**Ramapo College of New Jersey CMPS 240 Fall 2019**

**Assignment 1**

**Software for CMPS 240: Data Analytics**

Please follow the instructions:

1. Install Python 3.7.4 (Windows x86-64 executable installer) as it is the latest version as of 8/17/2019 from the following link:

<https://www.python.org/downloads/windows>

1. Double click the executable file to start installation, "Customize installation" will be an option that needs to be selected then click Next button. You need to select under "Optional Features": **pip**, tcl/tk and IDLE, and Python test suite, py launcher then click Next button.

Under Advanced Options, select "Add Python to environment variables".

You can check to see if **pip** is installed by typing the following command from command line prompt:

pip3 --version

1. Now you can install all the required modules and their dependencies using **pip** command as follows:

pip3 install --upgrade numpy

pip3 install --upgrade matplotlib

pip3 install --upgrade pandas

pip3 install --upgrade scipy

pip3 install --upgrade scikit-learn

pip3 install --upgrade jupyter

To check your installation, try to import every module like this:

python -c "import jupyter, matplotlib, numpy, pandas, scipy"

There should be no output and no error.

**Or alternatively,**

Fire up Jupyter by typing the following command from cmd:

jupyter notebook

import pandas as pd

medal\_data=pd.read\_csv('http://winterolympicsmedals.com/medals.csv')

medal\_data.shape

===Installation of more modules=====

pip3 install –upgrade seaborn

#To read Data from an HTML File and to write an HTML File

pip3 install html5lib

pip3 install BeautifulSoup4

# To read Excel file

 pip3 **install xlrd**

#To write to Excel file

pip3 install openpyxl

# To read and write HDF5 files

pip3 install tables

#

pip3 install Scrapy

pip3 install pypiwin32

pip3 install psycopg2

pip3 install SQLAlchemy

Test