Department of Mathematics
MTL 390 (Statistical Methods)
Tutorial Sheet No. 2

1. Show that the total sum of the distances between each data and its mean, namely d, $d_i = x_i - \bar{x}$, is zero. Test out with all continuous data provided in Exercises 9 and 10.

2. Show that the variance of any variable s^2 can be expressed as

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n-1}.$$

   Test out this result against all data provided in Exercises 9 and 10.

3. Marks obtained by a class of students (out of 100) in Introduction to Statistics course are as follows:

68	34	56	23	45	78	67	96	76	45
75	34	89	92	50	47	56	53	19	45
34	23	76	45	67	43	56	78	94	74
38	58	52	64	72	59	55	92	88	53
71	49	53	89	48	58	54	39	66	62

   (a) Write the frequency table for interval of ten marks i.e. 1-10, 11-20 and so on.

   (b) Draw the histogram of the distribution.

   (c) Comment on the symmetry and peakedness of the distribution after calculating appropriate measures.

4. Calculate the value of commonly used statistics to find measure of spread for the runs scored by the Indian Cricket Team based on their scores in last 15 One Day Internationals while batting first.

281	307	251	429	241	189	256	194	267	385	228	299	247	331	389

5. A mobile phone company examines the ages of 150 customers to start special plans for them. The Frequency table is as follows:

Age(years)	0-14	15-19	20-29	30-39	40-49	50-79
Frequency	14	40	28	27	24	17

   (a) Draw the histogram for the data.

   (b) Estimate the mean age for these policyholders.

   (c) Estimate the median age for these policyholders.

6. A sample of 10 claims in an insurance company had mean and variance of 5,478 and 1,723 respectively. On reconciliation, it was found that one claim of 3,250 was wrongly written as 4,250. Calculate the sample mean and standard deviation of the sample with correct values.

7. Government of Tamilnadu wants to analyze the number of children in families for improving their immunization program. They analyze a group of 200 families and report their findings in the form of a frequency distribution shown below:

Number of children	0	1	2	3	4	5	6	7
Number of families	18	30	72	43	25	8	3	1

(a) Draw the bar chart for the following data and calculate the total number of children.

(b) Calculate mean, mode and median of the data.

(c) Calculate sample mean, sample variance and coefficient of variation of the data.

(d) Calculate Coefficient of kurtosis and Coefficient of skewness in the above data.

8. An insurance company wants to analyze the claims for damage due to fire on its household content's policies. Following are the values for a sample of 50 claims in Rupees.

57000	115000	119000	131000	152000	167000	188000	190000	197000	201000
206000	209000	213000	217000	221000	229000	247000	250000	252000	253000
257000	257000	258000	259000	260000	261000	262000	263000	267000	271000
277000	285000	287000	305000	307000	309000	311000	313000	317000	321000
322000	327000	333000	351000	357000	371000	399000	417000	433000	499000

The table given below displays the grouped frequency distribution for the above data.

Claim Size (in 1000's of Rupees)	Frequency
50-99	1
100-149	3
150-199	5
200-249	8
250-299	16
300-349	10
350-399	4
400-449	2
450- 500	1

(a) What is the range of the above data?

(b) Draw a bar graph for the above data

(c) For the data given above if instead of equal sized groups we had a single group for all value below 250, how would this bar be represented?

(d) Calculate the mean, median and mode for the above data.

(e) Calculate the sample geometric mean.

(f) Calculate the sample standard deviation and sample variance.

(g) Calculate the coefficient of variation for the above data.

9. The Dean of an important college wants to know some basic information about the height of the students of the last six year groups. To do so, twenty students were selected from each group and their height was measured, so a total of 120 observations divided into six groups. The obtained results are presented in Table 1. The Dean needs to know if there is any evidence that some students have the required height to play sports such as basketball, volleyball and swimming. On the other hand, he is looking for some useful information to plan some requirements such as uniforms, shoes and caps. Provide a descriptive report of the information found in the provided data. If possible, also provide graphics and recommendations.

10. The department of analysis of a taxicab company has the records of 15 drivers, which are shown in Table 2. Those records includes the following information: distance (Km), amount of crashes, amount of fines received per driver, amount of visits to garage for repairs (V.T.G), and days of operation (D.O.O).

    Provide a statistical report of the information provided by data. The department of analysis of the company is looking for information useful to identify good drivers and the efficiency of the fleet.

Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
175	172	162	174	177	176
174	179	182	170	185	162
190	170	166	154	177	183
187	177	175	168	179	158
181	186	192	186	191	169
173	183	199	200	183	184
200	188	168	164	171	166
189	169	188	184	178	170
173	178	187	182	182	171
186	179	178	164	170	182
176	163	196	169	183	177
170	171	158	184	154	144
177	178	190	175	152	164
186	198	165	177	173	180
178	184	159	167	189	179
154	203	174	165	190	174
192	193	160	194	174	185
188	174	177	160	182	148
185	175	181	186	183	188
194	172	170	162	155	187

Table 1: Measured height for Exercise 9

11. Consider Exercise 10. The department of analysis of the company suggests to compare all variables to each others using graphical analysis. The idea is to find some patterns whose information should be useful to identify possible relationships among variables. Provide only interesting graphics alongside with a description of their findings.

Km	Crash	Fines	V.T.G	D.O.O
13,381	0	0	1	240
12,200	0	0	0	240
30,551	1	1	2	240
26,806	0	3	0	240
11,984	0	0	0	240
32,625	0	1	0	240
60,308	1	1	2	240
24,683	0	0	0	240
8,167	0	0	1	360
11,198	0	0	0	360
26,120	0	0	0	360
186,632	2	4	3	360
7,147	1	2	2	360
18,731	0	1	1	360
2,129	2	2	2	360

Table 2: Recorded data for Exercise 10