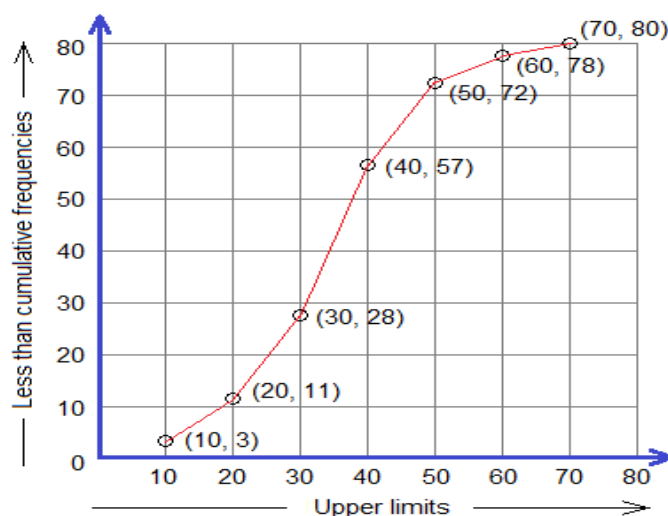# Assignment Mid Recovery 02

1. The following data represent the daily ticket sales at a park during three weeks.

    $$52, 73, 34, 85, 62, 79, 89, 50, 45, 83, 84, 91, 85, 84, 87, 44, 86, 41, 35, 73, 86$$

    (a). Construct a Stem-Leaf diagram to illustrate the data.
    (b). Use your diagram to find the median and mode of the data.
    (c). Find the interquartile range.
    (d). Comment about the skewness of the given data.
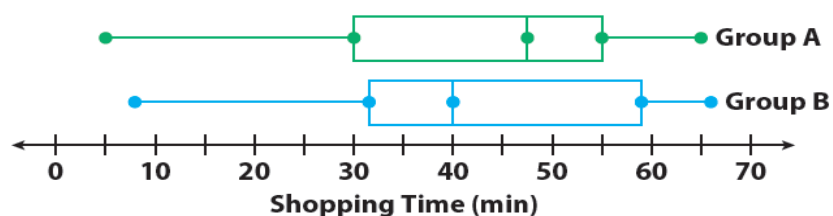   (e). Explain the kurtosis of the given data.

2. For the following Cumulative Frequency Polygon construct a Box-Whisker plot.

3. According to the following Box-Whisker plot write the relation among the Mode, Median, and Mean.
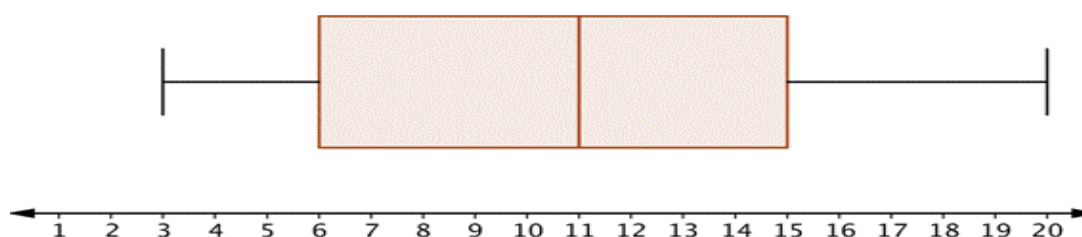

4. Distributions of the shopping time (in minutes) for two groups of people are given in the following Box-Whisker plot.

Assignment Mid Recovery 02

    (a). Find inter-quartile range (IQR) and hence investigate the consistency of the shopping time for the target groups.

    (b). Describe the nature (skewness) of the distributions.

5. Analyze the following dual Box-Whisker plot. You may comment about the central tendency, consistency, skewness, and kurtosis.

Height of Girls and Boys (inches)
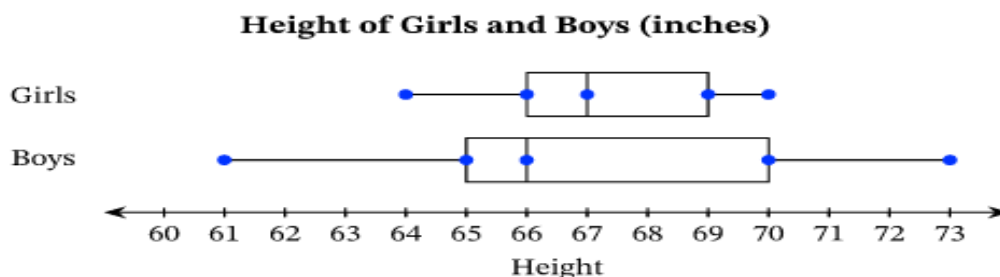
6. A survey was done to collect data on the number of hours per month that students in a class spent doing two activities: (i) playing computer games, (ii) watching TV. The data was put into back-to-back Stem-Leaf plot, as shown below.

Time Spent per Month by

    (a). If the target is to minimize the number of hours, which activity is better?

    (b). Which one of the activities is consistent?

    (c). Construct a Box-Whisker plot according to the given Stem-Leaf plot.

7. For the following sets of data design, a two-leaf Stem-Leaf diagram.

| Set A	33, 45, 21, 35, 41, 36, 37, 39, 35, 46, 47, 40, 20, 22, 35, 47, 34, 49, 23
Set B	34, 30, 31, 38, 42, 36, 27, 34, 37, 38, 19, 36, 36, 20, 32, 20, 30, 37, 38, 34, 41, 19

    (a). Find the average of the given sets.

    (b). Find the half way position of both of the data set and investigate the consistency.

    (c). Convert your Stem-Leaf diagram to the Box-Whisker plot.

    (d). Does your diagram have any outliers?

    (e). Could you find the modes of the data sets?

Assignment Mid Recovery 02

8. The back-to-back stem and leaf plot below shows the LDL cholesterol levels (in mg/dL) of two groups of people, smokers and non-smokers. The digits in the stem represent the hundreds and tens and the digit in the leaf represent the ones. So, for example 11|8 = 118 and so on.

smokers	stem	non smokers
	10	0 8
1	11	1 5 8 9
3 0	12	3 6 6 8 8 9 9
4 2 0	13	0 1 7 7 8 8 9
6 5 3 1	14	0 1 5 6 9 1
9 7 6 5 0	15	2 4 2
2 9 7 6 4 1	16	0 1 8
9 8 7 4 3	17	4 6
6 5 3 0	18	1
7 5 1	19	
5 0	20	

(a). People with a cholesterol level of 129 or less are said to have a near ideal level of cholesterol. How many people, in each group, have a near ideal level of cholesterol?

(b). People with a cholesterol level between 130 and 159 inclusive are said to be in the border high. How many people, in each group, are in the border high?

(c). People with a cholesterol level between 160 and 189 inclusive are said to have a high level of cholesterol. How many people, in each group, have a high level of cholesterol?

(d). People with a cholesterol level of 190 or above are said to have a very high level of cholesterol. How many people, in each group, have a very high level of cholesterol?

(e). Comparing the two groups, which group has more people with a higher level of cholesterol?

9. For the following Box-Whisker plot find the appropriate central tendency and check the consistency of the given samples.
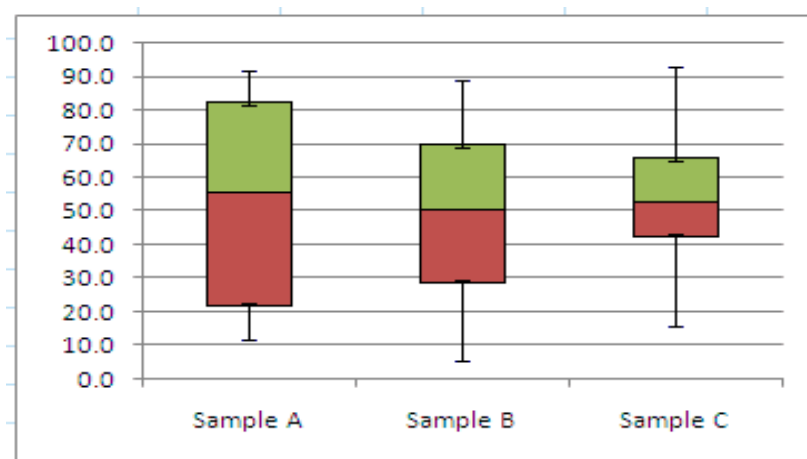
10. Compare the data sets represented through the Box-Whisker plots given below.
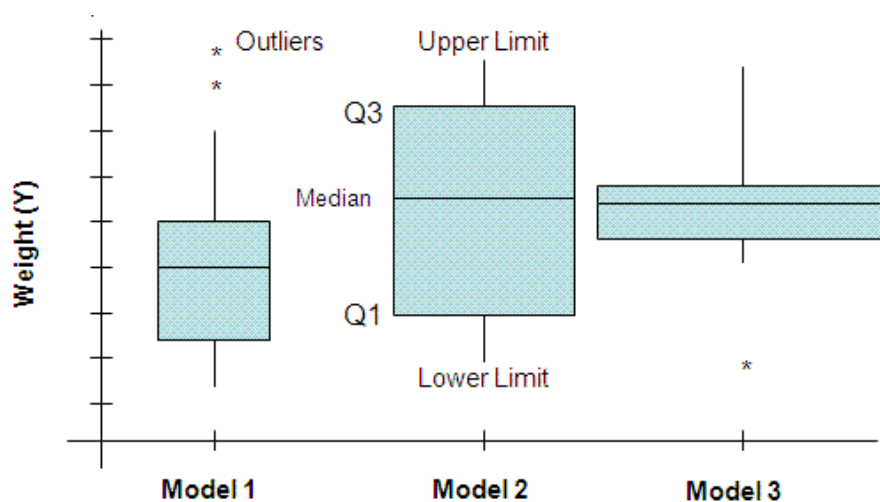
11. You are given that the lower quartile, median, and upper quartile of the weight of a county are 64, 74, 78, respectively. Draw the Box-Whisker plot of the given data. State the nature of the distribution of the given data.

12. Discuss the strength of correlation from the following Pearson' correlation coefficient.
    (i) $r = -0.15$ (ii) $r = -1$ (iii) $r = 0.65$ (iv) $r = 1$ (v) $r = 0$

13. If the correlation coefficient of two variables is 0.72 and regression coefficient of x on y is 1.08. If $\bar{x} = 29.2$ and $\bar{y} = 37.5$ find the regression line of y on x. Also, find the value of y when $x = 44$ graphically.

14. If the correlation coefficient of $x \& y$ is 0.75 and the corresponding standard deviations $1.25 \& 1.75$. Find the regression coefficient of y on x and x on y.

15. If the correlation coefficient of two variables is 0.65 and regression coefficient of y on x is 1.68. Also, $\bar{x} = 32.3$ and $\bar{y} = 45.6$.

    (a). Find the regression coefficient of x on y.
    (b). Find and sketch the regression line x on y.
    (c). Predict the value of x when y is 52. Also, verify your result graphically.

16. For the following data find the correlation co-efficient. How much y depends on x? Also, find the corresponding regression line.

x	5	12	18	23	27	30	26	22
y	18	16	13	11	9	7	10	13

Assignment Mid Recovery 02

17. Fit a least-squares line to the following set of data by using x as the dependent variable. Also, from the graph of least-squares line predict x for $y = 11$.

x	1	3	4	6	8	9	11	14
y	1	2	4	4	5	7	8	9

18. Find the rank correlation co-efficient between obtained places of 8 students in Mathematics and Physics.

Serial	1	2	3	4	5	6	7	8
Mathematics	3	1	6	5	7	4	8	2
Physics	8	3	1	2	6	5	4	7

19. Ten candidates were ranked as follows by two independent examiners, according to the score they obtained in an interview. Calculate the Spearman's rank correlation coefficient and interpret the result.

Candidate Number	1	2	3	4	5	6	7	8	9	10
Ranked by Ex. 1	7	9	1	3	8	4	10	5	6	2
Ranked by Ex. 2	9	5	1	4	6	7	8	2	10	3

20. A department store has the following statistics of sales (Y) for a period of 2 years of 10 salespersons who have varying years of experience (X) in sales promotion.

Experience (X) in Years	1	3	4	4	6	8	10	10	11	13
Average Annual sales (Y) in thousand	80	97	92	102	103	111	119	123	117	136

(a). Using the above set of data calculate the value of r (coefficient of correlation) and interpret the result.

(b). Find the regression line of Y on X in the form $Y = a + bX$.

(c). Sketch a scatter diagram in Years (X) vs Average Annual sales (Y).

(d). Verify your model found in question (b) with the tabular value for 6 years' experience.

(e). Predict the annual sales volume of persons what have 12 and 15-years' experience.

(f). Indicate why it may not be appropriate to use your question to predict the average annual sales at 30 years of experience.