

DAV 5400 Module 1 Assignment (30 Points)

A crucial aspect of working effectively with lists in Python is understanding how Python applies its internal indexing structure to lists. For this assignment you will be creating, searching, and sorting lists you will create from the contents of a small text file containing data values that are separated by commas (also commonly referred to as a 'CSV' file (Comma Separated Values)) using some of the concepts we are learning about this week.

Start by downloading the **M1_Data40.txt** file to your local environment.

Then, open a new Jupyter Notebook and copy in the following small Python code snippet:

```
import csv

# Be sure to update the file path contained within the 'open()' statement shown below
# to reflect the location of the file within your own computing environment!!
# Also be sure that the code is properly indented after you paste it!

with open('yourpath/M1_Data40.txt') as csvfile:
    readCSV = csv.reader(csvfile)
    for row in readCSV:
        # print each row as read by the csv.reader function
        print(row)
```

This code will read the contents of a file containing data values that are separated by commas in a line-by-line manner and print the contents of each line in the form of a Python list. Each item within that list will contain the corresponding element of the comma-separated line within the data file from which it came.

Once you've run the code snippet, you will see that each line of the file contains six distinct values. For example, a print of the result of reading the first line will show the following:

```
['FALSE', '63', 'MALE', 'TYPICAL_ANGINA', '145', '233']
```

These six distinct values represent attributes of hospital patients who had complained of chest pain. Specifically we have the following:

- Diagnosis (i.e., was the patient diagnosed with heart disease?)
- Age
- Gender
- Type of chest pain
- Blood pressure reading
- Cholesterol level

Now that you are familiar with the content of the file, **complete the following 6 tasks:**

(Assignment Continues on Next Page)

1. **(5 Points)** Your first task is to read the file again and extract these six attributes from each line of the file and create six distinct Python list objects (one for each of the six attributes) comprised solely of the values you extracted for a given attribute. In other words, you should have separate lists for diagnosis, age, gender, etc. For example, the first five "chest pain type" values should be as follows:

```
['TYPICAL_ANGINA', 'ASYMPTOMATIC', 'ASYMPTOMATIC', 'NON_ANGINAL', 'ATYPICAL_ANGINA']
```

One way to complete this task would be to use the code snippet from above as a starting point: you could replace the "print(row)" statement with whatever Python code you feel is necessary to create the required list objects. When finished, each of your list objects should contain 40 data values.

Note that the lists you have created thus far are composed solely of character strings. However, three of the attributes (Age, Blood Pressure, and Cholesterol) are clearly meant to be used as integer data values. As such, before proceeding any further, convert the contents of each of the 'Age', 'Blood Pressure', and 'Cholesterol' lists to integer values

2. **(5 Points)** Your second task entails using the updated Blood Pressure list to find the **list index** values of each hospital patient having a "**blood pressure**" reading **that is greater than 140**. Create a new list object with your result. ***HINT:*** *you can accomplish this task by searching the list of blood pressure values you created as part of the first task.* Be sure to print your results.
3. **(5 Points)** Your third task is to find the "**pain type**" value for each hospital patient having a "**blood pressure**" value **that is greater than 140**. Create a new list to store your findings and be sure to print your results.
4. **(5 Points)** Your fourth task is to find the index value for each hospital patient having a "**gender**" value of "**FEMALE**" and a "**pain type**" value that ***is not*** "**ASYMPTOMATIC**". Create a new list to store your findings and be sure to print your results.
5. **(5 Points)** Your fifth task is to find the index value for each hospital patient having an '**age**' value of **less than 56** and a "**diagnosis**" value of "**TRUE**". Create a new list to store your findings and be sure to print your results.
6. **(5 Points)** Finally, calculate the average **age**, the median **blood pressure** reading, and the standard deviation of the **cholesterol** readings using whichever of Python's built in functions you require. Print your results.

Be sure to **include some commentary within formatted Markdown cells** explaining your approach to solving each of the individual problems. When you are finished, save the Jupyter Notebook containing your work and commentary and submit it via the Module 1 Assignment page on Canvas. Be sure to save your Notebook using the following nomenclature: **first initial_last name_M1_assn**" (e.g., J_Smith_M1_assn_).