Aman Bhandari

Data Set: Customer Call List.xlsx

	A	B	C	D	E	F	G	H
1	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	Not_Useful_Column
2	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Yes	No	TRUE
3	1002	Abed	Nadir	123/643/9775	93 West Main Street	No	Yes	FALSE
4	1003	Walter	/White	7066950392	298 Drugs Driveway	N		TRUE
5	1004	Dwight	Schrute	123-543-2345	980 Paper Avenue, Pennsylvania, 18503	Yes	Y	TRUE
6	1005	Jon	Snow	876|678|3469	123 Dragons Road	Y	No	TRUE
7	1006	Ron	Swanson	304-762-2467	768 City Parkway	Yes	Yes	TRUE
8	1007	Jeff	Winger		1209 South Street	No	No	FALSE
9	1008	Sherlock	Holmes	876|678|3469	98 Clue Drive	N	No	FALSE
10	1009	Gandalf		N/a	123 Middle Earth	Yes		FALSE
11	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Yes	No	TRUE
12	1011	Samwise	Gamgee		612 Shire Lane, Shire	Yes	No	TRUE
13	1012	Harry	...Potter	7066950392	2394 Hogwarts Avenue	Y		TRUE
14	1013	Don	Draper	123-543-2345	2039 Main Street	Yes	N	FALSE
15	1014	Leslie	Knope	876|678|3469	343 City Parkway	Yes	No	FALSE
16	1015	Toby	Flenderson_	304-762-2467	214 HR Avenue	N	No	FALSE
17	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	No	N	FALSE
18	1017	Michael	Scott	123/643/9775	121 Paper Avenue, Pennsylvania	Yes	No	FALSE
19	1018	Clark	Kent	7066950392	3498 Super Lane	Y		TRUE
20	1019	Creed	Braton	N/a	N/a	N/a	Yes	TRUE
21	1020	Anakin	Skywalker	876|678|3469	910 Tatooine Road, Tatooine	Yes	N	TRUE
22	1020	Anakin	Skywalker	876|678|3469	910 Tatooine Road, Tatooine	Yes	N	TRUE
23								

Aman Bhandari

Python Code : index.py

index.py

index.py > ...

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

# Load the Excel file
file_path = "Customer Call List.xlsx"
df = pd.read_excel(file_path)

# Make a copy for cleaning
df_cleaned = df.copy()

# Step 1: Standardize column names (remove spaces)
df_cleaned.columns = df_cleaned.columns.str.strip().str.replace(' ', '_')

# Step 2: Drop completely empty rows
df_cleaned.dropna(how='all', inplace=True)

# Step 3: Handle missing values
df_cleaned['Phone_Number'] = df_cleaned['Phone_Number'].fillna('Unknown')
df_cleaned['Last_Name'] = df_cleaned['Last_Name'].fillna('Unknown')
df_cleaned['Do_Not_Contact'] = df_cleaned['Do_Not_Contact'].fillna('No')

# Step 4: Drop duplicates
df_cleaned.drop_duplicates(inplace=True)

# Step 5: Remove unnecessary columns
if 'Not_Useful_Column' in df_cleaned.columns:
    df_cleaned.drop(columns=['Not_Useful_Column'], inplace=True)

# Step 6: Save cleaned data to Excel
output_excel = "Refined_Customer_Call_List.xlsx"
df_cleaned.to_excel(output_excel, index=False)

# Step 7: Set seaborn style
sns.set(style="whitegrid")

```

Aman Bhandari

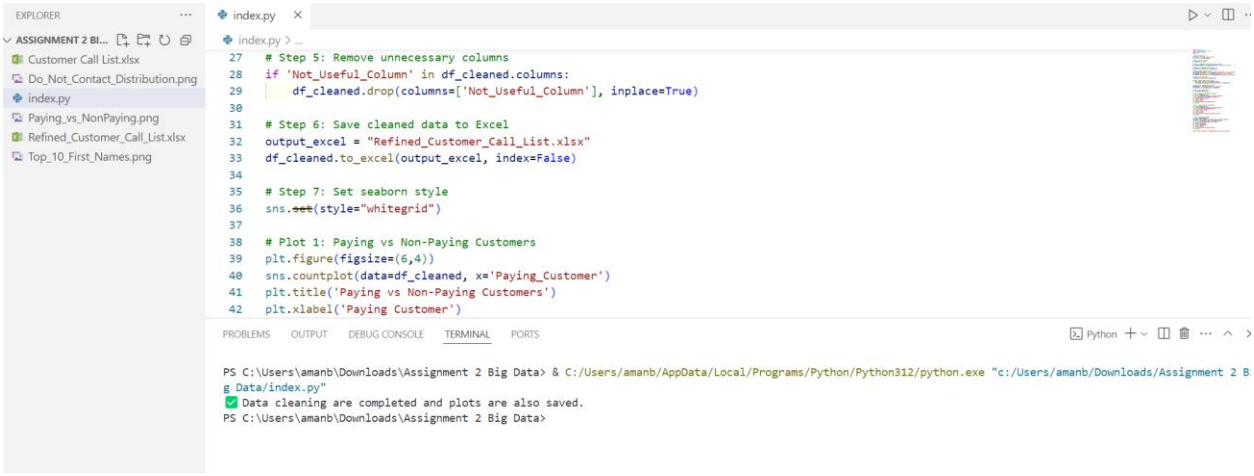
index.py

index.py > ...

```python

# Plot 1: Paying vs Non-Paying Customers
plt.figure(figsize=(6,4))
sns.countplot(data=df_cleaned, x='Paying_Customer')
plt.title('Paying vs Non-Paying Customers')
plt.xlabel('Paying Customer')
plt.ylabel('Count')
plt.tight_layout()
plt.savefig("Paying_vs_NonPaying.png")
plt.close()

# Plot 2: Do Not Contact Distribution
plt.figure(figsize=(6,4))
sns.countplot(data=df_cleaned, x='Do_Not_Contact')
plt.title('Do Not Contact Preference')
plt.xlabel('Do Not Contact')
plt.ylabel('Count')
plt.tight_layout()
plt.savefig("Do_Not_Contact_Distribution.png")
plt.close()

# Plot 3: Top 10 First Names
plt.figure(figsize=(8,5))
top_names = df_cleaned['First_Name'].value_counts().head(10)
sns.barplot(x=top_names.index, y=top_names.values)
plt.title('Top 10 Most Common First Names')
plt.xlabel('First Name')
plt.ylabel('Frequency')
plt.xticks(rotation=45)
plt.tight_layout()
plt.savefig("Top_10_First_Names.png")
plt.close()

print("✅ Data cleaning are completed and plots are also saved.")

```

Aman Bhandari

Output :



Excel Created: As mentioned on Step 6 : Line no 32 in index.py File

Refined_Customer_Call_List.xlsx

	A	B	C	D	E	F	G
1	**CustomerID**	**First_Name**	**Last_Name**	**Phone_Number**	**Address**	**Paying_Customer**	**Do_Not_Contact**
2	1001	Frodo	Baggins	123-545-5421	123 Shire Lane, Shire	Yes	No
3	1002	Abed	Nadir	123/643/9775	93 West Main Street	No	Yes
4	1003	Walter	/White	7066950392	298 Drugs Driveway	N	No
5	1004	Dwight	Schrute	123-543-2345	980 Paper Avenue, Pennsylvania, 18503	Yes	Y
6	1005	Jon	Snow	876|678|3469	123 Dragons Road	Y	No
7	1006	Ron	Swanson	304-762-2467	768 City Parkway	Yes	Yes
8	1007	Jeff	Winger	Unknown	1209 South Street	No	No
9	1008	Sherlock	Holmes	876|678|3469	98 Clue Drive	N	No
10	1009	Gandalf	Unknown	N/a	123 Middle Earth	Yes	No
11	1010	Peter	Parker	123-545-5421	25th Main Street, New York	Yes	No
12	1011	Samwise	Gamgee	Unknown	612 Shire Lane, Shire	Yes	No
13	1012	Harry	...Potter	7066950392	2394 Hogwarts Avenue	Y	No
14	1013	Don	Draper	123-543-2345	2039 Main Street	Yes	N
15	1014	Leslie	Knope	876|678|3469	343 City Parkway	Yes	No
16	1015	Toby	Flenderson_	304-762-2467	214 HR Avenue	N	No
17	1016	Ron	Weasley	123-545-5421	2395 Hogwarts Avenue	No	N
18	1017	Michael	Scott	123/643/9775	121 Paper Avenue, Pennsylvania	Yes	No
19	1018	Clark	Kent	7066950392	3498 Super Lane	Y	No
20	1019	Creed	Braton	N/a	N/a	N/a	Yes
21	1020	Anakin	Skywalker	876|678|3469	910 Tatooine Road, Tatooine	Yes	N
22							

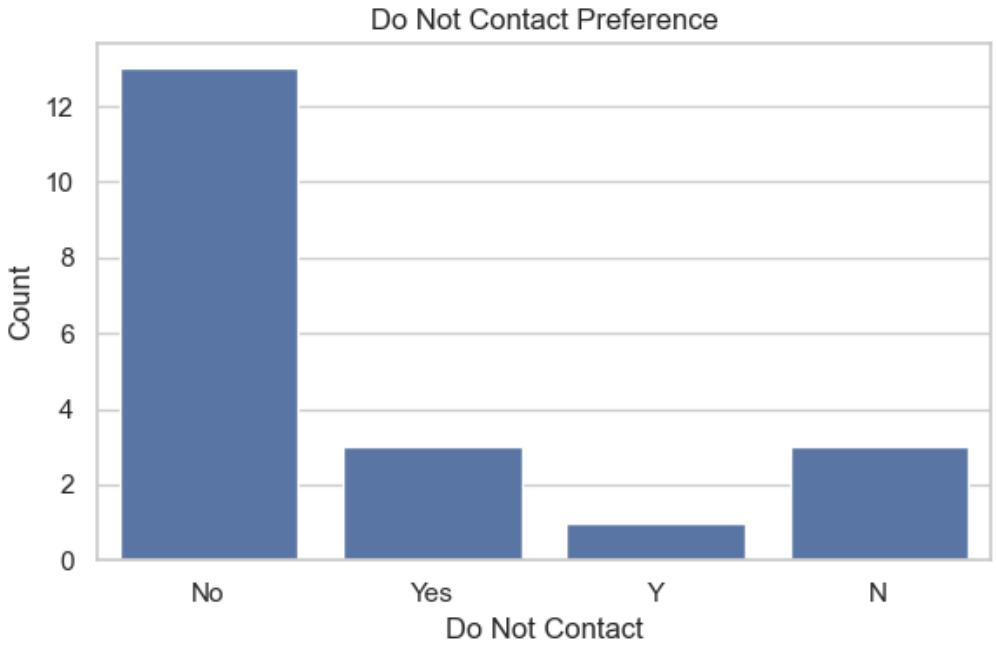
Aman Bhandari

Plot 1: As mentioned in line 38 of index.py file

Paying_vs_NonPaying.png

Plot 2: As mentioned in line 48 of index.py file

Do_Not_Contact_Distribution.png



Aman Bhandari

Plot 3: As mentioned in line 58 of index.py file

Top_10_First_Names.png