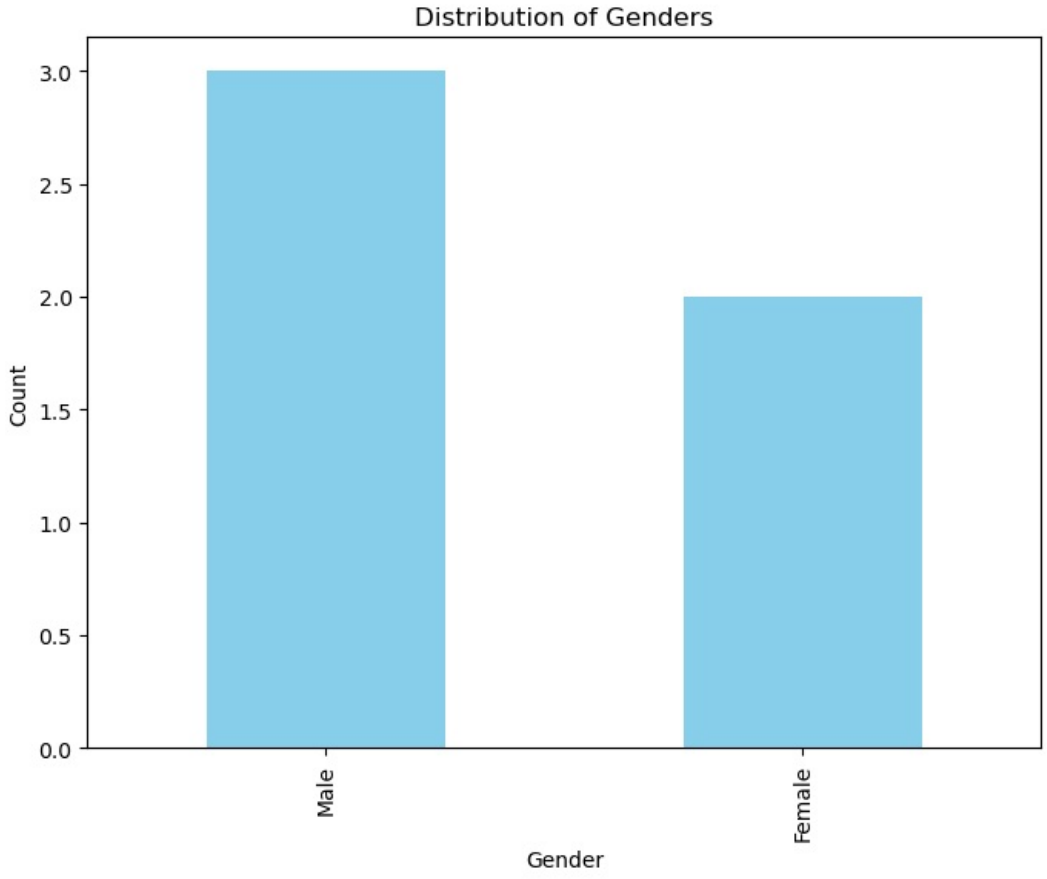
In [2]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [21]:
```
# Create a sample dataframe
data = {'Age': [25, 30, 35, 40, 45], 'Gender': ['Male', 'Female', 'Male', 'Female', 'Male']}
df = pd.DataFrame(data)
df
```

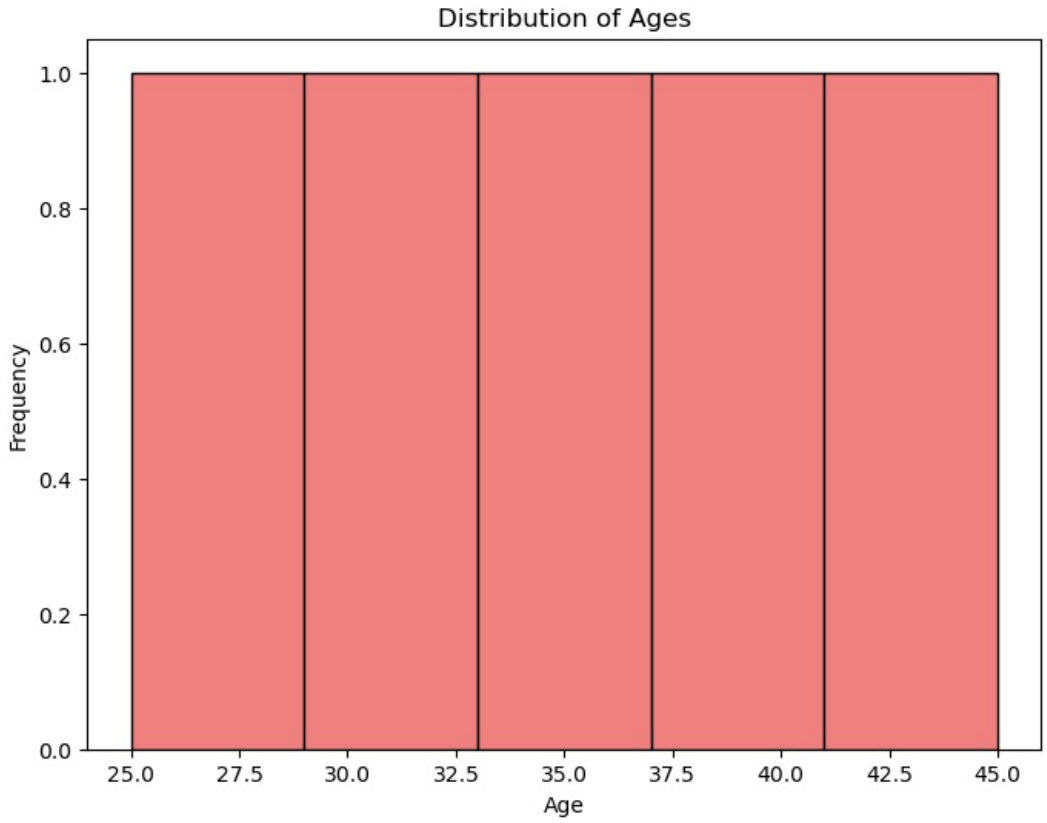
Out[21]:

	Age	Gender
0	25	Male
1	30	Female
2	35	Male
3	40	Female
4	45	Male

In [22]:
```
# Plot a bar chart for the distribution of genders
plt.figure(figsize=(8, 6))
df['Gender'].value_counts().plot(kind='bar', color='skyblue')
plt.title('Distribution of Genders')
plt.xlabel('Gender')
plt.ylabel('Count')
plt.show()
```



In [23]:
```
# Plot a histogram for the distribution of ages
plt.figure(figsize=(8, 6))
plt.hist(df['Age'], bins=5, color='lightcoral', edgecolor='black')
plt.title('Distribution of Ages')
plt.xlabel('Age')
plt.ylabel('Frequency')
plt.show()
```



In [8]:
```
# Load the data
df = pd.read_csv("C:\\Users\\rakhi\\Downloads\\API_SP.POP.TOTL_DS2_en_csv_v2_85\\Metadata_Country_API_SP.POP.TO
df
```

Out[8]:

	Country Code	Region	IncomeGroup	SpecialNotes	TableName	Unnamed: 5
0	ABW	Latin America & Caribbean	High income	NaN	Aruba	NaN
1	AFE	NaN	NaN	26 countries, stretching from the Red Sea in t...	Africa Eastern and Southern	NaN
2	AFG	South Asia	Low income	The reporting period for national accounts dat...	Afghanistan	NaN
3	AFW	NaN	NaN	22 countries, stretching from the westernmost ...	Africa Western and Central	NaN
4	AGO	Sub-Saharan Africa	Lower middle income	The World Bank systematically assesses the app...	Angola	NaN
...	...	...	...	...	...	...
260	XKX	Europe & Central Asia	Upper middle income	NaN	Kosovo	NaN
261	YEM	Middle East & North Africa	Low income	The World Bank systematically assesses the app...	Yemen, Rep.	NaN
262	ZAF	Sub-Saharan Africa	Upper middle income	Fiscal year end: March 31; reporting period fo...	South Africa	NaN
263	ZMB	Sub-Saharan Africa	Lower middle income	National accounts data were rebased to reflect...	Zambia	NaN
264	ZWE	Sub-Saharan Africa	Lower middle income	National Accounts data are reported in Zimbabw...	Zimbabwe	NaN

265 rows × 6 columns

In [10]:
```
df.head()
```

Out[10]:

	Country Code	Region	IncomeGroup	SpecialNotes	TableName	Unnamed: 5
0	ABW	Latin America & Caribbean	High income	NaN	Aruba	NaN
1	AFE	NaN	NaN	26 countries, stretching from the Red Sea in t...	Africa Eastern and Southern	NaN
2	AFG	South Asia	Low income	The reporting period for national accounts dat...	Afghanistan	NaN
3	AFW	NaN	NaN	22 countries, stretching from the westernmost ...	Africa Western and Central	NaN
4	AGO	Sub-Saharan Africa	Lower middle income	The World Bank systematically assesses the app...	Angola	NaN

In [11]:
```
df.tail()
```

Out[11]:

	Country Code	Region	IncomeGroup	SpecialNotes	TableName	Unnamed: 5
260	XKX	Europe & Central Asia	Upper middle income	NaN	Kosovo	NaN
261	YEM	Middle East & North Africa	Low income	The World Bank systematically assesses the app...	Yemen, Rep.	NaN
262	ZAF	Sub-Saharan Africa	Upper middle income	Fiscal year end: March 31; reporting period fo...	South Africa	NaN
263	ZMB	Sub-Saharan Africa	Lower middle income	National accounts data were rebased to reflect...	Zambia	NaN
264	ZWE	Sub-Saharan Africa	Lower middle income	National Accounts data are reported in Zimbabw...	Zimbabwe	NaN

In [12]:
```
df.describe()
```

Out[12]:

	Unnamed: 5
count	0.0
mean	NaN
std	NaN
min	NaN
25%	NaN
50%	NaN
75%	NaN
max	NaN

In [13]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 265 entries, 0 to 264
Data columns (total 6 columns):
 #   Column        Non-Null Count  Dtype  
---  ------        --------------  -----  
 0   Country Code  265 non-null    object 
 1   Region        217 non-null    object 
 2   IncomeGroup   216 non-null    object 
 3   SpecialNotes  126 non-null    object 
 4   TableName     265 non-null    object 
 5   Unnamed: 5    0 non-null      float64
dtypes: float64(1), object(5)
memory usage: 12.6+ KB
```

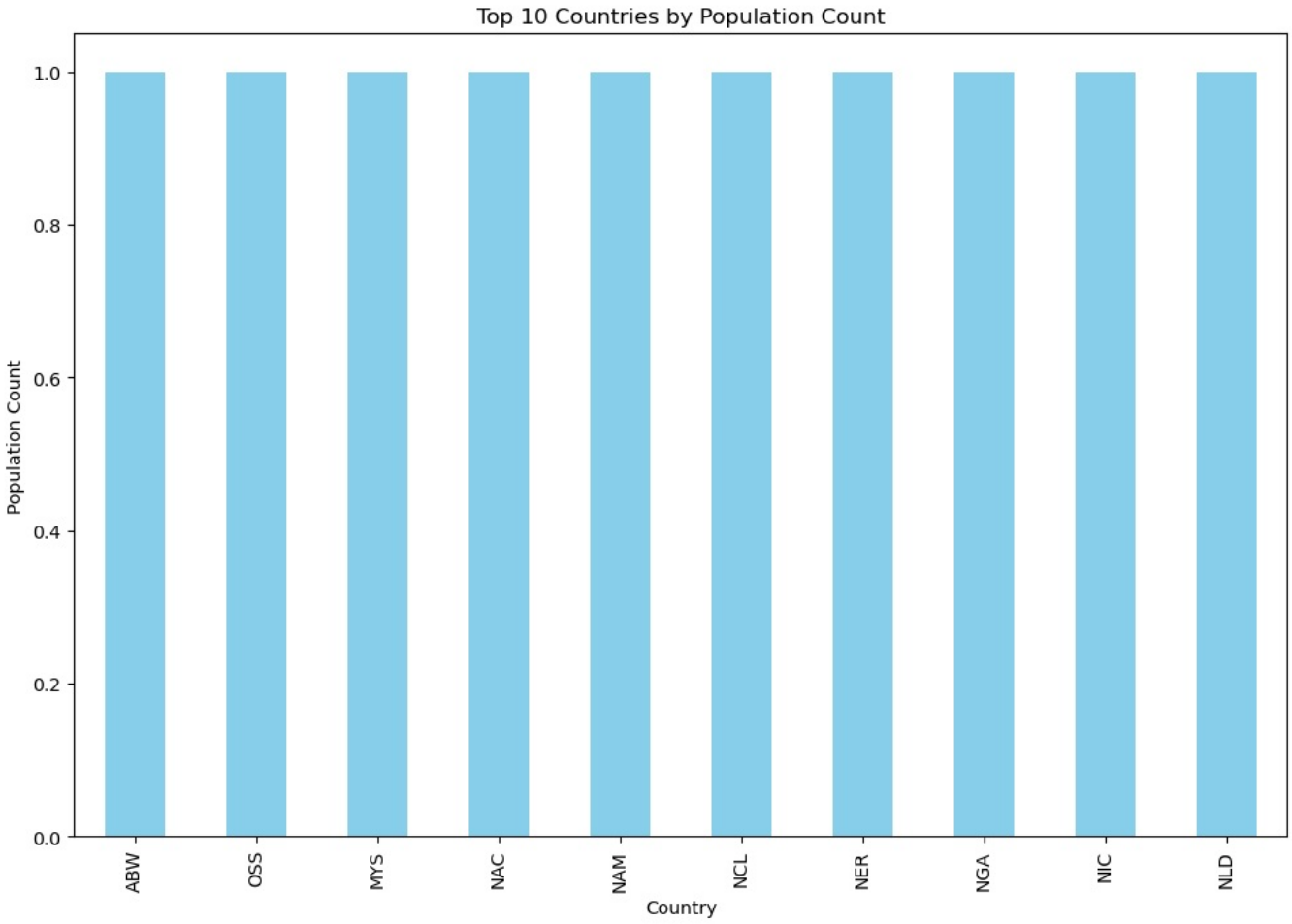
In [14]:
```
df.isnull().sum()
```

Out[14]:
```
Country Code      0
Region           48
IncomeGroup      49
SpecialNotes    139
TableName         0
Unnamed: 5      265
dtype: int64
```

In [16]:
```
df.columns
```

Out[16]:
```
Index(['Country Code', 'Region', 'IncomeGroup', 'SpecialNotes', 'TableName',
       'Unnamed: 5'],
      dtype='object')
```

In [15]:
```
# Plot a bar chart for the distribution of countries
plt.figure(figsize=(12, 8))
df['Country Code'].value_counts().head(10).plot(kind='bar', color='skyblue')
plt.title('Top 10 Countries by Population Count')
plt.xlabel('Country')
plt.ylabel('Population Count')
plt.show()
```



In [17]:
```
# Load the Metadata_Indicator file to examine its contents
metadata_indicator_df = pd.read_csv("C:\\Users\\rakhi\\Downloads\\API_SP.POP.TOTL_DS2_en_csv_v2_85\\Metadata_In
metadata_indicator_df
```

Out[17]:

	INDICATOR_CODE	INDICATOR_NAME	SOURCE_NOTE	SOURCE_ORGANIZATION	Unnamed: 4
0	SP.POP.TOTL	Population, total	Total population is based on the de facto defi...	(1) United Nations Population Division. World ...	NaN

In [18]:
```
# Display the first few rows of the dataframe
display(metadata_indicator_df.head())
```

	INDICATOR_CODE	INDICATOR_NAME	SOURCE_NOTE	SOURCE_ORGANIZATION	Unnamed: 4
0	SP.POP.TOTL	Population, total	Total population is based on the de facto defi...	(1) United Nations Population Division. World ...	NaN

In [19]:
```
metadata_indicator_df.columns = metadata_indicator_df.iloc[0]
metadata_indicator_df = metadata_indicator_df[1:]
metadata_indicator_df
```

Out[19]:

	SP.POP.TOTL	Population, total	Total population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship. The values shown are midyear estimates.	(1) United Nations Population Division. World Population Prospects: 2022 Revision. (2) Census reports and other statistical publications from national statistical offices, (3) Eurostat: Demographic Statistics, (4) United Nations Statistical Division. Population and Vital Statistics Reprot (various years), (5) U.S. Census Bureau: International Database, and (6) Secretariat of the Pacific Community: Statistics and Demography Programme.	NaN

In [ ]: