BUSI650
Assignment1 – by Robin Teotia
Total Marks - 100

Part-A | Marks-40

Quetion1. Read the attached "Vehicle_data" csv file, on google colab. [Marks-20]

 Answer the following questions given below: [Marks-10]

   a.   What is the shape of the data?
   b.   List the columns in the data.
   c.   Use python code to summarize the data by showing the details such as Mean, Median etc. of the columns.
   d.   List the quantitative and qualitative variables in the data set.
   e.   Provide the details regarding the data type of the variables.

Note: The answers to the questions must be given using the required python coding.

 Question2. Take the column "TIME_GAP_MIN" from the above data set, convert it into python set. Divide each element in the set by 5 using the iteration through "for loop" and save the remainder in new list named as "rem". [Marks-10]

Hint: To convert a DataFrame column to list.

TIME_GAP_MIN_list=df["TIME_GAP_MIN"].tolist()

Submission: In pdf format with a given link to .pynb file.

Part-B | Marks-60

Question3: What is variance and standard deviation? Provide their formula with the explanation. What do you understand by the data if the variance is large and vice versa. [Marks-20]

Answer limit: 1/2 page maximum.

Question4: Explain Bernoulli, Binomial and Poison distribution by giving practical example of each. State the mean and variance of each distribution. [Marks-30]

Answer limit: 1 page maximum.

Question5: Explain the difference between Normal distribution and t-distribution. [Marks-10]

Answer limit: 2 lines maximum.

Note:

   1.   Try to answer the questions to the point, beating around the bush can lead to negative marking.
   2.   Students must maintain academic integrity, if found guilty, It can lead to 0 marks or the decisive decision taken by the Academic Integrity Committee, UCW.

Submission: Handwritten pdf file is preferred for part-B. Merge both part-A and Part-B pdfs and submit.