## Statistics for Decision Making - I

Full Marks: 40 Time: 2 hrs 30 minutes

Answer 1, 2 and 3 and any two from the rest.

 Draw a suitable diagram for following data on tea production (in '000 metric tons) in different countries in a particular year: (4)

Country	Production
India	405
Sri Lanka	225
Japan	85
Indonesia	40
Other Countries	275

The following data pertains to the age distribution of the employees of a firm. Compute the value of one suitable measure for each of i) central tendency, ii) dispersion, iii) skewness. Draw the boxplot and comment. (3+3+2) + 4=12

Age	Below	20-25	25-30	30-35	35-40	40-45	45-55	55 and
(Years)	20							above
No. of	13	29	46	60	112	94	45	21
Employees								

(2)

3. i) Consider the following equations.

Y = 4 + 0.9x and X = 3.7 + 3.1y.

Can you consider these as regression equations? Give reasons.

ii) If at a later stage it is found that, the actual expression for the second equation is X = 3.7 - 0.31y, will your decision change? Give reasons. (2)

- 4. a) How do A.M., S.D. and correlation coefficient react to change of origin and scale? Show necessary derivations and make your comment. (5)
  - b) Write briefly about the scales of measurement. (5)
- 5. a) Prove that Spearman's rank correlation coefficient lies between -1 and +1. (6)b) Prove or disprove:

$$\frac{|b_{yx}| + |b_{xy}|}{r} \ge |r|, \text{ where the symbols have their usual meanings.}$$
(4)

a) 'Pearson's Correlation coefficient measures the strength of linear inter-relationship' - prove.
(6)

b) The A.M., median and mode of the marks of a group of 75 students in a test were calculated as 27, 29, 31. It was later discovered that, one observation was earlier wrongly recorded as 45 instead of 55. Examine, to what extent, the calculated values of the three averages will be affected by this? (4)