**CMSC 462 – Introduction to Data Science**

**Assignment 1**

**Due: 9/15/2024**

**Total Points - 30**

The assignment consists of **three** parts

**First,** the attached Excel file StudentsTestScore shows the points obtained by students in a test sorted by student number. Please do the following-

1. Choose a specific BIN and calculate the frequency distribution and draw a histogram.
2. Calculate the following measures of location – mean, median, mode.
3. Calculate Variance and Standard Deviation.
4. From the above data find out the probability that a student will obtain a grade 71. Explain as a note how did you do that?
5. Estimate the number of students who would get points between 79 and 89. Does it match with the actual data?

**Second,** the attached Excel file Sales\_Data shows the number of customers who made a purchase in a store among the first 1000 people who visited the store.

Please do the following

1. Using the data find out the probability that among 1000 customers there will be between 5 and 8 purchases in a day.
2. Using the data find out the probability that among 1000 customers there will be exactly 6 purchases in a day.
3. Using the data find out the probability that among 1000 customers there will be at most 4 purchases in a day.
4. Is it possible to use Poisson distribution in this case? Explain.
   1. Not applicable because poisson only has a single parameter, lambda, and in this dataset, the variance is not equal to the mean. In poisson, the population mean must be equal or similar enough to the population variance

**Third,** the attached excel file restaurant consists of number of customers that visit a restaurant on a specific day. Based on the data, calculate how much food (in person) the restaurant should prepare so that the restaurant won’t run out of food 90% of the days. Did it match the data? Explain.

What to submit?

Please submit all the answers nicely formatted in a document. Also submit all the supporting files (R, or Python or Excel). In the document, please refer to the supporting files, as necessary.