**BUAN 6340: Assignment No.2**

**Python – Numpy, Matplotlib, pandas**

**Total Points: 30**

Download the following datasets from eLearning

us\_hurricane\_data.csv

world\_city\_temp.csv

sales\_data.csv

worker\_education\_data.csv

1. Read the file called us\_hurricane\_data.csv file. Draw the Bar graph as the x-axis has a year and the YY-axis axis has the number of occurrences. Try to use the options like linewidth, size, and other options we have discussed in matplotlib. Your name should be attached to the Title of the graph.
2. Read the file called world\_city\_temp.csv. Plot the histogram of the temperatures over this period for the cities of San Jose and Zurich. Do the same for Zurich and Sidney. Your name should be attached to the Title of the graph.
3. Read the file called sales\_date.csv. Load the data using panda and show full, head and tail of the data as we discussed in the class. Convert the .csv file to excel, text and load it again.
4. From sales\_data.csv file, Calculate count, mean, std, max, min of unit price.
5. From the sales\_data.csv file. Calculate the total sales price by the customer and draw the graph that shows the company on the x-axis and total sales on the y-axis.
6. Draw a scatter graph of any 50 random values of the x and y-axis. Use matplotlib library to do this. Increase the font size, Title should have your name and other cosmetic changes to graph to make it more presentable with all labels.
7. We have the following data

First\_names = Jason,Molly,Tina,Jake,Amy

Last Names = Miller,Jackson,fey,albert

gender = m,f,f,m,f

Age = 18,20,22,19

ACT = 29,32,31,35

SAT = 1280,1420,1360,1520

* Load the data as five columns and five rows
* Draw a Scatterplot of ACT and SAT scores, with the size of each point determined by age
* Compare SAT and ACT scores for each person with the percentage of each type.

SAT max score is 1600, ACT Max score is 36

1. Numpy arrays have the following numbers

2 0 3 1 2 4 6 8 7 2 0 7 4 5 6 8 5 5 6

Display following:

* Size of the array
* Dimension of the array
* numberer of times each digit is repeated
* Total of all the numbers

1. Create a NumPy array with a size of 12 X 12 with random values

Display the following from the array you created

* Minimum
* Maximum
* Sum
* Average

1. Create a random array of 4 rows and 4 columns.

Display values by

* Sort by the first column
* Sort by the third column
* Apply a few mathematical functions ( minimum 3)

Use Worker\_education\_data.csv to answer following questions

1. Load the data from each column to separate the numpy array

Display each column with the title.

1. Using the above NumPy arrays

Display the following information from the arrays

* Total Undergraduates
* Total Graduates
* Total Ph.D. Holders
* Develop a graph with these numbers

1. Using the above NumPy arrays

Display the following information from the arrays

* Only Male with no college degree
* Only females with Ph.D. holders
* Develop a graph with Age and Degree

1. Create a random vector of size 20 and find the mean values
2. Create a three-dimensional array and get the sum of the last two axes at once

**Expected Output :**

* Assignment jupyter notebook file
* One screenshot covering your work. **MUST HAVE A TOP PORTION.**

Example:

Question: Combine two strings and display both strings and the combined string.

The output should look like the below.

If you do not have a screenshot showing your name, laptop, and time, then you will be graded for 70%.

