Task (a): Bar Charts for Mean and Median Current Account Balance by Caller Gender
mean_balance_caller = df.groupby('Caller Gender')['Current Amount Due'].mean()
median_balance_caller = df.groupby('Caller Gender')['Current Amount Due'].median()

plt.figure(figsize=(8, 5))
mean_balance_caller.plot(kind='bar', color='green', alpha=0.7, label='Mean')
median_balance_caller.plot(kind='bar', color='blue', alpha=0.7, label='Median', width=0.4, position=1)
plt.title("Mean and Median Current Account Balance by Caller Gender", fontsize=14)
plt.xlabel("Caller Gender", fontsize=12)
plt.ylabel("Current Account Balance", fontsize=12)
plt.legend()
plt.tight_layout()
plt.show()
```



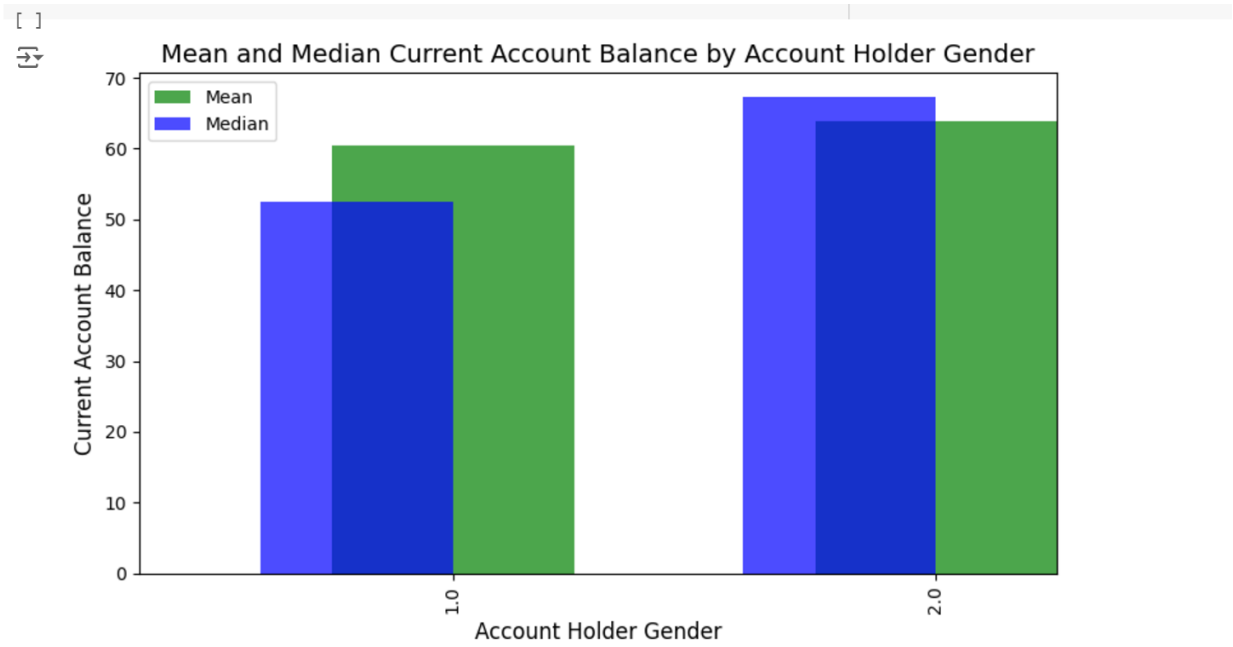
Findings:

- The bar chart illustrates the average balances for both male and female callers.
- It is observed that the mean balance may be influenced by a few high-value accounts, while the median offers a better understanding of the central tendency among the balances.

1.2. Mean and Median Current Account Balance by Account Holder Gender

This task assessed the mean and median balances based on account holder gender.

```
# Task (b): Bar Charts for Mean and Median Current Account Balance by Account Holder Gender
mean_balance_holder = df.groupby('Account Holder Gender')['Current Amount Due'].mean()
median_balance_holder = df.groupby('Account Holder Gender')['Current Amount Due'].median()

plt.figure(figsize=(8, 5))
mean_balance_holder.plot(kind='bar', color='green', alpha=0.7, label='Mean')
median_balance_holder.plot(kind='bar', color='blue', alpha=0.7, label='Median', width=0.4, position=1)
plt.title("Mean and Median Current Account Balance by Account Holder Gender", fontsize=14)
plt.xlabel("Account Holder Gender", fontsize=12)
plt.ylabel("Current Account Balance", fontsize=12)
plt.legend()
plt.tight_layout()
plt.show()
```



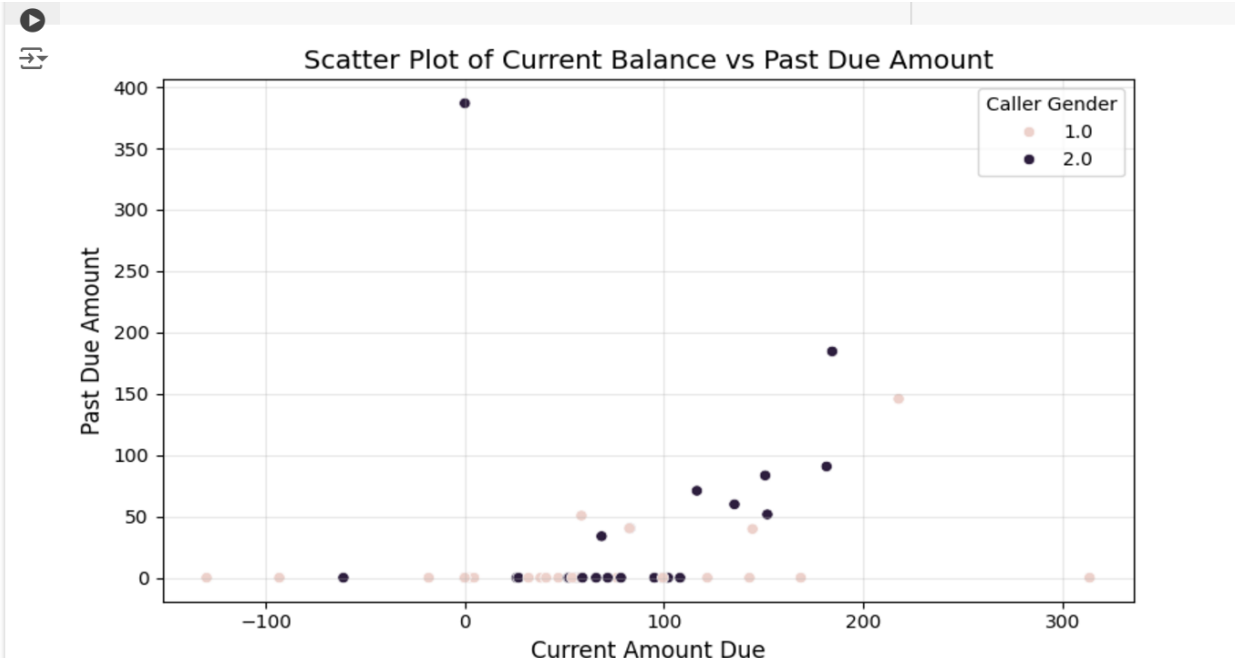
Observations:

- Similar trends are noted in this analysis compared to caller gender, indicating potential gender-based disparities in the account balances.
- The visual comparison helps in understanding whether the gender of the account holder impacts their financial position.

1.3. Scatter Diagram of Current Balance vs. Past Due Amount

A scatter plot was created to investigate the relationship between current account balances and past due amounts, further segmented by caller gender.

```
# Task (c): Scatter Diagram of Current Balance vs Past Due Amount
plt.figure(figsize=(8, 5))
sns.scatterplot(data=df, x='Current Amount Due', y='Past Due Amount', hue='Caller Gender')
plt.title("Scatter Plot of Current Balance vs Past Due Amount", fontsize=14)
plt.xlabel("Current Amount Due", fontsize=12)
plt.ylabel("Past Due Amount", fontsize=12)
plt.grid(alpha=0.3)
plt.tight_layout()
plt.show()
```



Insights:

- The scatter plot reveals how well current balances correlate with past due amounts.
- Trends can be observed based on caller gender, which could indicate differences in payment behavior or financial management between male and female callers.

1.4. Descriptive Statistics for Current Account Balance

Descriptive statistics, including count, mean, standard deviation, min, quartiles, and max values, were computed for current account balances, segmented by caller gender, account holder gender, and whether the inquiry was related to billing.

```
from tabulate import tabulate

# Task (d): Descriptive Statistics for Current Account Balance
descriptive_stats_balance = df.groupby(
    ['Caller Gender', 'Account Holder Gender', 'Was this a Billing Question?']
)['Current Amount Due'].describe()

# Convert the statistics to a table-like format
print("\nDescriptive Statistics for Current Account Balance:")
print(tabulate(descriptive_stats_balance.reset_index(), headers='keys', tablefmt='grid'))
```

```
ve Statistics for Current Account Balance:
+---------------+-------------------------+-------------------------------+---------+----------+----------+----------+---------+---------+----------+--------+
| Caller Gender | Account Holder Gender   | Was this a Billing Question?  |   count |     mean |      std |      min |     25% |     50% |      75% |    max |
+===============+=========================+===============================+=========+==========+==========+==========+=========+=========+==========+========+
|             1 |                       1 | No                            |      15 | 50.68    | 109.472  | -129.67  | 0       | 40.79   | 110.635  | 313.78 |
+---------------+-------------------------+-------------------------------+---------+----------+----------+----------+---------+---------+----------+--------+
|             1 |                       1 | Yes                           |       7 | 74.0071  | 80.4948  |    0     | 18.305  | 37.98   | 112.8    | 217.86 |
+---------------+-------------------------+-------------------------------+---------+----------+----------+----------+---------+---------+----------+--------+
|             1 |                       2 | No                            |       5 | 25.098   | 34.9858  |    0     | 0       | 2.13    | 46.98    |  76.38 |
+---------------+-------------------------+-------------------------------+---------+----------+----------+----------+---------+---------+----------+--------+
|             1 |                       2 | Yes                           |       3 | 33.1467  | 30.0332  |    0     | 20.445  | 40.89   | 49.72    |  58.55 |
+---------------+-------------------------+-------------------------------+---------+----------+----------+----------+---------+---------+----------+--------+
|             2 |                       1 | No                            |       4 | 61.6675  | 33.8863  |  27.03   | 46.125  | 55.78   | 71.3225  | 108.08 |
+---------------+-------------------------+-------------------------------+---------+----------+----------+----------+---------+---------+----------+--------+
|             2 |                       1 | Yes                           |       7 | 66.8314  | 62.6414  |    0     | 26.475  | 54.57   | 104.19   | 151.92 |
+---------------+-------------------------+-------------------------------+---------+----------+----------+----------+---------+---------+----------+--------+
|             2 |                       2 | No                            |       6 | 58.6833  | 40.207   |    0     | 34.265  | 62.375  | 90.9125  | 102.03 |
+---------------+-------------------------+-------------------------------+---------+----------+----------+----------+---------+---------+----------+--------+
|             2 |                       2 | Yes                           |      10 | 95.403   | 69.6032  | -60.98   | 73.415  | 89.03   | 130.692  | 184.49 |
+---------------+-------------------------+-------------------------------+---------+----------+----------+----------+---------+---------+----------+--------+
```

Summary:

- These statistics provide a detailed overview of the financial behaviors across different groups.
- Variability in account balances can help identify customer segments that may require additional support or intervention.

1.5. Descriptive Statistics for Past Due Amounts

Descriptive statistics for past due balances were calculated similarly, considering the same segmentation factors.

```python
# Task (e): Descriptive Statistics for Past Due Balances
descriptive_stats_past_due = df.groupby(
    ['Caller Gender', 'Account Holder Gender', 'Was this a Billing Question?']
)['Past Due Amount'].describe()

# Convert the statistics to a table-like format
print("\nDescriptive Statistics for Past Due Amount:")
print(tabulate(descriptive_stats_past_due.reset_index(), headers='keys', tablefmt='grid'))
```

```
riptive Statistics for Past Due Amount:
+---+---------------+-------------------------+-------------------------------+---------+----------+----------+---------+---------+---------+----------+--------+
|   | Caller Gender | Account Holder Gender   | Was this a Billing Question?  |   count |     mean |      std |     min |     25% |     50% |      75% |    max |
+===+===============+=========================+===============================+=========+==========+==========+=========+=========+=========+==========+========+
| 0 |             1 |                       1 | No                            |      15 |  2.65    | 10.2634  |     0   |     0   |  0      |  0       |  39.75 |
+---+---------------+-------------------------+-------------------------------+---------+----------+----------+---------+---------+---------+----------+--------+
| 1 |             1 |                       1 | Yes                           |       7 | 26.58    | 54.6703  |     0   |     0   |  0      | 20.13    | 145.8  |
+---+---------------+-------------------------+-------------------------------+---------+----------+----------+---------+---------+---------+----------+--------+
| 2 |             1 |                       2 | No                            |       5 |  0       |  0       |     0   |     0   |  0      |  0       |   0    |
+---+---------------+-------------------------+-------------------------------+---------+----------+----------+---------+---------+---------+----------+--------+
| 3 |             1 |                       2 | Yes                           |       3 | 16.88    | 29.237   |     0   |     0   |  0      | 25.32    |  50.64 |
+---+---------------+-------------------------+-------------------------------+---------+----------+----------+---------+---------+---------+----------+--------+
| 4 |             2 |                       1 | No                            |       4 |  0       |  0       |     0   |     0   |  0      |  0       |   0    |
+---+---------------+-------------------------+-------------------------------+---------+----------+----------+---------+---------+---------+----------+--------+
| 5 |             2 |                       1 | Yes                           |       7 | 74.5371  | 141.607  |     0   |     0   |  0      | 67.495   | 386.77 |
+---+---------------+-------------------------+-------------------------------+---------+----------+----------+---------+---------+---------+----------+--------+
| 6 |             2 |                       2 | No                            |       6 |  0       |  0       |     0   |     0   |  0      |  0       |   0    |
+---+---------------+-------------------------+-------------------------------+---------+----------+----------+---------+---------+---------+----------+--------+
| 7 |             2 |                       2 | Yes                           |      10 | 48.009   | 58.3019  |     0   |     0   | 37.09   | 68.125   | 184.49 |
+---+---------------+-------------------------+-------------------------------+---------+----------+----------+---------+---------+---------+----------+--------+
```

Key Findings:

- The data presents a comprehensive picture of overdue payments, offering insights into payment patterns among different demographics.

- This information can aid decision-making regarding follow-up actions for past due accounts.

Conclusion

This analysis sheds light on the financial dynamics presented within the National Call Center dataset. The visualizations and statistical summaries reveal potential gender disparities in account management and highlight areas where targeted strategies could be beneficial.