IEN310 - Fall 2019 – Assignment 1

1. For this exercise you will need to download X\_normal.csv and X\_non\_normal.csv (click [here](https://miamiedu-my.sharepoint.com/%3Af%3A/g/personal/cxl985_miami_edu/Em4uquax6BxMnOrpDf3BG1kBBVhclxIn7g1c2YylU3bMwA?e=NSJdM5) to be directed to the dataset) two data sets and complete the following tasks on both.
	1. Obtain a graphical representation of the variable in the form of a histogram. Plot the histogram with three different sizes of bin (10,15,25,50).
	2. Build a function in MATLAB to compute the mean $μ$, median $med$, and standard deviation $σ^{2}$ of variable X. Do not use MATLAB built-in functions but your own code.
	3. Use the results of your functions on b. to compare against the results obtained to the corresponding built-in MATLAB functions.

Next, by means of simulations, you will verify the results of the Central Limit Theorem (CLT). From each of the two data set:

* 1. Extract 50 samples of size *100* from variable X. Use your function implemented in part b. to compute the mean (call it *sample\_mean\_X*).
	2. Build a histogram of *sample\_mean\_X*. What do you observe? Write down your observations.
	3. Now, use the function you implemented in part ‘b’ to compute the variance of *sample\_mean\_X*. Compare this result against $\frac{σ^{2}}{n}$, (figure out what value of $ n$ is needed)
	4. Comment on the results comparing those from the two different data set (max 200 word).
1. Suppose you have a box with candies. Note that 60% of the candies are sour, the rest are sweet. You extract 10 candies at random.
	1. Code in MATLAB a simulation of the candy box and the extraction process.
	2. Run 500 simulations of a. On each simulation, compute the number of sour candies you obtain.
	3. Compute the theoretical value of the expected value of sour candies in b. and compare with the one you obtained.
	4. Use the results of your simulation in b. to compute the variance.
	5. Compute the theoretical value of the variance and compare with the obtained in d.