In [4]:
```
from urllib.request import urlretrieve
import pandas as pd
urlretrieve('https://gist.githubusercontent.com/aakashns/28b2e504b3350afd9bdb
countries_df = pd.read_csv('countries.csv')
countries_df
```

	location	continent	population	life_expectancy	hospital_beds_per_thousand	gdp_p
0	Afghanistan	Asia	38928341.0	64.83	0.50	
1	Albania	Europe	2877800.0	78.57	2.89	
2	Algeria	Africa	43851043.0	76.88	1.90	
3	Andorra	Europe	77265.0	83.73	NaN	
4	Angola	Africa	32866268.0	61.15	NaN	
...	...	...	...	...	...	
205	Vietnam	Asia	97338583.0	75.40	2.60	
206	Western Sahara	Africa	597330.0	70.26	NaN	
207	Yemen	Asia	29825968.0	66.12	0.70	
208	Zambia	Africa	18383956.0	63.89	2.00	
209	Zimbabwe	Africa	14862927.0	61.49	1.70	

Q1: How many countries does the dataframe contain?

In [5]:
```
len(countries_df['location'])
```

Out[5]: 210

Q2 : Retrieve a list of continents from the dataframe?

In [6]:
```
countries_df['continent']
```

Out[6]:
```
0        Asia
1      Europe
2      Africa
3      Europe
4      Africa
        ...
205      Asia
206    Africa
207      Asia
208    Africa
209    Africa
Name: continent, Length: 210, dtype: object
```

Q3: What is the total population of all the countries listed in this dataset?

In [7]:
```
countries_df['population'].sum()
```

Out[7]: 7757980095.0

Q4: What is the overall life expectancy across the world?

In [8]:
```
(countries_df['life_expectancy']*countries_df['population']).sum() / countrie
```

Out[8]: 72.74135064806133

Q5: Create a dataframe containing 10 countries with the highest population.
(Hint: Chain thesort_values and head methods.)

In [11]:
```
countries_df.sort_values('population', ascending=False).head(10)
```

Out[11]:

	location	continent	population	life_expectancy	hospital_beds_per_thousand	gdp_pe
41	China	Asia	1.439324e+09	76.91	4.34	1
90	India	Asia	1.380004e+09	69.66	0.53	
199	United States	North America	3.310026e+08	78.86	2.77	5
91	Indonesia	Asia	2.735236e+08	71.72	1.04	1
145	Pakistan	Asia	2.208923e+08	67.27	0.60	
27	Brazil	South America	2.125594e+08	75.88	2.20	1
141	Nigeria	Africa	2.061396e+08	54.69	NaN	
15	Bangladesh	Asia	1.646894e+08	72.59	0.80	
157	Russia	Europe	1.459345e+08	72.58	8.05	2
125	Mexico	North America	1.289328e+08	75.05	1.38	1

Q6: Add a new column in countries_df to record the overall GDP per country
(product ofpopulation & per capita GDP).

In [17]:
```
countries_df['Overall_GDP'] = countries_df['population']*countries_df['gdp_pe
countries_df
```

Out[17]:

	location	continent	population	life_expectancy	hospital_beds_per_thousand	gdp_per_
0	Afghanistan	Asia	38928341.0	64.83	0.50	18
1	Albania	Europe	2877800.0	78.57	2.89	118
2	Algeria	Africa	43851043.0	76.88	1.90	139
3	Andorra	Europe	77265.0	83.73	NaN	
4	Angola	Africa	32866268.0	61.15	NaN	58
...	...	...	...	...	...	
205	Vietnam	Asia	97338583.0	75.40	2.60	61
206	Western Sahara	Africa	597330.0	70.26	NaN	
207	Yemen	Asia	29825968.0	66.12	0.70	14
208	Zambia	Africa	18383956.0	63.89	2.00	36
209	Zimbabwe	Africa	14862927.0	61.49	1.70	18

210 rows × 7 columns

Q7: Create a dataframe containing 10 countries with the lowest GDP per capita, among thecounties with population greater than 100 million.

In [20]:
```
lowest_gdp_countries = countries_df[countries_df.population > 100000000].sort
lowest_gdp_countries
```

Out[20]:

	location	continent	population	life_expectancy	hospital_beds_per_thousand	gdp_pe
63	Ethiopia	Africa	1.149636e+08	66.60	0.30	
15	Bangladesh	Asia	1.646894e+08	72.59	0.80	
145	Pakistan	Asia	2.208923e+08	67.27	0.60	
141	Nigeria	Africa	2.061396e+08	54.69	NaN	
90	India	Asia	1.380004e+09	69.66	0.53	
151	Philippines	Asia	1.095811e+08	71.23	1.00	
58	Egypt	Africa	1.023344e+08	71.99	1.60	1
91	Indonesia	Asia	2.735236e+08	71.72	1.04	1
27	Brazil	South America	2.125594e+08	75.88	2.20	1
41	China	Asia	1.439324e+09	76.91	4.34	1

```
# THE END #
```