BF550: Fall 2020
Problem Set 2 is due by 12 pm on Tuesday, September 29

Reading Assignment
Using Python documentation (`https://matplotlib.org`) or any other source familiarize yourself with line plots, scatter plots, bar plots, and histograms.

Submission instructions
Follow the same submission instructions as before. Send completed assignments to Howard and cc Kirill. Submit this assignment as a Python notebook that shows the code and the final plots.

Problem 1
Use the oil price data either from the dat or csv file to make an informative and aesthetically pleasing plot of oil prices vs. time.

Problem 2
This problem follows up on what you did last week with the Human Microbiome Project data.

- Make a bar plot showing the number of samples for each body site.
- Update the previous plot to show the number of samples for male and female subjects. The bars for different genders should be in different color and grouped by body sites.

Problem 3
Download *E. coli* genome from the course website and use it to simulate a random walk. Assume that 'A' and 'T' correspond to a forward step $\Delta x = +1$ while 'C' and 'G' to a backward step $\Delta x = -1$. Make a plot of "position" x vs. the number of base pairs read/traveled for 3 different initial positions within the genome. Each random walk should travel 100 consecutive base pairs. Let us assume that $x = 0$ at the start of the random walk. For more information on random walks, see Wikipedia: `https://en.wikipedia.org/wiki/Random_walk`. Then, make a histogram of the random walk displacements over one walk for 200 walks; each consecutive random walk starts where the previous one finishes. Finally, combine the two figures made in this problem into a single figure with two subplots arranged horizontally. Add panel labels: "(a)" and "(b)".