## McKesson Assignment 02 Deep Azure

## 

### Handed out: 10/17/2017 Due by 11:59 PM (CST) on Tuesday, 10/24/2017

Download notebook file python\_fundametals.ipynb from the class site.. If you are new to Python, our recommendation is to go through every cell of the notebook and execute it. Save your notebook.

**Problem 1**. Install Anaconda for Python 2.7. We do not care whether it is the very latest release. If you already have a working Python 2.7 on your machine please be free to use it. Show to use what versions of Python, conda, pip and jupyter apps you have on your system. If you insist you can work with Python 3.

(10%)

**Problem 2**. Write a Python function that, when passed a single number as its argument, will return a tuple containing that number, square of that number and sin() of that number. Place this function in a loop going through a sequence of integers starting with 0 and ending with 9. Pass each integer to the function and print results to an output cell of your jupyter notebook.

(20%)

**Problem 3.** Copy the code for problem #2 from your notebook into a standalone script. Name the script as you please but add suffix .py to it. You can modify that script in a text editor like Notepad. Do not do it in MS Word, Power Point or WordPad. If you have successfully install Anaconda, it came with a Python script editor called Spyder. You can choose to use Spyder. Demonstrate that you can run your script on the command prompt by typing:

C:..> python script.py

and obtain the same results as in the notebook. Copy back the content of your script into a cell of your notebook. Run it in that cell.

(20%)

**Problem 4**. Modify previous script so that it reads a sequence of integers from a simple text file and write its results into another text file. Prepare the input file with a single integer on every line. Write outputs into the output file similarly, one output per line.

(25%)

**Problem 5**. Consider attached file small\_car\_data.xlsx. Import data into a panda DataFrame. Use panda machinery to calculate basic statistics for all numerical columns, min, max, median, average and standard deviation. Determine and present graphically statistical distributions of values in those columns. Try to establish correlations between values in different columns. Which two columns are mutually most correlated?

(25%)

SUBMISSION INSTRUCTIONS:

Your main submission should be an MS Word or PDF document containing your code, results produced by that code and brief textual descriptions of what you did and why. Typically, you copy important snippets of your code as ASCII text and the results into this Word document. Please copy ASCII text rather than take a snapshot. If your MS Word document is larger than 1 MB save it as a MINIMIZED PDF. Describe the purpose of every code snippet and the significance of the results. Start with the text of this homework assignment as the template. Please add any other files that you might have used or generated. Please do not provide ZIP or RAR or any other archives. Canvas cannot open those archives and they turn into a nuisance for us.