**CMSC 462 – Introduction to Data Science**

**Assignment 1**

**Due: 2/13/2023**

**Total Points - 30**

The assignment consists of **three** parts

**First,** the attached Excel file Exam\_Score 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.
   1. Mean: 75.44954128
   2. Median: 75
   3. Mode: 73
3. Calculate Variance and Standard Deviation.
   1. Variance: 124.3608563
   2. Standard deviation: 11.10044733
4. From the above data find out the probability that a student will obtain a grade 76. Explain as a note how did you do that?
   1. 0.519774995
   2. Did it by getting the normal distribution
5. Estimate the number of students who would get points between 79 and 93. Does it match with the actual data?
   1. 37

**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 6 and 9 purchases in a day.
   1. 0.379938969
2. Using the data find out the probability that among 1000 customers there will be exactly 7 purchases in a day.
   1. 0.086543703
3. Using the data find out the probability that among 1000 customers there will be at least 7 purchases in a day.
   1. 0.878912604
4. Is it possible to use Poisson distribution in this case? Explain.
   1. Yes, because Poisson distribution can be used to determine x successes in n trials within a given interval of time or sample size

**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 the restaurant should prepare so that the restaurant won’t run out of food 85% of the days.

112.5102969 is the answers

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.