Chapter 6 - exercise 3: Year - Pop - gdp_cap - life_exp

Cho dữ liệu year, pop, gdp_cap, life_exp từ tập tin du_lieu_year_gap_pop_life.txt, col từ tập tin col.txt

Hãy thực hiện các yêu cầu sau:

1. Tạo year, pop. In item cuối của year và pop
2. Vẽ biểu đồ line của year và pop với x-axis: year, y-axis: pop
3. Tạo gdp_cap, life_exp. In item cuối của gdp_cap, life_exp
4. Vẽ biểu đồ line của gdp_cap, life_exp với x-axis: gdp_cap, y-axis: life_exp. Biểu đồ này có thể xem được không? Nếu không thì đề xuất một loại biểu đồ phù hợp
5. Vẽ histogram của life_exp, màu cột xanh, viền đỏ
6. Vẽ histogram với bins = 5, màu cột xanh dương, viền đỏ
7. Vẽ histogram với bins = 20, màu cột xanh dương, viền đỏ
8. Từ life_exp1950 trong tập tin => Tạo life_exp1950. Vẽ histogram life_exp1950, với bins = 15, màu cột xanh dương, viền đỏ
9. Tạo scatter plot của gdp_gap và life_exp, sử dụng plt.xscale('log'). Thiết lập xlabel, ylabel, title
10. Tạo Scatter plot của gdp_gap và life_exp, sử dụng plt.xscale('log'). Thiết lập xlabel, ylabel, title. plt.scatter(gdp_cap, life_exp). Với: tick_val = [1000,10000,100000] và tick_lab = ['1k','10k','100k'] => plt.xticks(tick_val, tick_lab)
11. Đưa pop vào array np_pop. Vẽ scatter plot của gdp_cap, life_exp, với s = np_pop * 2, màu magenta. Thiết lập xlabel, ylabel, title và plt.xticks([1000, 10000, 100000],['1k', '10k', '100k'])
12. Tạo col từ col trong tập tin col.txt. Vẽ scatter plot của gdp_cap, life_exp, với s = np.array(pop) * 2, màu c = col, alpha=0.8. Thiết lập xlabel, ylabel, title và plt.xticks([1000, 10000, 100000],['1k', '10k', '100k'])
13. Tạo col từ col trong tập tin col.txt. Vẽ scatter plot của gdp_cap, life_exp, với s = np.array(pop) * 2, màu c = col, alpha=0.8. Thiết lập xlabel, ylabel, title và plt.xticks([1000, 10000, 100000],['1k', '10k', '100k']). Thêm text cho 2 nơi trên biểu đồ là India và China: plt.text(1550, 71, 'India'), plt.text(5700, 80, 'China'). Thêm lưới cho biểu đồ

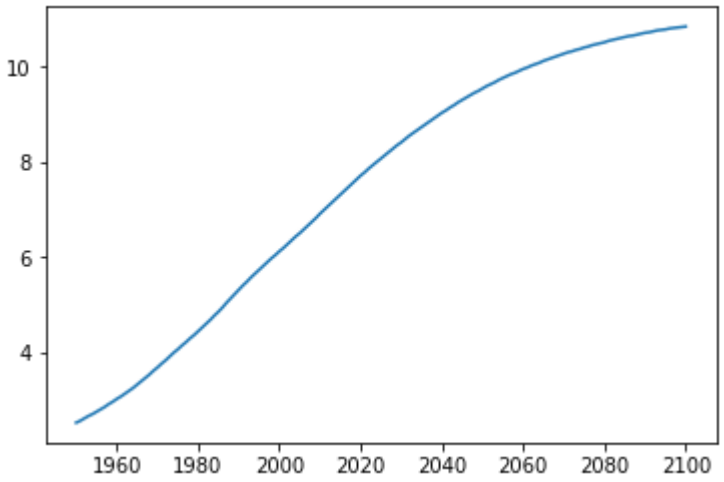
In [10]:
```
# Câu 1:
year = [1950, 1951, 1952, 1953, 1954, 1955, 1956, 1957, 1958, 1959, 1960, 1961, 19
```

In [11]:
```
pop = [2.53, 2.57, 2.62, 2.67, 2.71, 2.76, 2.81, 2.86, 2.92, 2.97, 3.03, 3.08, 3.1
```

In [12]:
```
# Print the last item from year and pop
print(year[-1])
print(pop[-1])
```

```
2100
10.85
```

In [13]:
```
# Câu 2: Vẽ biểu đồ line của year và pop với x-axis: year, y-axis: pop
import matplotlib.pyplot as plt

# Make a line plot: year on the x-axis, pop on the y-axis
plt.plot(year, pop)

# Display the plot with plt.show()
plt.show()
```



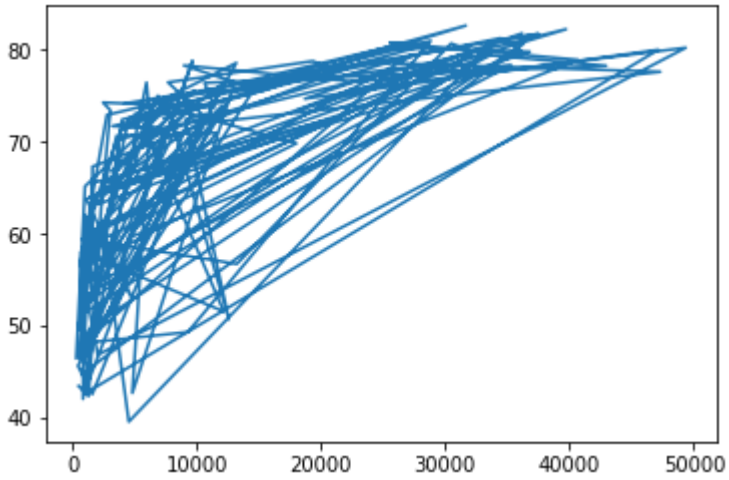
In [14]:
```
# Câu 3:
gdp_cap = [974.5803384, 5937.029525999998, 6223.367465, 4797.231267, 12779.37964,
```

In [15]:
```
life_exp = [43.828, 76.423, 72.301, 42.731, 75.32, 81.235, 79.829, 75.635, 64.062,
```

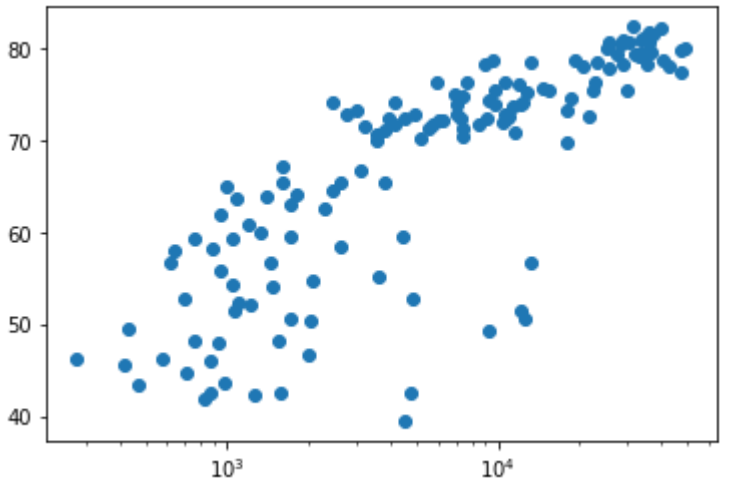
In [16]:
```
# Print the last item of gdp_cap and life_exp
print(gdp_cap[-1])
print(life_exp[-1])
```

```
469.70929810000007
43.487
```

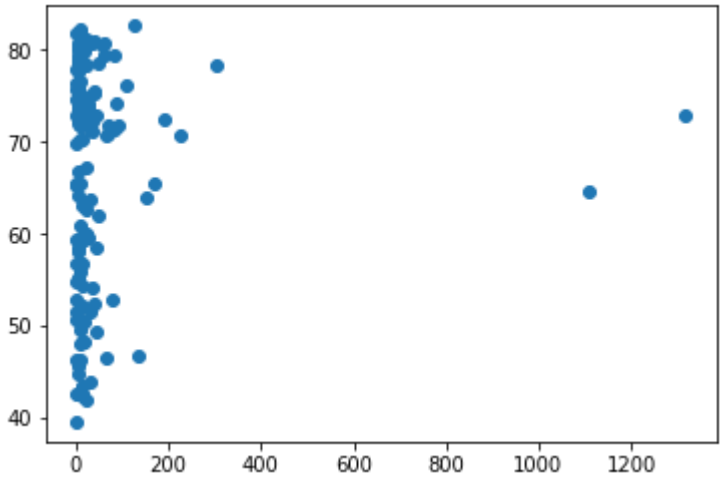
In [17]:
```
# Câu 4: Vẽ biểu đồ line của gdp_cap, life_exp với x-axis: gdp_cap, y-axis: life_e
# Biểu đồ này có thể xem được không? Nếu không thì đề xuất một loại biểu đồ phù hợ
# Make a line plot, gdp_cap on the x-axis, life_exp on the y-axis
plt.plot(gdp_cap, life_exp)

# Display the plot
plt.show()
```



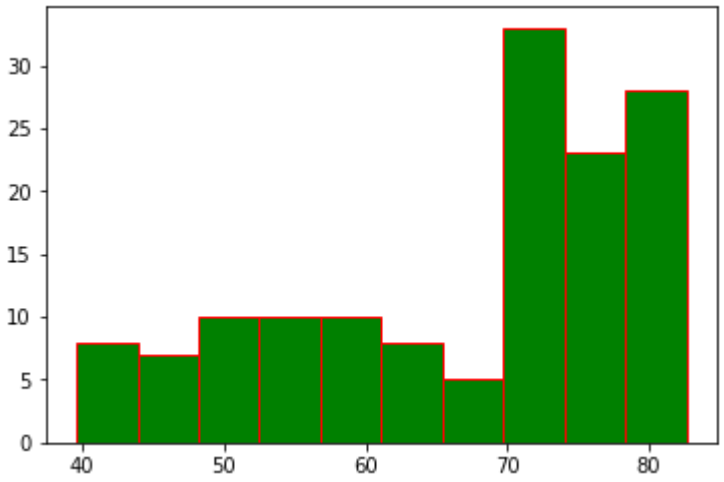
In [18]:
```
# Change the line plot below to a scatter plot
plt.scatter(gdp_cap, life_exp)

# Put the x-axis on a logarithmic scale
plt.xscale('log')

# Show plot
plt.show()
```



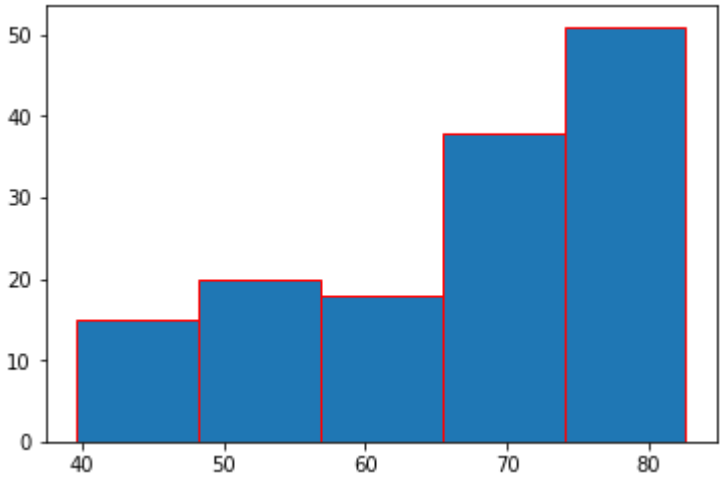
In [19]:
```
pop = [31.889923, 3.600523, 33.333216, 12.420476, 40.301927, 20.434176, 8.199783,
# Build Scatter plot
plt.scatter(pop, life_exp)

# Show plot
plt.show()
```



In [20]:
```
# Câu 5: Vẽ histogram của life_exp, màu cột xanh, viền đỏ
# Create histogram of life_exp data
plt.hist(life_exp, color='g', edgecolor='r')

# Display histogram
plt.show()
```

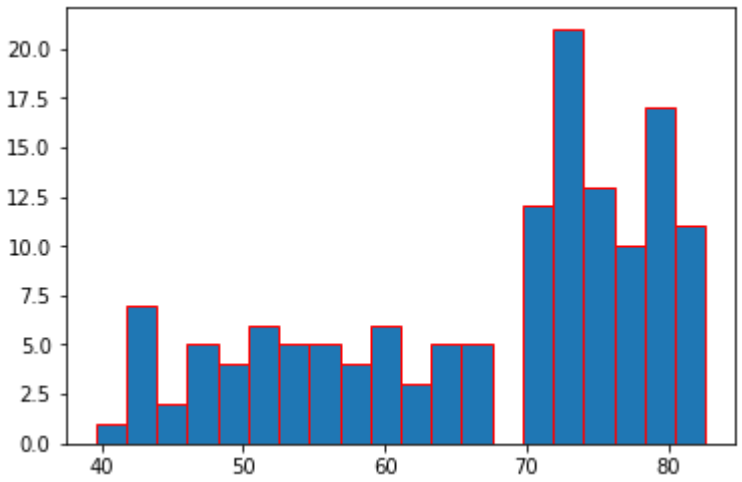


In [21]:
```
# Câu 6: Vẽ histogram với bins = 5, màu cột xanh dương, viền đỏ
plt.hist(life_exp, bins=5, edgecolor='r')

# Show and clean up plot
plt.show()
plt.clf()
```




```
<Figure size 432x288 with 0 Axes>
```

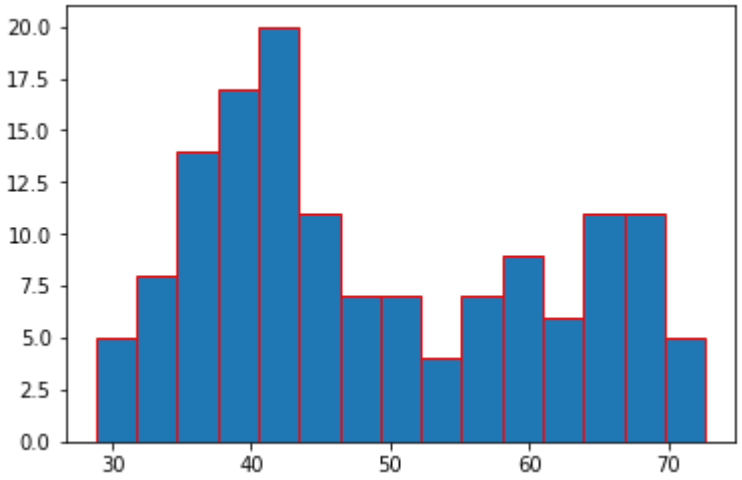
In [22]:
```
# Câu 7: Vẽ histogram với bins = 20, màu cột xanh dương, viền đỏ
plt.hist(life_exp, bins=20,  edgecolor='r')

# Show and clean up again
plt.show()
plt.clf()
```




```
<Figure size 432x288 with 0 Axes>
```

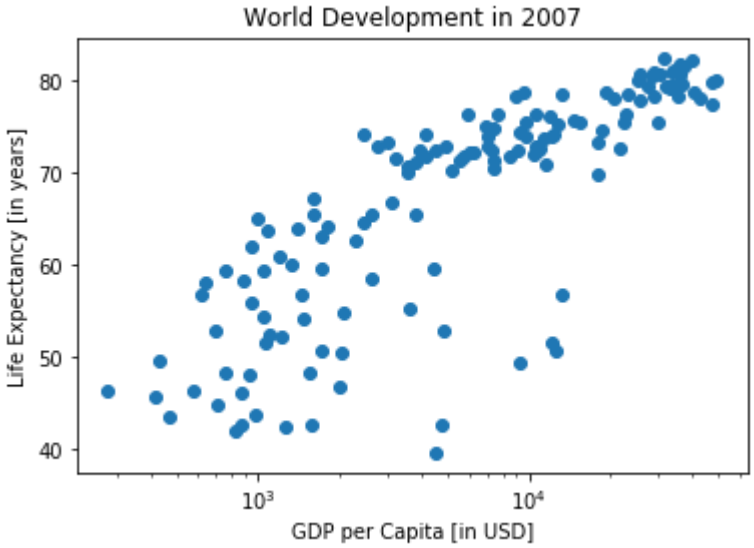
In [23]:
```
# Câu 8: Vẽ histogram life_exp1950, với bins = 15, màu cột xanh dương, viền đỏ
life_exp1950 = [28.8, 55.23, 43.08, 30.02, 62.48, 69.12, 66.8, 50.94, 37.48, 68.0,
```

In [24]:
```
# Histogram of life_exp1950, 15 bins
plt.hist(life_exp1950,  bins= 15,  edgecolor='r')

# Show and clear plot again
plt.show()
plt.clf()
```

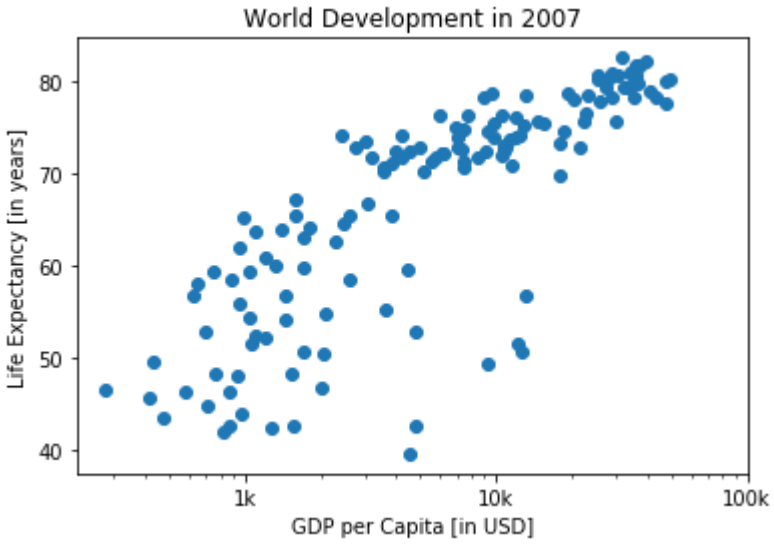



```
<Figure size 432x288 with 0 Axes>
```

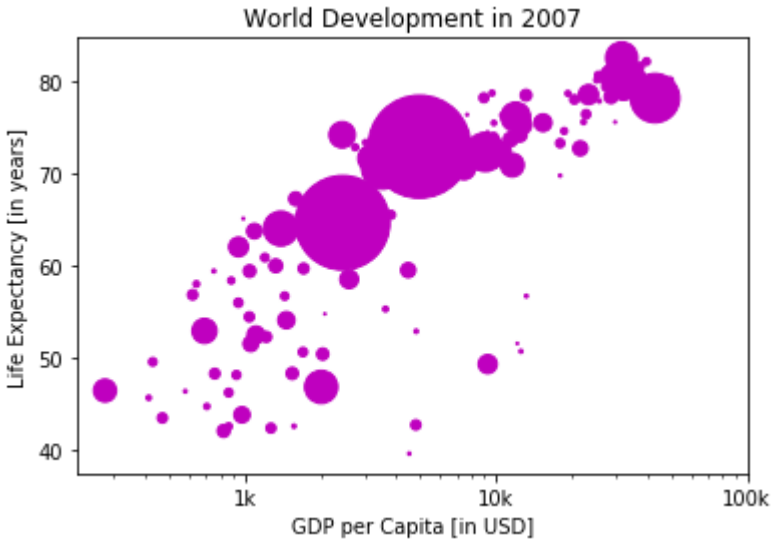
In [25]:
```
# Câu 9: Tạo scatter plot của gdp_gap và life_exp, sử dụng plt.xscale('log'). Thiế
plt.scatter(gdp_cap, life_exp)
plt.xscale('log')

# 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)

# Add title
plt.title(title)

# After customizing, display the plot
plt.show()
```



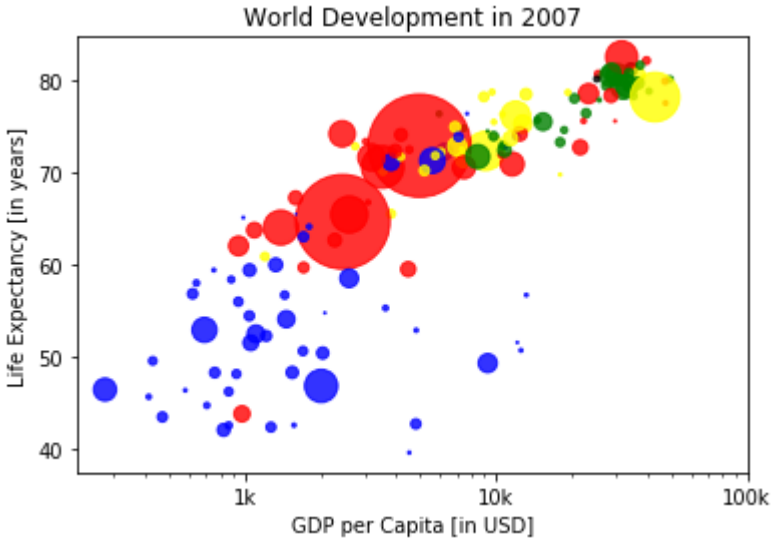
In [30]:
```
''' Câu 10: Tạo Scatter plot của gdp_gap và life_exp, sử dụng plt.xscale('log'). T
plt.scatter(gdp_cap, life_exp). Với:
tick_val = [1000,10000,100000]
tick_lab = ['1k','10k','100k']
plt.xticks(tick_val, tick_lab)
'''
plt.scatter(gdp_cap, life_exp)
# Previous customizations
plt.xscale('log')
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)

# After customizing, display the plot
plt.show()
```



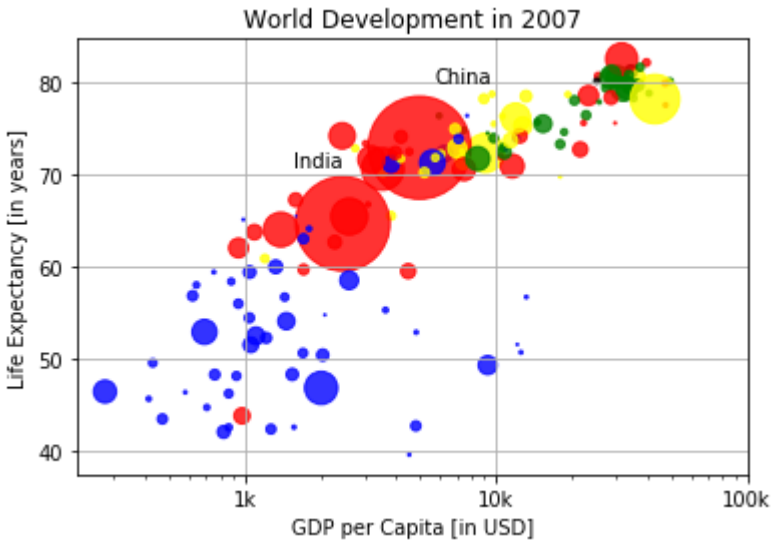
In [27]:
```
'''
Câu 11: Đưa pop vào array np_pop
Vẽ scatter plot của gdp_cap, life_exp, với s = np_pop * 2, màu magenta
Thiết lập xlabel, ylabel, title và plt.xticks([1000, 10000, 100000],['1k', '10k',
'''
# Import numpy as np
import numpy as np

# Store pop as a numpy array: np_pop
np_pop = np.array(pop)

# Update: set s argument to np_pop
plt.scatter(gdp_cap, life_exp, s = np_pop * 2, color='m')

# 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 [28]:
```
'''
Câu 12: Tạo col từ col trong tập tin col.txt
Vẽ scatter plot của gdp_cap, life_exp, với s = np.array(pop) * 2, màu c = col, alp
Thiết lập xlabel, ylabel, title và plt.xticks([1000, 10000, 100000],['1k', '10k',
'''
col = ['red', 'green', 'blue', 'blue', 'yellow', 'black', 'green', 'red', 'red', '
       'yellow', 'green', 'blue', 'yellow', 'green', 'blue', 'blue', 'red', 'blue'
       'blue', 'yellow', 'red', 'yellow', 'blue', 'blue', 'blue', 'yellow', 'blue'
# Specify c and alpha inside plt.scatter()
plt.scatter(x = gdp_cap, y = life_exp, s = np.array(pop) * 2, c= col, alpha=0.8)

# 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'])

# Show the plot
plt.show()
```



In [29]:
```
'''
Câu 13: Tạo col từ col trong tập tin col.txt
Vẽ scatter plot của gdp_cap, life_exp, với s = np.array(pop) * 2, màu c = col, alp
Thiết lập xlabel, ylabel, title và plt.xticks([1000, 10000, 100000],['1k', '10k',
Thêm text cho 2 nơi là India và China:
plt.text(1550, 71, 'India')
plt.text(5700, 80, 'China')
Thêm lưới cho biểu đồ
'''
plt.scatter(x = gdp_cap, y = life_exp, s = np.array(pop) * 2, c = col, alpha = 0.8

# 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()

# Show the plot
plt.show()
```



In [ ]: