Technique Assignment 1: Python
Cogs 109 Spring 2020
Due: Sunday April 12 11:59pm
75 points total

Submit your completed assignment as a single pdf report containing all graphs and answers to questions. Include all Python code you generated, either as part of a Jupyter notebook or as an appendix to your report. Include comments for clarity.

Your pdf should have your name at the top and should be easy to read. Each question should be clearly marked and answers to all questions should be clearly indicated (use Markdown cells in your Jupyter notebook to clearly display text).

1. **Basic programming (10)** Without using `range()` or other built-in functions, write Python code to add (summate) every third number, starting with 3, under 100. That is, you are finding the result of 3 + 6 + … + 99, but without typing out each individual number. *Hint*: you may use a loop.
   Show your code, including the result, with comments in your report.

2. **List operations (10)** Rewrite your solution to Question 1 using `range()`.
   Show your code, including the result, with comments in your report.

3. **Playing with strings (10)** Create a string and store it in a variable. Use indexing to extract a portion of the string from your variable and store it in a new variable. Print the result.

4. **Datasets (15)** Load the Fisher's iris data set in Python from the sklearn library.
   a. List and describe the data that has been loaded (hint: there are two arrays)
   b. How many **variables** are being measured in this data set?
   c. How many **samples** are there of each measurement?
   d. Is there a label for each sample?
   e. Looking at the numbers in the `iris.data` array, can you spot any trends in the data? If so, describe any trends (patterns) you notice.

5. **Plotting data (20)** Create a figure with 4 subplots in a 2x2 pattern. Use the `iris.data` array from the iris dataset to create the following plots (plot labels are helpful but not required):
    a. In the first plot window, display the variables in the array with Matplotlib using `plot(iris.data)`. Describe the plot – What is the x axis? What is the y axis? What do the different lines represent?
    b. Looking at this plot, can you spot any trends in the data? If so, describe any trends (patterns) you notice.
    c. In the second plot window, plot column 1 of `iris.data` on the x axis and column 2 on the y axis. Plot each sample as a single point with no connecting lines.
    d. Looking at this plot, can you spot any trends in the data?
    e. In the third plot window, plot a sample of the dataset from the second plot window by plotting every 5^{th} row of column 1 and column 2.
    f. In the fourth plot window, plot a histogram of the column 2 data. Describe the pattern shown by the histogram.

6. **Plotting with labels (10)** Notice that the `iris.target` array contains 3 different labels. Create a plot that you think is interesting using the data from the `iris.data` array. Use three different colors to label each point with a color representing its species. Correctly label the x and y axes and include a legend for species.

7. **Extra credit** (Up to 5 points) Create a short video (~1-3 minutes) of you working on a portion of this assignment (coding in Python). Show work in progress—it's ok if your process is messy! Submit your video as a link in your pdf (linking to YouTube, Google drive, etc.) and indicate whether it is ok to share with the class.