Statistical Methods in Data Science
Assignment 1

Due on 5th July 24　　　　　　　　　　　　　　　　　　　　　　　***Marks 100***

Note: You are required to submit **Handwritten** solution in hard form.

Question - 1

Football: Age of Professional Players How old are professional football players? The 11th Edition of *The Pro Football Encyclopedia* gave the following information. Random sample of pro football player ages in years:

24	23	25	23	30	29	28	26	33	29
24	37	25	23	22	27	28	25	31	29
25	22	31	29	22	28	27	26	23	21
25	21	25	24	22	26	25	32	26	29

a)   Compute the mean, median and mode of ages
b)   Calculate the Standard Deviation of the ages data
c)   Find the 77th percentile of this data

Question – 2

A grand statistics test was conducted in a university in which 50 candidates appeared. The total marks of the test were 150 and the highest attained marks were 97. The marks of 50 candidates are shown

60	89	78	65	97	65	45	78	65	34
23	34	55	67	88	90	76	34	32	33
54	67	87	90	45	56	24	34	56	45
76	87	91	65	65	74	53	57	84	88
32	34	36	33	76	50	76	65	45	56

a)   Calculate the range of marks
b)   Find the 3rd quartile of marks
c)   Find the 80th percentile of this data

Question – 3

Two groups of 7 students each were selected from a class of Probability and Statistics and given to solve a mathematical challenge. The scores of groups A and B are given below. Which group shows more variability?

A.   20, 50, 35, 44, 60, 57, 51
B.   40, 45, 42, 60, 50, 62, 30

Question – 4

If A and B are the mean and median of a data set respectively, what will happen to them (will they change or not, any increase/ decrease in value, explain with logic) if we do the following:

a)   Increase the highest value of the dataset
b)   Increase the lowest value of the dataset so that it becomes greater than B

Question – 5

Consider the following data showing the weights of 50 students of grade 6 (rounded to nearest kg)

38	40	30	35	39	40	48	36	31	36
47	35	34	43	41	36	41	43	48	40
32	34	41	30	46	35	40	30	46	37
55	39	33	32	32	45	42	41	36	50
42	50	37	39	33	45	38	46	36	31

a) Construct a stem and leaf plot for the above data
b) Build frequency distributable table, show relative and cumulative frequencies
c) Construct Box plot for this data

Question – 6

Each bit transmitted over the network can be classified as either valid or corrupt. If a four-bit message is to be sent; formulate the sample space
a) What is the probability of the Event that exactly 2 bits were corrupt?
b) What is the probability of the Event that the entire message was transmitted corretly?
c) What is the probability that at least 1 bit was valid?

Question – 7

If two questions have been asked from a student regarding satistifaction with the university infrastructure standards and they can pick an option from "excellent, very good, good, satisfactry, poor" for each of the questions, What is the sample space
a) What is the probability of the Event that a randomly selected student chose "excellent" for both questions?

Question – 8

An order for an automobile can specify either an automatic or mannual transmission, either with or without ac and any of the 4 color choices (red, white, back and silver). List down all possible orders (sample space) for this experiment
a) What is the probability that a randomly selected order specified a red automobile?

Question – 9

A sports club trains players for national tournaments and for each tournament players enrolled at the club are chosen randomly. Each player can be classified as Beginner, Intermediate or Expert. If 3 players are to be chosen for a tournament, list the sample spacce.
a) What is the probability that all chosen players are beginners?
b) What is the probability that at least 1 expert player is chosen?
c) What is the probability that no expert player got chosen?

Question – 10

A department of university consisting of 22 faculty members must assign the following roles to 3 faculty members, "Head of Department, Deputy, Finance Officer". How many different ways can this be done.
a) If 9 of the faculty members are women, what is the probability that women get selected for all 3 positions?

Question – 11

A computer password consists of eight characters.
a) How many different passwords are possible if each character may be any lowercase letter or digit?
b) Howmany different passwords are possible if each character may be any lowercase letter or digit, and at least one character must be a digit?
c) A computer system requires that passwords containat least one digit. If eight characters are generated at random, and each is equally likely to be any of the 26 letters or 10 digits, what is the probability that a valid password will be generated?

Question – 12

A famous cosmetic brand conducted a poll for 2 new shades of lipsticks. 100 girls visiting the outlet were shown the two shades and asked whether they liked them or not. 65 girls said, they liked shade 1, 59 liked shade 2 while 33 liked both. Consider A = [likes shade 1], B=[likes shade 2].
a) How many girls didn't like any of the shades?
b) If a girl picked randomly liked shade 2 what is the probability that she likes shade 1 also?