CS422/622- HW 1

In HW1, we will play with data visualization tools: Principal Component Analysis (PCA) and Boxplot.

- Data:

- Download the MINST_100.csv and housing_training.csv attached to HW1.

- Task 1: Visualize and interpret the MNIST data using PCA. Reduce the data dimension to two or three and plot the data of reduced dimension. Must plot all the data of ten groups (0 to 9). (40 points)

- Task 2: Visualize and interpret the columns of K, M, and N in the housing data using Boxplot or Violin Plot. The first column is denoted as A, and K, M, and N represent 11-th, 13-th, and 14-th columns, respectively. (30 points)

- Task 3: Visualize and interpret the column of A (1st column) in the housing data using histogram. (30 points)

Grading guideline:

- For all tasks, class labels should be indicated clearly using different colors or texts, and characteristics of the dataset should be discussed.

Submission instructions:

You must submit the following to the assignment folder in UNLV WebCampus:

1. MS word file that describes if available:
   - Used libraries and their purposes
   - Test codes and execution results (some screenshots)
   - Characteristics of the dataset shown through the visualization
   - And anything more you want to describe
2. Source code file(s)
   - Must be well organized (function name, indentation, …)
   - **You need to upload the python text file (*.py.txt). Simply add ".txt" to the py extension. Don't upload jupyter notebook files**

You must submit the files SEPARATELY. DO NOT compress into a ZIP file. If you fail to provide all required information or files, you may be given zero score without grading.

Once you submit, Webcampus will perform similarity check for your submission and show you the result. Your similarity score must be lower than 50% unless something essential is described in the report. Otherwise, (the score -50%) will be deducted. Detecting any attempts to bypass the similarity check may result in receiving zero points.

Deadline:

Please follow the deadline specified in Webcampus. Late submissions will not be accepted.