task-1-2

August 3, 2024

[4]:
```
import pandas as pd
import matplotlib.pyplot as plt
```

[6]:
```
data = pd.read_excel('WorldBank Data.xlsx')
data.head()
```

[6]:

	Country Name	Country Code	Region	IncomeGroup	Year	Birth rate, crude (per 1,000 people)	Death rate, crude (per 1,000 people)	Electric power consumption (kWh per capita)	GDP (USD)	GDP per capita (USD)	Individuals using the Internet (% of population)	Infant mortality rate (per 1,000 live births)
0	Afghanistan	AFG	South Asia	Low income	2018	NaN	NaN	NaN	1.936300e+10	520.897	NaN	47.9
1	Afghanistan	AFG	South Asia	Low income	2017	33.211	6.575	NaN	2.019180e+10	556.302	13.50	49.5
2	Afghanistan	AFG	South Asia	Low income	2016	33.981	6.742	NaN	1.936260e+10	547.228	11.20	
3	Afghanistan	AFG	South Asia	Low income	2015	34.809	6.929	NaN	1.990710e+10	578.466	8.26	
4	Afghanistan	AFG	South Asia	Low income	2014	35.706	7.141	NaN	2.048490e+10	613.856	7.00	

```
2                                                 51.2
3                                                 53.1
4                                                 55.1

   Life expectancy at birth (years)  \
0                                NaN
1                             64.130
2                             63.763
3                             63.377
4                             62.966

   Population density (people per sq. km of land area)  \
0                                            56.9378
1                                            55.5960
2                                            54.1971
3                                            52.7121
4                                            51.1148

   Unemployment (% of total labor force) (modeled ILO estimate)
0                                              1.542
1                                              1.559
2                                              1.634
3                                              1.679
4                                              1.735
```

[8]:
```
data.tail()
```

[8]:
```
       Country Name Country Code             Region IncomeGroup  Year  \
12444     Zimbabwe          ZWE  Sub-Saharan Africa  Low income  1964
12445     Zimbabwe          ZWE  Sub-Saharan Africa  Low income  1963
12446     Zimbabwe          ZWE  Sub-Saharan Africa  Low income  1962
12447     Zimbabwe          ZWE  Sub-Saharan Africa  Low income  1961
12448     Zimbabwe          ZWE  Sub-Saharan Africa  Low income  1960

       Birth rate, crude (per 1,000 people)  \
12444                                 47.770
12445                                 47.876
12446                                 47.950
12447                                 47.988
12448                                 47.996

       Death rate, crude (per 1,000 people)  \
12444                                 13.083
12445                                 13.419
12446                                 13.762
12447                                 14.104
12448                                 14.441
```

```
       Electric power consumption (kWh per capita)     GDP (USD)  \
12444                                          NaN  1.217138e+09
12445                                          NaN  1.159512e+09
12446                                          NaN  1.117602e+09
12447                                          NaN  1.096647e+09
12448                                          NaN  1.052990e+09

       GDP per capita (USD)  Individuals using the Internet (% of population)  \
12444               281.558                                                NaN
12445               277.480                                                NaN
12446               276.689                                                NaN
12447               280.829                                                NaN
12448               278.814                                                NaN

       Infant mortality rate (per 1,000 live births)  \
12444                                           83.2
12445                                           85.7
12446                                           88.1
12447                                           90.5
12448                                           92.8

       Life expectancy at birth (years)  \
12444                            54.849
12445                            54.403
12446                            53.946
12447                            53.483
12448                            53.019

       Population density (people per sq. km of land area)  \
12444                                            11.1745
12445                                            10.8019
12446                                            10.4413
12447                                            10.0944
12448                                                NaN

       Unemployment (% of total labor force) (modeled ILO estimate)
12444                                                NaN
12445                                                NaN
12446                                                NaN
12447                                                NaN
12448                                                NaN
```

[10]:
```
data.columns
```

[10]:
```
Index(['Country Name', 'Country Code', 'Region', 'IncomeGroup', 'Year',
       'Birth rate, crude (per 1,000 people)',
```

```
       'Death rate, crude (per 1,000 people)',
       'Electric power consumption (kWh per capita)', 'GDP (USD)',
       'GDP per capita (USD)',
       'Individuals using the Internet (% of population)',
       'Infant mortality rate (per 1,000 live births)',
       'Life expectancy at birth (years)',
       'Population density (people per sq. km of land area)',
       'Unemployment (% of total labor force) (modeled ILO estimate)'],
      dtype='object')
```

[12]:
```
data.describe()
```

[12]:
```
              Year  Birth rate, crude (per 1,000 people)  \
count  12449.00000                          11440.000000
mean    1989.00000                             28.643276
std       17.03007                             13.131893
min     1960.00000                              6.900000
25%     1974.00000                             16.600000
50%     1989.00000                             27.545500
75%     2004.00000                             40.881250
max     2018.00000                             58.227000

       Death rate, crude (per 1,000 people)  \
count                          11416.000000
mean                              10.588539
std                                5.489382
min                                1.127000
25%                                6.863750
50%                                9.200000
75%                               12.687000
max                               54.444000

       Electric power consumption (kWh per capita)     GDP (USD)  \
count                                  5848.000000  9.578000e+03
mean                                   3175.294686  1.700740e+11
std                                    4467.139298  8.979866e+11
min                                       0.000000  8.824450e+06
25%                                     390.385750  1.393010e+09
50%                                    1541.895000  7.275305e+09
75%                                    4313.767500  4.857782e+10
max                                   54799.200000  2.050000e+13

       GDP per capita (USD)  Individuals using the Internet (% of population)  \
count           9575.000000                                        5064.000000
mean            8231.812259                                          23.334471
std            16173.539954                                          28.319388
min               34.790600                                           0.000000
```

```
25%              513.145500                                                0.594949
50%             1852.810000                                                8.406225
75%             7774.565000                                               41.295950
max           189171.000000                                              100.000000

       Infant mortality rate (per 1,000 live births)  \
count                                     9984.000000
mean                                        51.704437
std                                         46.131039
min                                          1.400000
25%                                         14.475000
50%                                         37.000000
75%                                         78.200000
max                                        279.400000

       Life expectancy at birth (years)  \
count                      11176.000000
mean                          64.044692
std                           11.491087
min                           18.907000
25%                           55.917750
50%                           67.276000
75%                           72.692250
max                           85.417100

       Population density (people per sq. km of land area)  \
count                                       11845.000000
mean                                          318.861370
std                                          1593.406041
min                                             0.098625
25%                                            19.783400
50%                                            64.007500
75%                                           144.823000
max                                         21389.100000

       Unemployment (% of total labor force) (modeled ILO estimate)
count                                        5208.000000
mean                                            8.295079
std                                             6.290703
min                                             0.140000
25%                                             3.687000
50%                                             6.775000
75%                                            11.212250
max                                            37.940000
```

[14]:
```
data1=data.dropna()
data1.head()
```

[14]:

	Country Name	Country Code	Region	IncomeGroup
63	Albania	ALB	Europe & Central Asia	Upper middle income
64	Albania	ALB	Europe & Central Asia	Upper middle income
65	Albania	ALB	Europe & Central Asia	Upper middle income
66	Albania	ALB	Europe & Central Asia	Upper middle income
67	Albania	ALB	Europe & Central Asia	Upper middle income

	Year	Birth rate, crude (per 1,000 people)
63	2014	12.259
64	2013	12.257
65	2012	12.197
66	2011	12.100
67	2010	12.001

	Death rate, crude (per 1,000 people)
63	7.219
64	7.096
65	6.996
66	6.915
67	6.841

	Electric power consumption (kWh per capita)	GDP (USD)
63	2309.37	1.322820e+10
64	2533.25	1.277630e+10
65	2118.33	1.231980e+10
66	2205.70	1.289090e+10
67	1943.34	1.192700e+10

	GDP per capita (USD)	Individuals using the Internet (% of population)
63	4578.67	60.100
64	4413.08	57.200
65	4247.61	54.656
66	4437.18	49.000
67	4094.36	45.000

	Infant mortality rate (per 1,000 live births)
63	8.9
64	9.5
65	10.2
66	11.0
67	11.9

	Life expectancy at birth (years)
63	77.813
64	77.554
65	77.252
66	76.914

```
67                          76.562

    Population density (people per sq. km of land area)  \
63                                             105.442
64                                             105.660
65                                             105.854
66                                             106.029
67                                             106.315

    Unemployment (% of total labor force) (modeled ILO estimate)
63                                              17.490
64                                              15.866
65                                              13.376
66                                              13.481
67                                              14.086
```

[16]:
```
data1.isna().sum()
```

[16]:
```
Country Name                                                    0
Country Code                                                    0
Region                                                          0
IncomeGroup                                                     0
Year                                                            0
Birth rate, crude (per 1,000 people)                            0
Death rate, crude (per 1,000 people)                            0
Electric power consumption (kWh per capita)                     0
GDP (USD)                                                       0
GDP per capita (USD)                                            0
Individuals using the Internet (% of population)                0
Infant mortality rate (per 1,000 live births)                   0
Life expectancy at birth (years)                                0
Population density (people per sq. km of land area)             0
Unemployment (% of total labor force) (modeled ILO estimate)    0
dtype: int64
```

[18]:
```
data1.info()
```

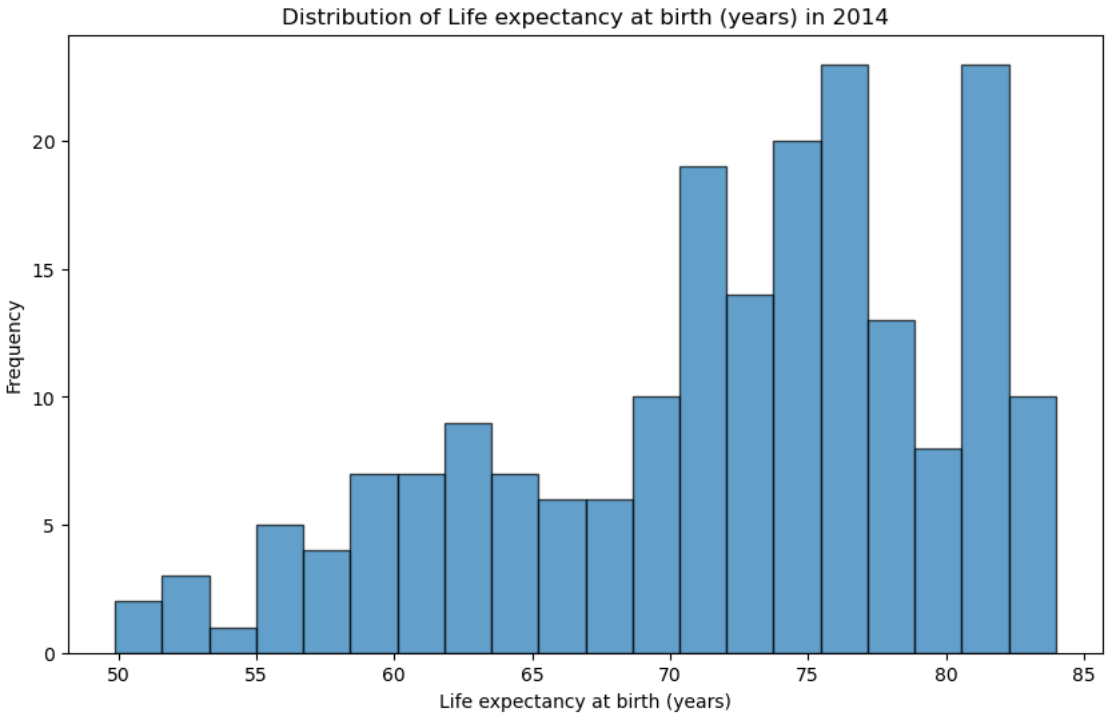
```
<class 'pandas.core.frame.DataFrame'>
Index: 2775 entries, 63 to 12414
Data columns (total 15 columns):
 #   Column                                                      Non-Null
Count  Dtype
---  ------
--------------  -----
 0   Country Name                                                2775 non-null
object
 1   Country Code                                                2775 non-null
```

```
object
 2   Region                                                    2775 non-null
object
 3   IncomeGroup                                               2775 non-null
object
 4   Year                                                      2775 non-null
int64
 5   Birth rate, crude (per 1,000 people)                      2775 non-null
float64
 6   Death rate, crude (per 1,000 people)                      2775 non-null
float64
 7   Electric power consumption (kWh per capita)               2775 non-null
float64
 8   GDP (USD)                                                 2775 non-null
float64
 9   GDP per capita (USD)                                      2775 non-null
float64
 10  Individuals using the Internet (% of population)          2775 non-null
float64
 11  Infant mortality rate (per 1,000 live births)             2775 non-null
float64
 12  Life expectancy at birth (years)                          2775 non-null
float64
 13  Population density (people per sq. km of land area)       2775 non-null
float64
 14  Unemployment (% of total labor force) (modeled ILO estimate)  2775 non-null
float64
dtypes: float64(10), int64(1), object(4)
memory usage: 346.9+ KB
```

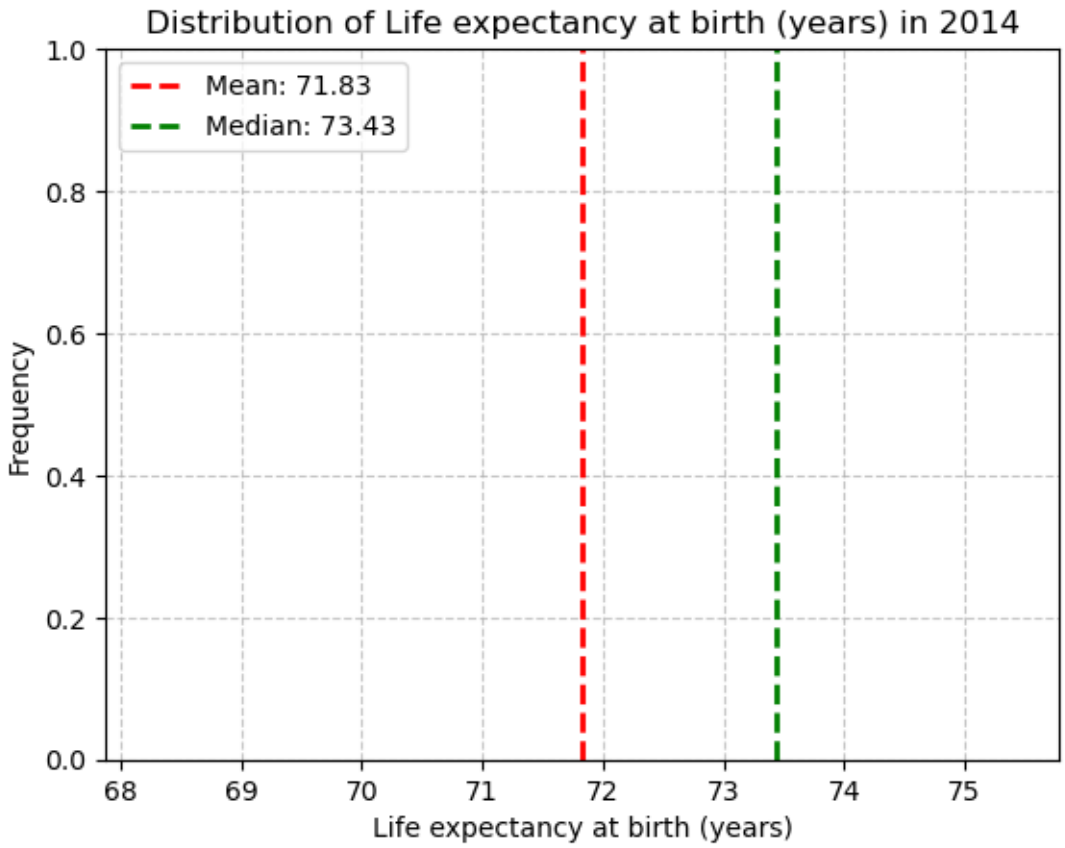
[20]:
```
year = 2014
variable = 'Life expectancy at birth (years)'
```

[22]:
```
data_year = data[data['Year'] == year]
```

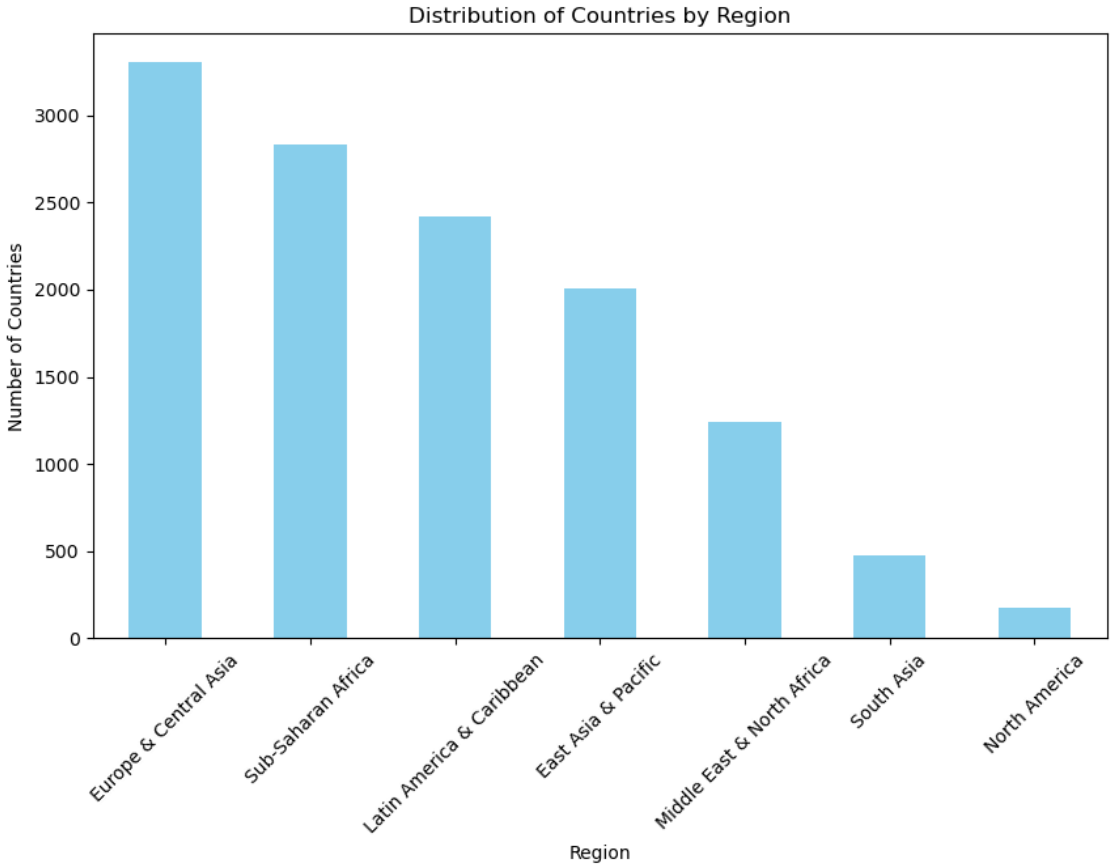
[24]:
```
plt.figure(figsize=(10, 6))
plt.hist(data_year[variable], bins=20, edgecolor='k', alpha=0.7)

plt.title(f'Distribution of {variable} in {year}')
plt.xlabel(variable)
plt.ylabel('Frequency')
plt.show()
```



[26]:
```
plt.title(f'Distribution of {variable} in {year}')
plt.xlabel(variable)
plt.ylabel('Frequency')
plt.grid(True, linestyle='--', alpha=0.7)
mean_value = data_year[variable].mean()
median_value = data_year[variable].median()
plt.axvline(mean_value, color='red', linestyle='dashed', linewidth=2,␣
 ↪label=f'Mean: {mean_value:.2f}')
plt.axvline(median_value, color='green', linestyle='dashed', linewidth=2,␣
 ↪label=f'Median: {median_value:.2f}')
plt.legend()
plt.show()
```



[28]:
```
region_counts = data['Region'].value_counts()
plt.figure(figsize=(10, 6))
region_counts.plot(kind='bar', color='skyblue')
plt.xlabel('Region')
plt.ylabel('Number of Countries')
plt.title('Distribution of Countries by Region')
plt.xticks(rotation=45)
plt.show()
```


Distribution of Countries by Region
Number of Countries
0
500
1000
1500
2000
2500
3000
Europe & Central Asia
Sub-Saharan Africa
Latin America & Caribbean
East Asia & Pacific
Middle East & North Africa
South Asia
North America
Region

[ ]: