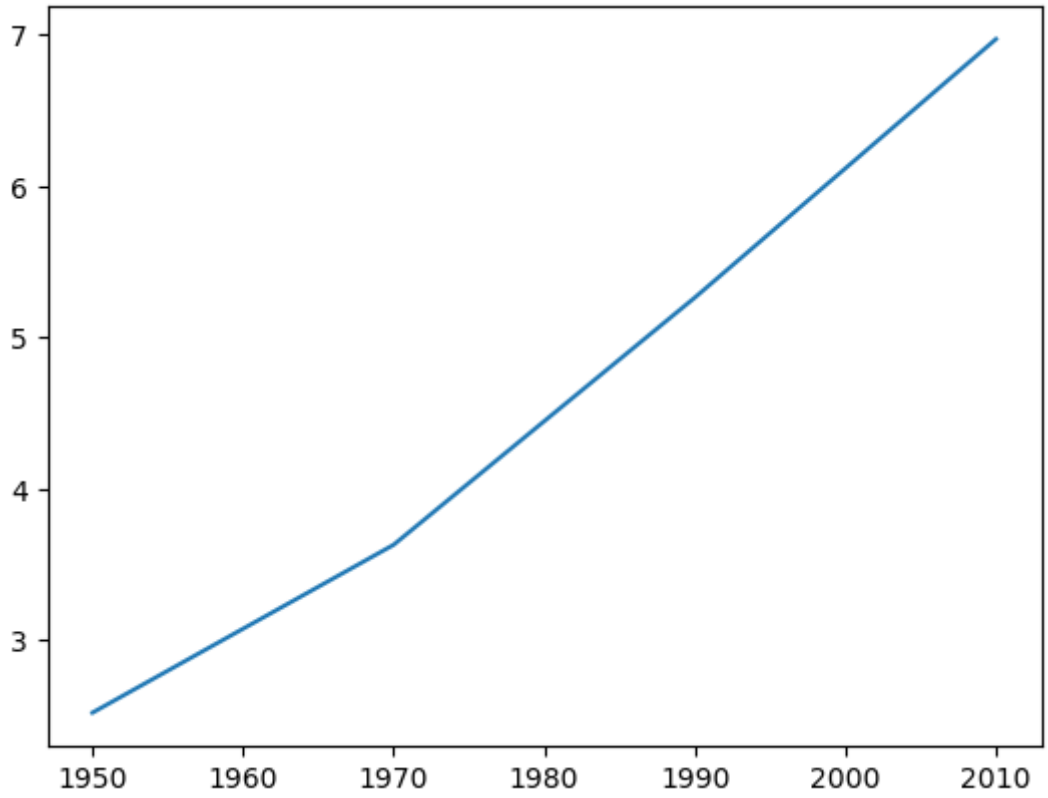
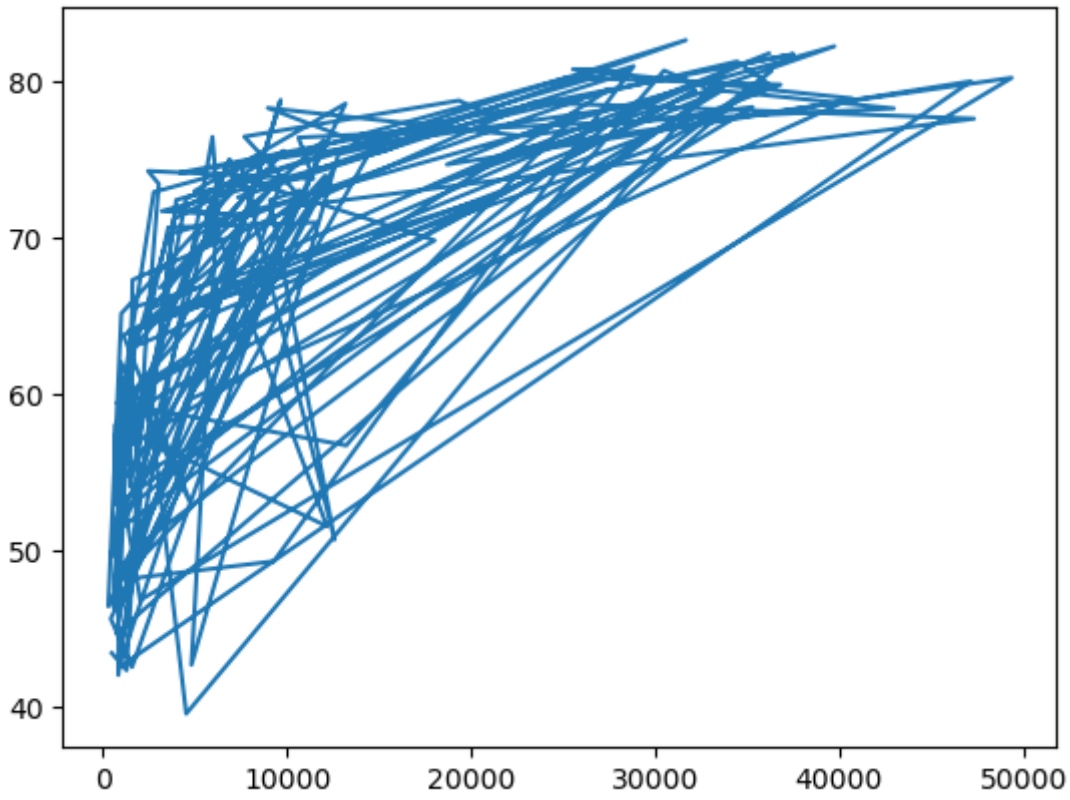
In [1]:
```
import pandas as pd
```

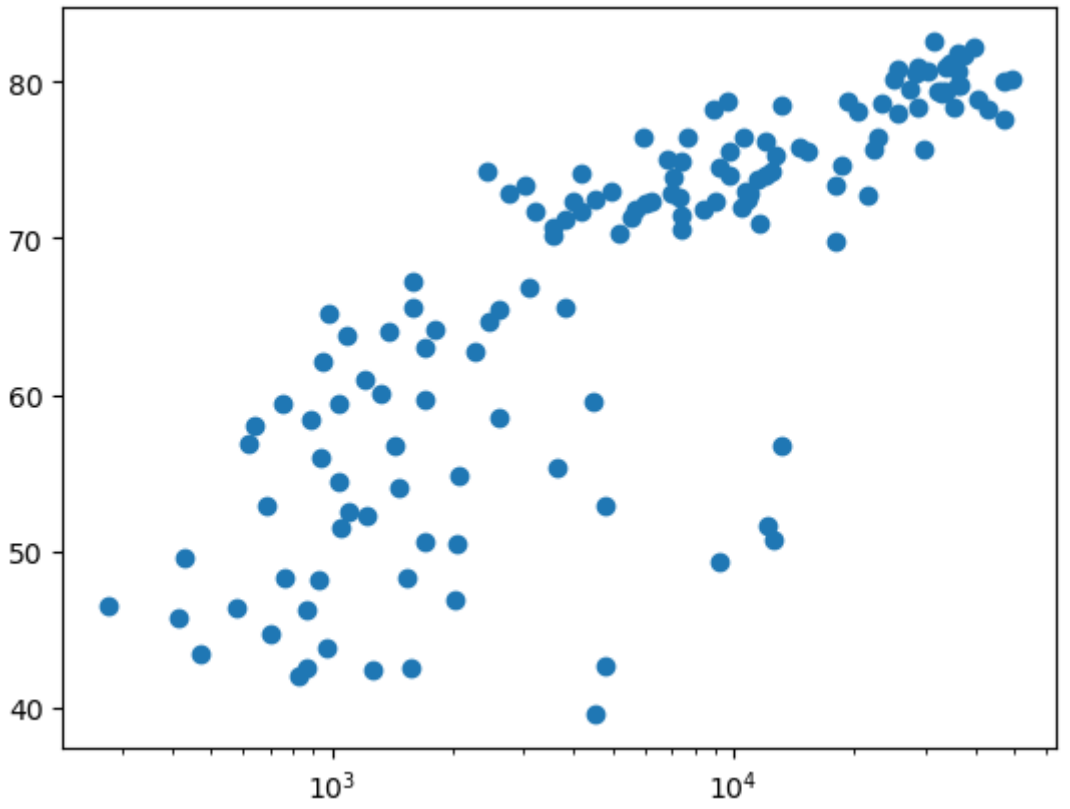
In [3]:
```
import matplotlib.pyplot as plt
%matplotlib inline
year = [1950, 1970, 1990, 2010]
pop = [2.519, 3.629, 5.263, 6.972 ]
plt.plot(year, pop)
plt.show()
```



In [14]:
```
gdp_cap =  [974.5803384, 5937.029525999999, 6223.367465, 4797.231267, 12779.
life_exp = [43.828, 76.423, 72.301, 42.731, 75.32, 81.235, 79.829, 75.635, 6
plt.plot(gdp_cap, life_exp)
plt.show()
```

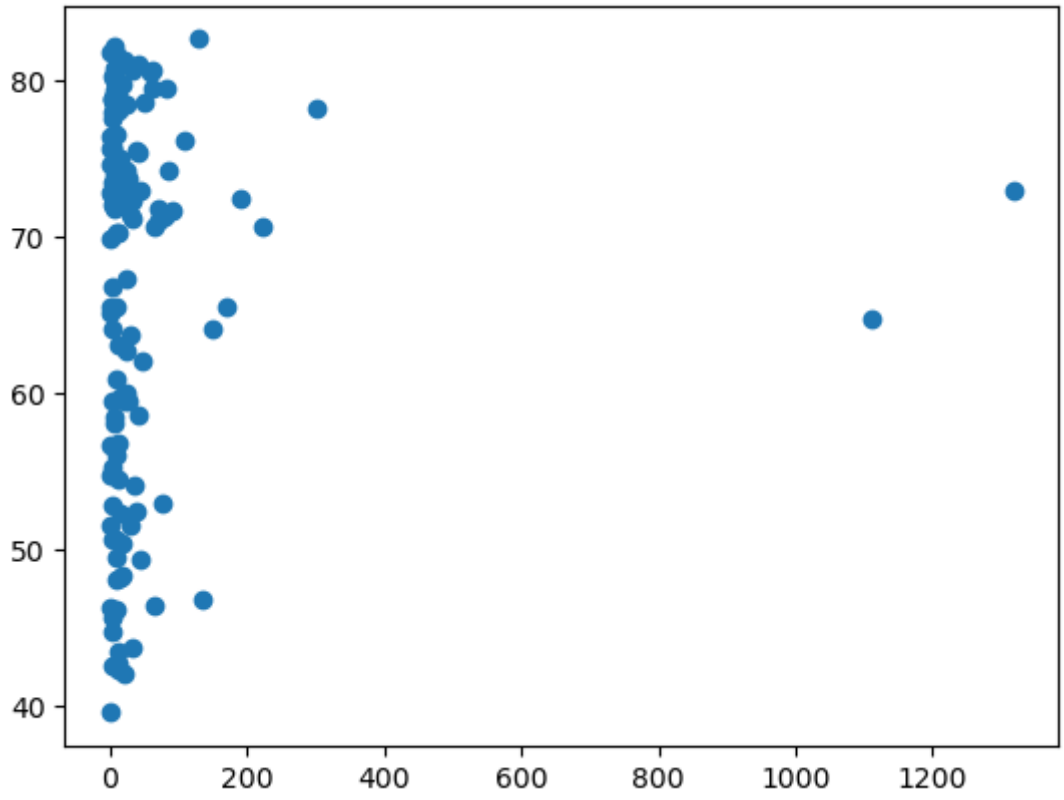


In [12]:
```
gdp_cap =  [974.5803384, 5937.029525999999, 6223.367465, 4797.231267, 12779.
life_exp = [43.828, 76.423, 72.301, 42.731, 75.32, 81.235, 79.829, 75.635, 6
plt.scatter(gdp_cap, life_exp)
plt.xscale('log') #log scale
plt.show()
```



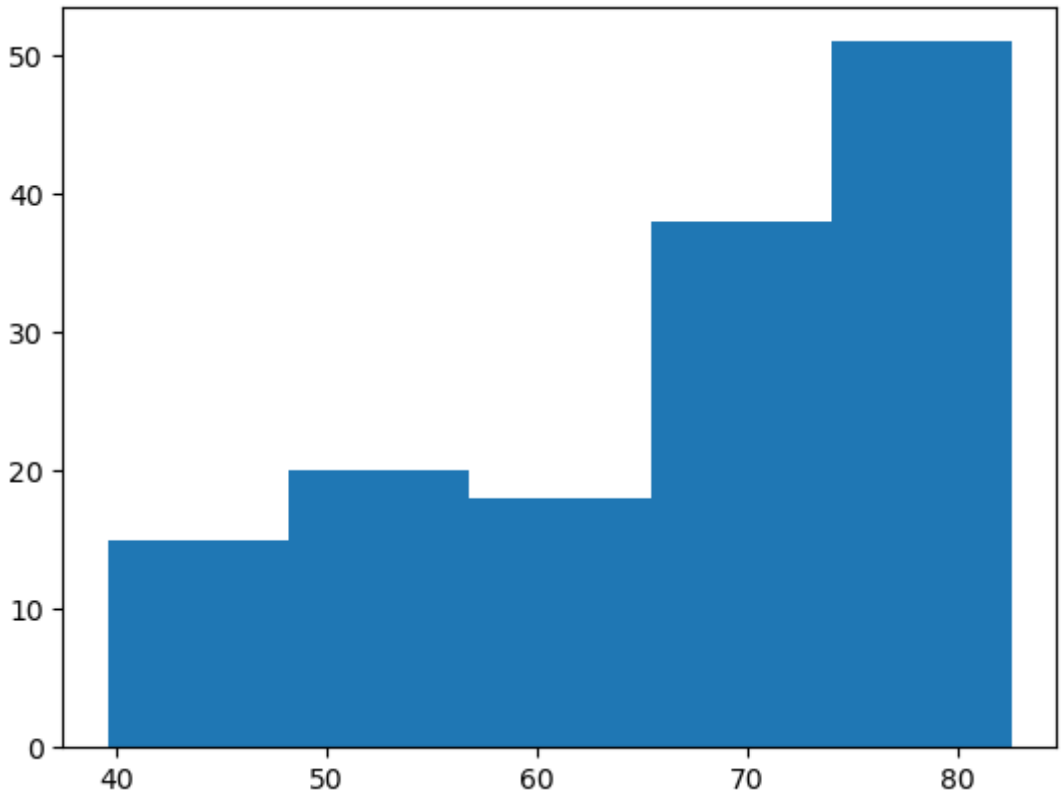
Exercises 1

In [36]:
```
#Exercise 1
pop = [31.889923, 3.600523, 33.333216, 12.420476, 40.301927, 20.434176, 8.19
life_exp = [43.828, 76.423, 72.301, 42.731, 75.32, 81.235, 79.829, 75.635, 6
plt.scatter(pop, life_exp)
plt.show()
```

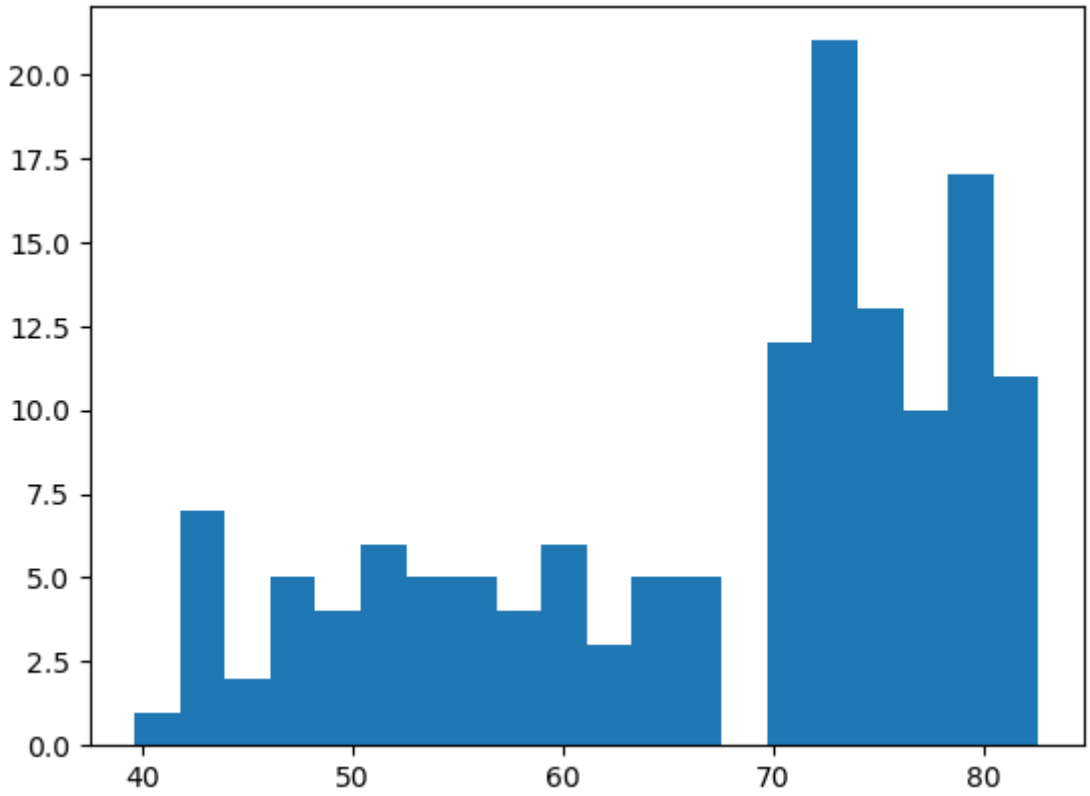


Exercise 2

In [21]:
```
life_exp = [43.828, 76.423, 72.301, 42.731, 75.32, 81.235, 79.829, 75.635, 6
plt.hist(life_exp,bins=5)
plt.show()
```



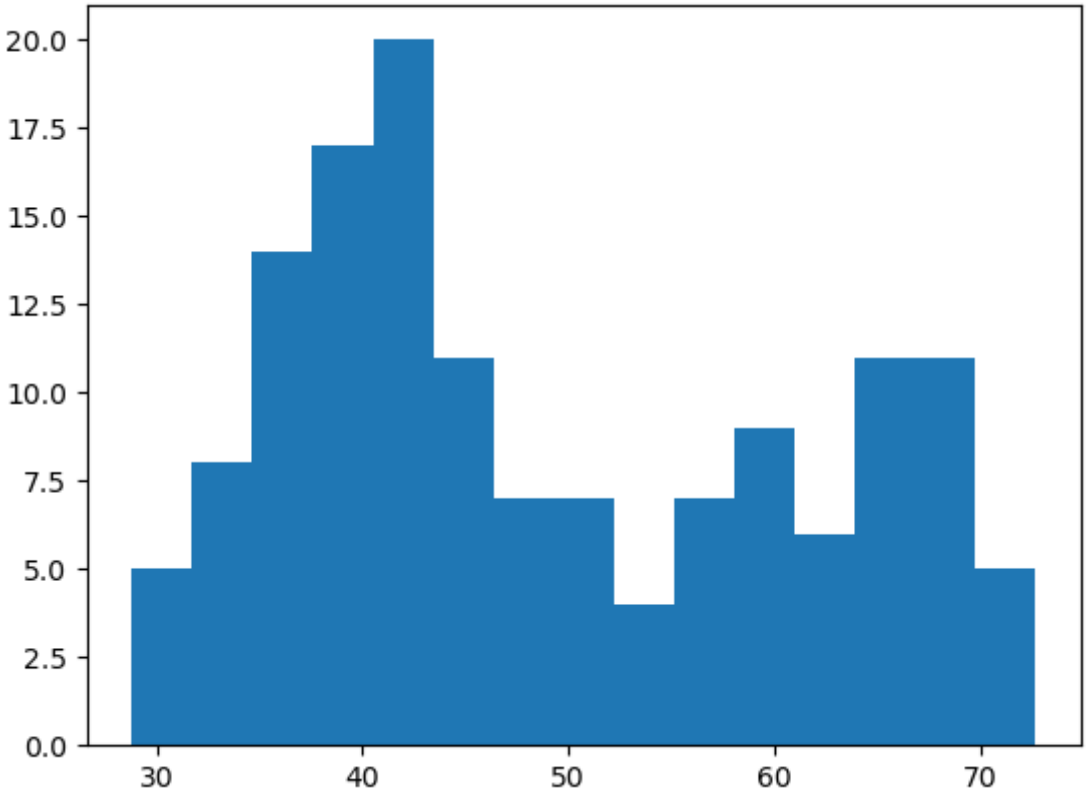
In [25]:
```
life_exp = [43.828, 76.423, 72.301, 42.731, 75.32, 81.235, 79.829, 75.635, 6
plt.hist(life_exp,bins=20)
plt.show()
```



Exercise 3 Ans

In [29]:
```
#Exercise 3
life_exp1950 = [28.8, 55.23, 43.08, 30.02, 62.48, 69.12, 66.8, 50.94, 37.48,
```

```
plt.hist(life_exp1950,bins=15)
plt.show()
```



The histogram shows that the life expectancy in 2007 is much higher than in 1950. The bin with the highest frequency in the histogram for life_exp1950 is the bin from 40 to 45 years old, with a frequency of 35. The bin with the highest frequency in the histogram for life_exp is the bin from 70 to 75 years old, with a frequency of 43. This shows that life expectancy has increased significantly since 1950.

Exercise 4 Ans

To visually assess if the grades on your exam follow a particular distribution, you would typically use a histogram plot. A histogram provides a visual representation of the distribution of a dataset by dividing it into bins and displaying the frequency or count of values within each bin.

Exercise 5

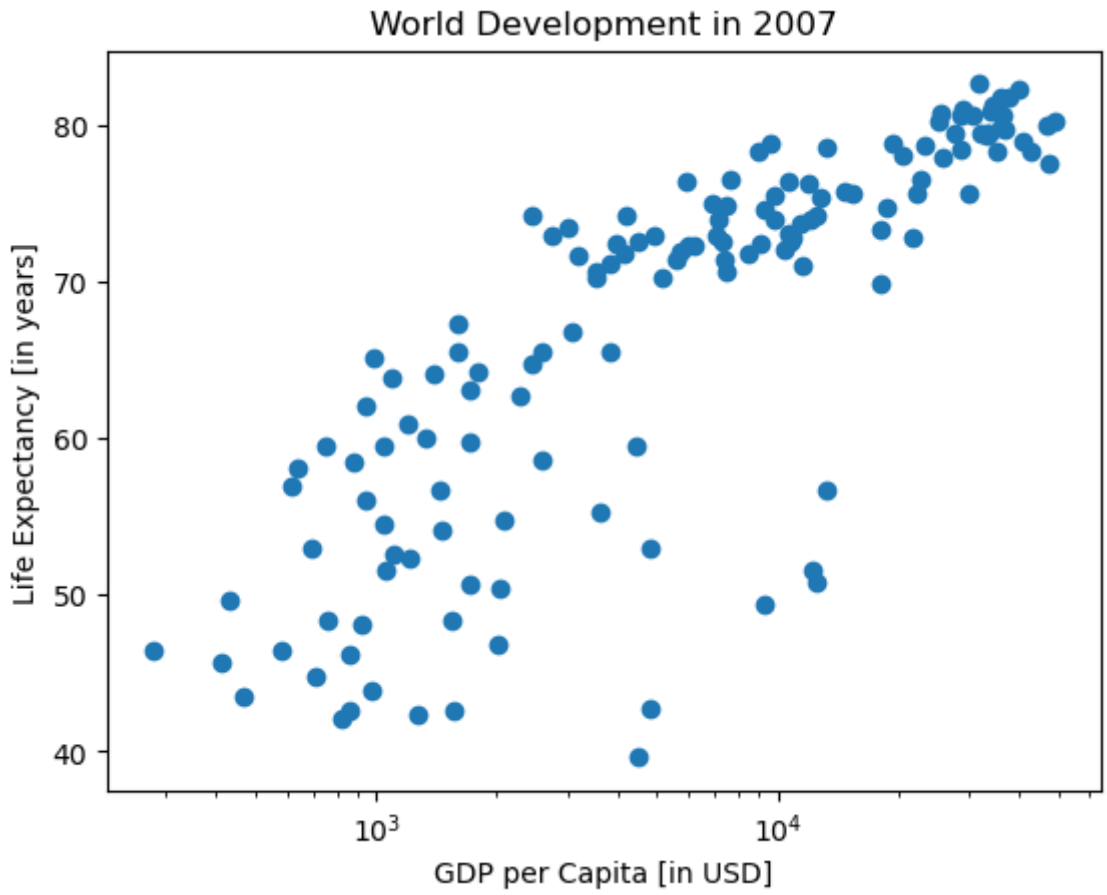
To visually assess if longer answers on exam questions lead to higher grades, you would typically use a scatter plot. A scatter plot is suitable for comparing two continuous variables and allows you to examine the relationship or correlation between them.

Exercise 6

In [31]:
```
# Basic scatter plot, log scale
plt.scatter(gdp_cap, life_exp)
```

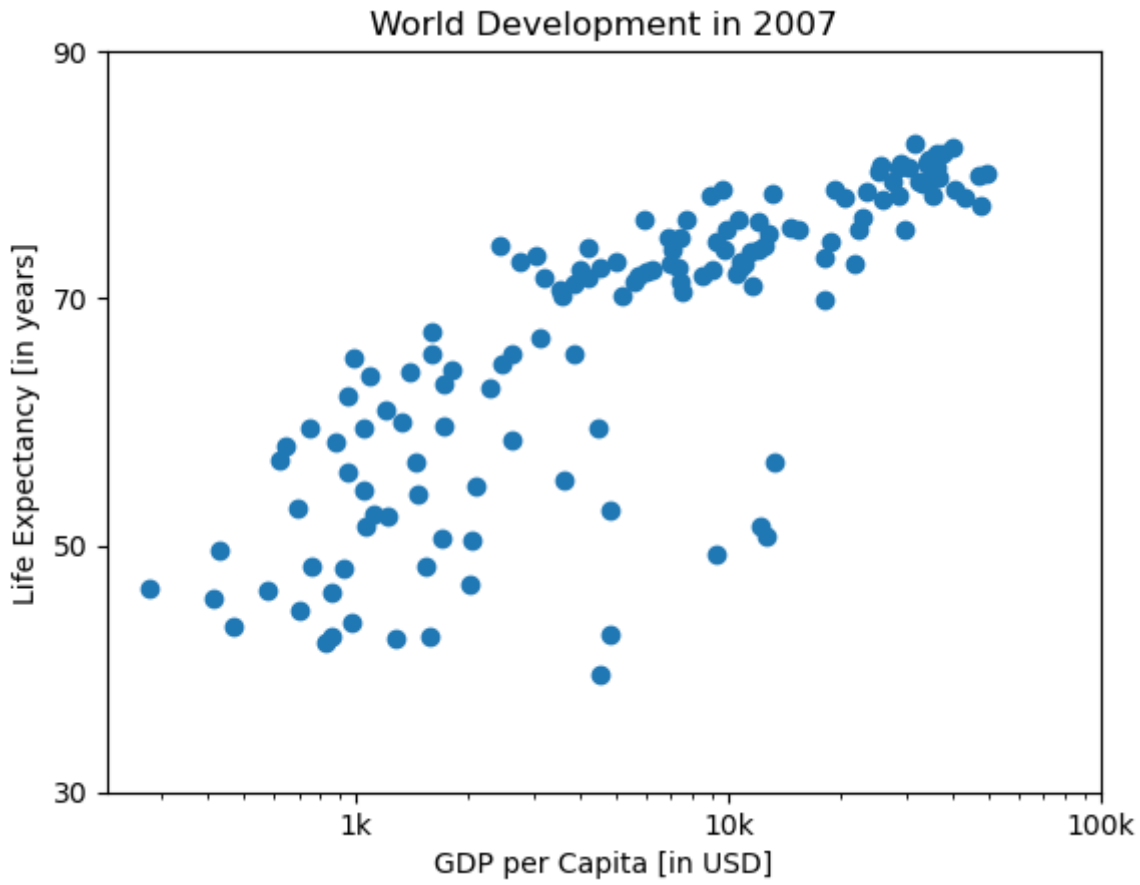
```
plt.xscale('log')
# Strings
xlab = 'GDP per Capita [in USD]'
ylab = 'Life Expectancy [in years]'
title = 'World Development in 2007'
# Add axis labels
plt.xlabel(xlab)
plt.ylabel(ylab)
plt.title(title)
```

Out[31]:
```
Text(0.5, 1.0, 'World Development in 2007')
```

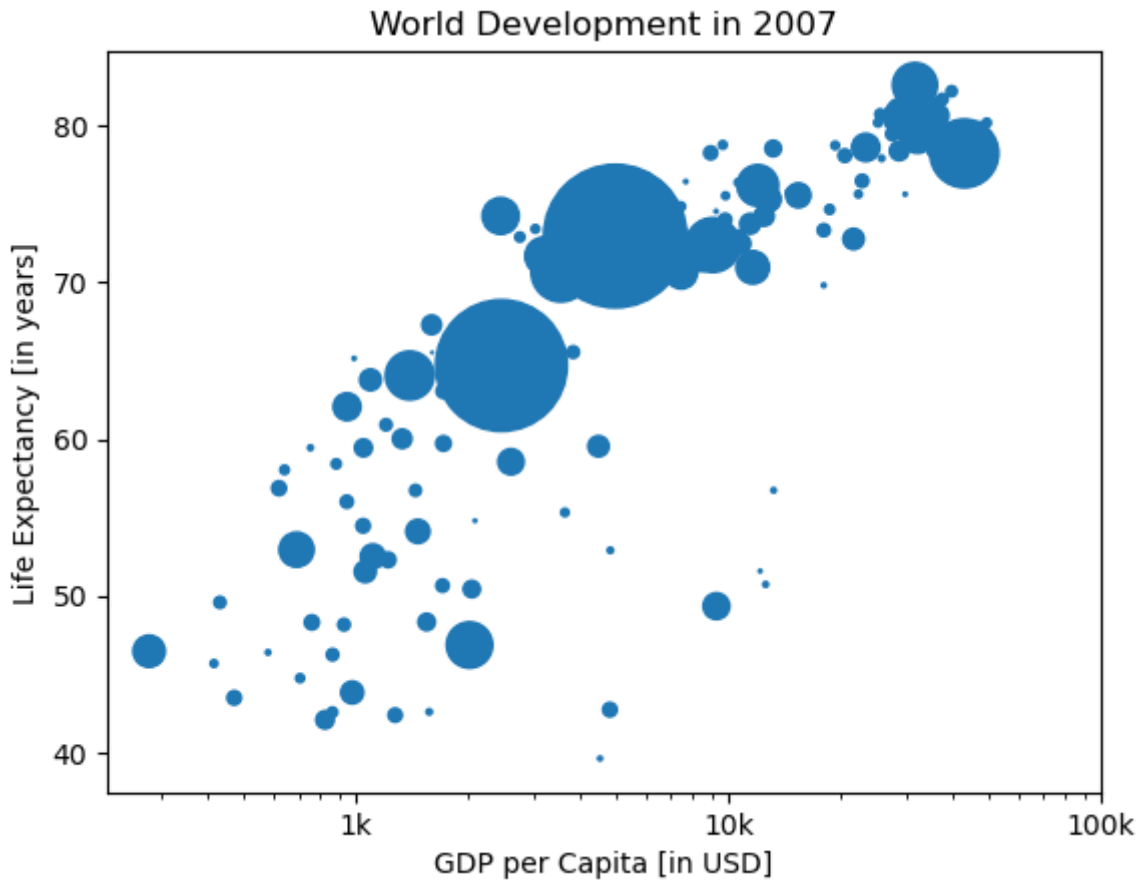


Exercise 7

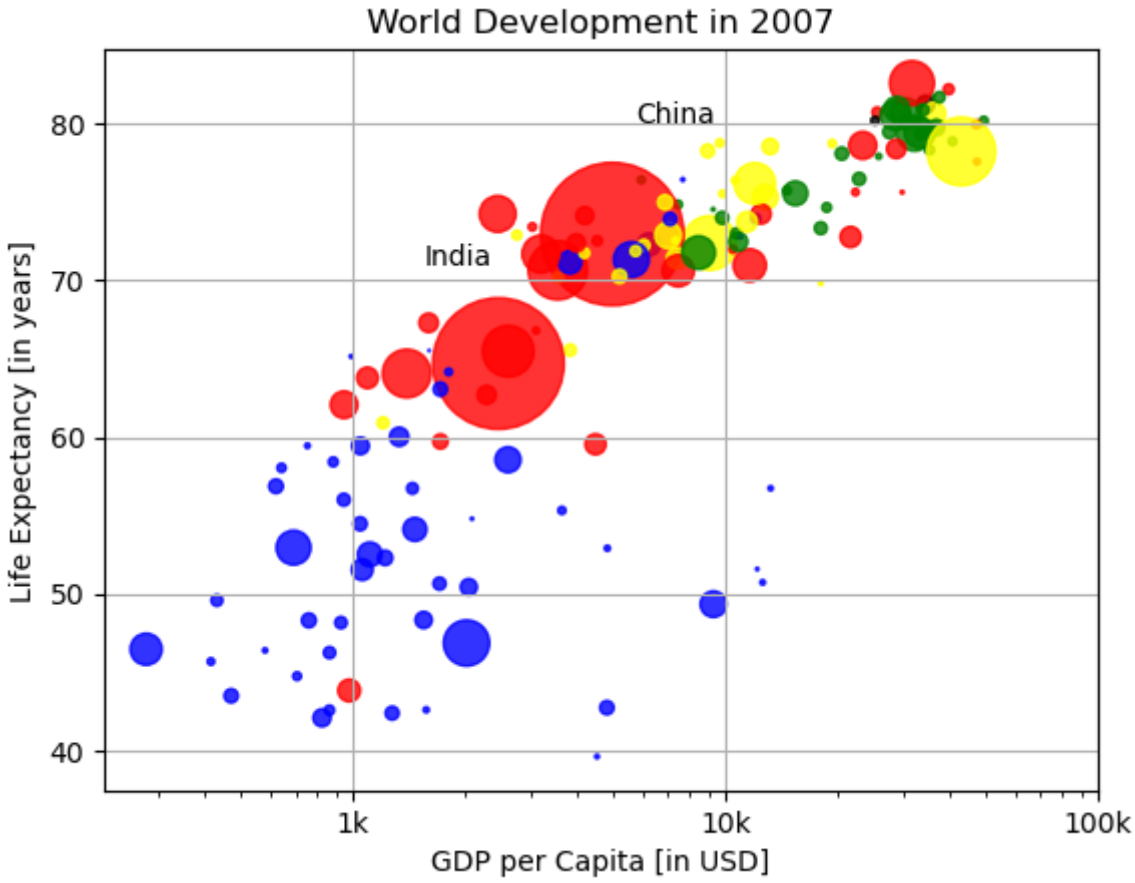
In [37]:
```
# Scatter plot
plt.scatter(gdp_cap, life_exp)
# Previous customizations
plt.xscale('log')
plt.xlabel('GDP per Capita [in USD]')
plt.ylabel('Life Expectancy [in years]')
plt.title('World Development in 2007')
# Definition of tick_val and tick_lab
tick_val = [1000,10000,100000]
tick_lab = ['1k','10k','100k']
# Adapt the ticks on the x-axis
plt.xticks(tick_val,tick_lab)
plt.yticks([30, 50, 70, 90]) #setting tick
# After customizing, display the plot
plt.show()
```



In [38]:
```
# Import numpy as np
import numpy as np
# Store pop as a numpy array: np_pop
np_pop = np.array(pop)
# Double np_pop
np_pop = np_pop * 2
# Update: set s argument to np_pop
plt.scatter(gdp_cap, life_exp, s = np_pop)
# Previous customizations
plt.xscale('log')
plt.xlabel('GDP per Capita [in USD]')
plt.ylabel('Life Expectancy [in years]')
plt.title('World Development in 2007')
plt.xticks([1000, 10000, 100000],['1k', '10k', '100k'])
# Display the plot
plt.show()
```



In [40]:
```
col = ['red', 'green', 'blue', 'blue', 'yellow', 'black', 'green', 'red', 'r
# Scatter plot
plt.scatter(x = gdp_cap, y = life_exp, s = np.array(pop) * 2, c = col, alpha
# Previous customizations
plt.xscale('log')
plt.xlabel('GDP per Capita [in USD]')
plt.ylabel('Life Expectancy [in years]')
plt.title('World Development in 2007')
plt.xticks([1000,10000,100000], ['1k','10k','100k'])
# Additional customizations
plt.text(1550, 71, 'India')
plt.text(5700, 80, 'China')
# Add grid() call
plt.grid(1)
# Show the plot
plt.show()
```



Exercise 8

8) What can you say about the plot? Which one is True?

- The countries in blue, corresponding to Africa, have both low life expectancy and a low GDP per capita.

In [ ]: