San Diego State University – Department of Electrical and Computer Engineering
EE300 - Computational and Statistical Methods of Electrical Engineers
Summer 2020 - MATLAB Project

The following dataset was obtained from the Organization for Economic Cooperation and Development (OECD) website on the *Gross Domestic Product* (GDP) per capita and the *Life Satisfaction Index* (LSI) of the following countries:

Index	Country	GDP in $1000 USD	Life Satisfaction Index (LSI)
1.	Australia	33.4	7.3
2.	Austria	32.5	7
3.	Belgium	30.0	6.9
4.	Canada	29.9	7.3
5.	Chile	16.6	6.7
6.	Czech Republic	21.1	6.6
7.	Denmark	29.0	7.5
8.	Estonia	18.7	5.6
9.	Finland	29.5	7.5
10.	France	31.1	6.4
11.	Germany	33.7	7
12.	Greece	17.0	5.2
13.	Hungary	16.8	5.3
14.	Iceland	30.5	7.5
15.	Ireland	25.4	7
16.	Israel	24.0	7.2
17.	Italy	26.1	5.9
18.	Japan	28.6	5.9
19.	Korea	21.7	5.9
20.	Latvia	15.3	5.9
21.	Luxembourg	41.3	6.9
22.	Mexico	13.9	6.6
23.	Netherlands	28.8	7.4
24.	New Zealand	24.4	7.3
25.	Norway	35.7	7.5
26.	Poland	18.9	6
27.	Portugal	20.5	5.2
28.	Slovak Republic	20.2	6.1
29.	Slovenia	20.5	5.8
30.	Spain	23.1	6.4
31.	Sweden	30.6	7.3
32.	Switzerland	36.4	7.5
33.	Turkey	17.1	5.5
34.	United Kingdom	28.4	6.7
35.	United States	44.0	6.9
36.	Brazil	12.2	6.6
37.	Russia	16.7	6
38.	South Africa	10.9	4.8

You should define x as a vector recording the data for the GDP and y as a vector recording the life satisfaction index. You can save your data into a 'mat' file by the MATLAB function save('data.mat','x','y'). Next time when you retrieve the data, use the MATLAB function load('data.mat').

Perform the following tasks:

1. *(2pt)* Find the (sample) means of the GDP and the LSI and store them as m_X and m_Y.

   - *(1pt)* You may wish to try the MATLAB functions $mean(x)$ and $mean(y)$.
   - *(1pt)* Write your own function $my\_mean(x)$ that takes input as a data vector x and returns its mean m_X. This function should provide the following result

   $$m_X = \frac{1}{N}\sum_{i=1}^{N} x_i$$

   where N is the length of the data vector x and x_i is its ith element. Compare the result you get from $my\_mean(x)$ and $my\_mean(y)$ with $mean(x)$ and $mean(y)$

2. *(2pt)* Find the (sample) standard deviation of the GDP and the LSI and store them as std_X and std_Y.

   - *(1pt)* You may wish to try the MATLAB functions $std(x)$ and $std(y)$.
   - *(1pt)* Write your own function $my\_std(x)$ that takes input as a data vector x and returns its standard deviation m_X. This function should provide the following result

   $$std_X = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - my\_mean(x))^2}.$$

   You need to use your own function $my\_mean(x)$ to get the mean of x. Compare the result you get from $my\_std(x)$ and $my\_std(y)$ with $std(x)$ and $std(y)$.

3. *(2pt)* Determine the median of the GDPs and store it as $median_X$

   - *(1pt)* You may wish to try the MATLAB function $median(x)$.
   - *(1pt)* Write your own function $my\_median(x)$ that takes input as a data vector x and returns its median $median_X$. You must sort the data vector x in an ascending order into x_a first, and then proceed to find the median (the mid point). You must write a script that is usable for both odd and even number of samples. Print your result and compare it with the result obtained from the MATLAB function $median(x)$.

4. *(1pt)* Illustrate the data in the table using a scatter plot with the horizontal axis representing the GDP and the vertical axis representing the LSI. From the plot, determine whether the GDP and the LSI have positive correlation, negative correlation or no correlation.

   Suggestion: You may wish to try the MATLAB function $scatter(x, y)$. Attach the plot in your report.

5. *(1pt)* Find the sample covariance of the GDP and the LSI and store it as cov_{XY}.

   Suggestion: You may wish to try the MATLAB command $N/(N-1) * mean((x - mx). * (y - my))$.

6. *(2pt)* Find the correlation coefficient from the sample covariance and the standard deviations of the GDP and the LSI. Print the obtained correlation coefficient, and determine if the calculation confirms with your observation from the scatter plot.

Please turn in your MATLAB script in one section and the results/plots in another section. Attach all in one document and upload to the Project submission link on Blackboard.

Deadline: Sunday Aug 16, 2020 at 9:00AM. No late submission is allowed. The grade of the course needs to be completed by the following Monday!