Task 1

Importing dataset and libraries

In [3]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [32]:
```
import pandas as pd
df = pd.read_csv(r"C:\Users\ADMIN\Downloads\Metadata_Country_API_SP.POP.TOTL_DS2
df
```

Out[32]:

	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	Middle East & North Africa	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

Data reading and cleaning

In [8]:
```
df.head(5)
```

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	Middle East & North Africa	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 [9]:
```
df.tail(5)
```

Out[9]:

	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 [10]:
```
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   215 non-null    object 
 3   SpecialNotes  130 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 [11]:
```
df.duplicated().sum()
```

Out[11]:
```
np.int64(0)
```

In [12]:
```
df.isna().sum()
```

Out[12]:
```
Country Code      0
Region           48
IncomeGroup      50
SpecialNotes    135
TableName         0
Unnamed: 5      265
dtype: int64
```

In [27]:
```
df['Region'].fillna('Unknown', inplace=True)
df['IncomeGroup'].fillna('Unknown', inplace=True)
df
```

```
C:\Users\ADMIN\AppData\Local\Temp\ipykernel_12780\331918434.py:1: FutureWarning:
A value is trying to be set on a copy of a DataFrame or Series through chained as
signment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work becau
se the intermediate object on which we are setting values always behaves as a cop
y.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.meth
od({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to pe
rform the operation inplace on the original object.


  df['Region'].fillna('Unknown', inplace=True)
C:\Users\ADMIN\AppData\Local\Temp\ipykernel_12780\331918434.py:2: FutureWarning:
A value is trying to be set on a copy of a DataFrame or Series through chained as
signment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work becau
se the intermediate object on which we are setting values always behaves as a cop
y.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.meth
od({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to pe
rform the operation inplace on the original object.


  df['IncomeGroup'].fillna('Unknown', inplace=True)
```

Out[27]:

	Country Code	Region	IncomeGroup	TableName
0	ABW	Latin America & Caribbean	High income	Aruba
1	AFE	Unknown	Unknown	Africa Eastern and Southern
2	AFG	Middle East & North Africa	Low income	Afghanistan
3	AFW	Unknown	Unknown	Africa Western and Central
4	AGO	Sub-Saharan Africa	Lower middle income	Angola
...	...	...	...	...
260	XKX	Europe & Central Asia	Upper middle income	Kosovo
261	YEM	Middle East & North Africa	Low income	Yemen, Rep.
262	ZAF	Sub-Saharan Africa	Upper middle income	South Africa
263	ZMB	Sub-Saharan Africa	Lower middle income	Zambia
264	ZWE	Sub-Saharan Africa	Lower middle income	Zimbabwe

265 rows × 4 columns

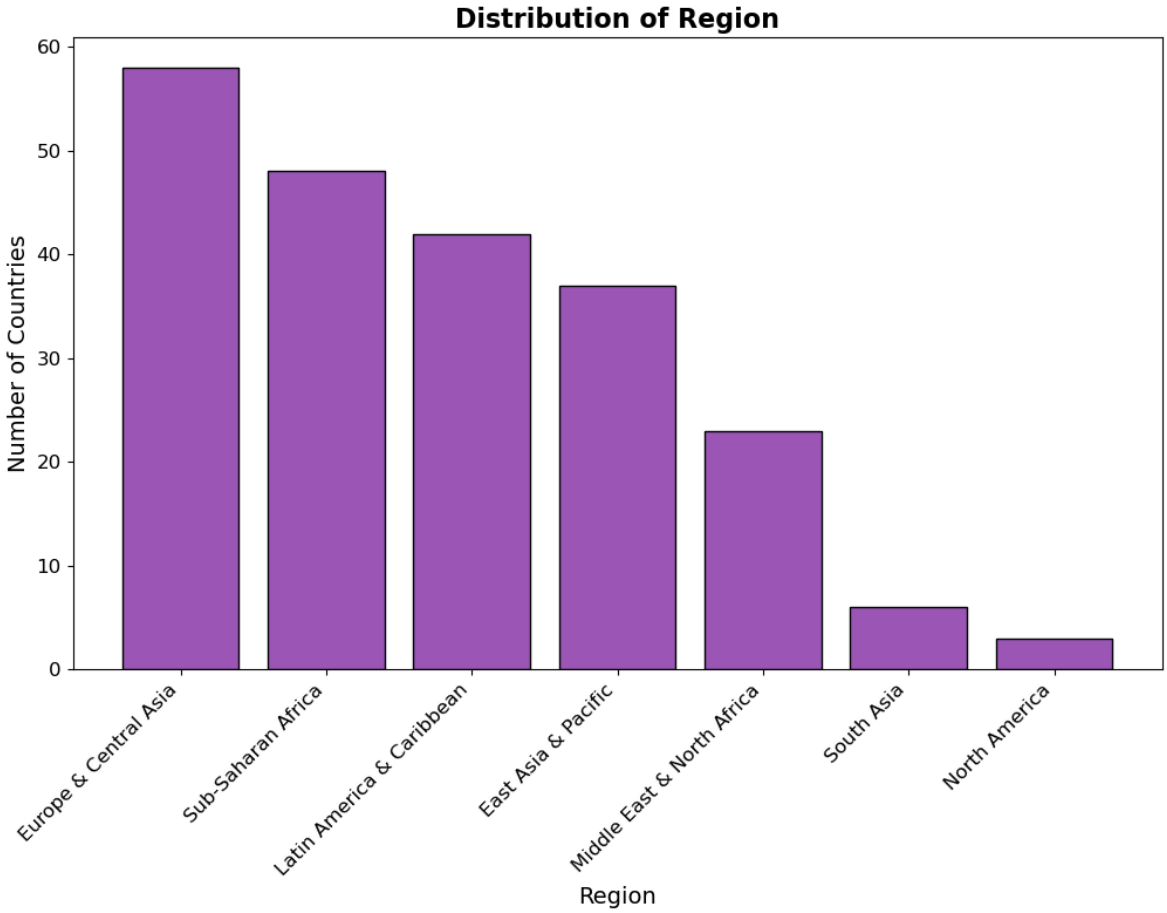
In [30]:

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

# Count how many countries belong to each region
region_counts = df2['Region'].value_counts().sort_values(ascending=False)

# Plot bar chart
plt.figure(figsize=(12,7))
bars = plt.bar(region_counts.index, region_counts.values,
               color='#9B59B6' , edgecolor='black')

plt.title('Distribution of Region', fontsize=16, fontweight='bold')
plt.xlabel('Region', fontsize=14)
plt.ylabel('Number of Countries', fontsize=14)
plt.xticks(rotation=45, ha='right', fontsize=12)
plt.yticks(fontsize=12)
plt.show()
```


Distribution of Region
Number of Countries
0
10
20
30
40
50
60
Europe & Central Asia
Sub-Saharan Africa
Latin America & Caribbean
East Asia & Pacific
Middle East & North Africa
South Asia
North America
Region

In [ ]: