

```
In [10]: import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
```

```
In [11]: df = pd.read_csv('F:\Git-Repos\PRODIGY_DS_01\DATA_2.csv')
```

```
In [12]: df
```

Out[12]:

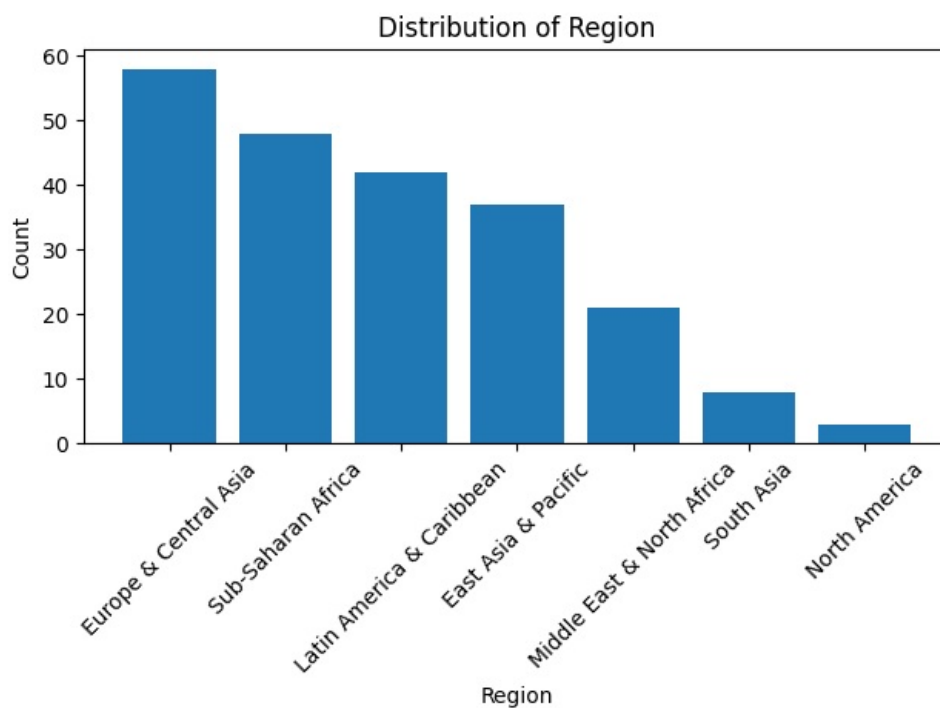
	Country Code	Region	IncomeGroup	SpecialNotes	TableName	Unnamed: 5
0	ABW	Latin America & Caribbean	High income	NaN	Aruba	NaN
1	AFE	NaN	NaN	26 countries, stretching from the Red Sea in t...	Africa Eastern and Southern	NaN
2	AFG	South Asia	Low income	The reporting period for national accounts dat...	Afghanistan	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 [13]: gender_counts = df['Region'].value_counts()
bar_width = 0.9
x=range(len(gender_counts.index))

plt.bar(gender_counts.index,gender_counts.values)
plt.xlabel('Region')
plt.ylabel('Count')
plt.title('Distribution of Region')

plt.xticks(x,gender_counts.index,rotation=45)
plt.tight_layout()
plt.show()
```



```
In [14]: df.shape
```

```
Out[14]: (265, 6)
```

```
In [15]: 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    127 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 [16]: df.describe()
```

```
Out[16]:
```

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

```
In [17]: df.isnull().sum()
```

```
Out[17]: Country Code    0
Region          48
IncomeGroup     49
SpecialNotes    138
TableName       0
Unnamed: 5      265
dtype: int64
```

```
In [18]: 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    127 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
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js