SUDHARSAN SRINIVASAN

10017559 19

ASSIGNMENT - 2

8.1 Blocked Intrusion Attempts during 1st 2 weeks:

56, 47, 49, 37, 38, 60, 50, 43, 43, 59, 50, 56, 54, 58

After firewall settings changed:-

53, 21, 32, 49, 45, 38, 44, 33, 32, 43, 53, 46, 36, 48, 39, 35, 37, 36, 39, 45

Blocked intrusion during 1st 2 weeks, [sorted in Ascending Order]

Before change:

37, 38, 43, 43, 47, 49, 50, 50, 54, 56, 56, 58, 59, 60

After change: [sorted in ascending order]

21, 32, 32, 33, 35, 36, 36, 37, 38, 39, 39, 43, 44, 45, 45, 46, 48, 49, 53, 53

(a) Side-by-Side stem and leaf plots:

	2	1
7 8	3	2 2 3 5 6 6 7 8 9 9
3 3 7 9	4	3 4 5 5 6 8 9
0 0 4 6 6 8 9	5	3 3
0	6	

(b) Five-point Summaries:

The formula for five-point summary

$= (\min X_i, \hat{Q}_1, \hat{M}, \hat{Q}_3, \max X_i)$

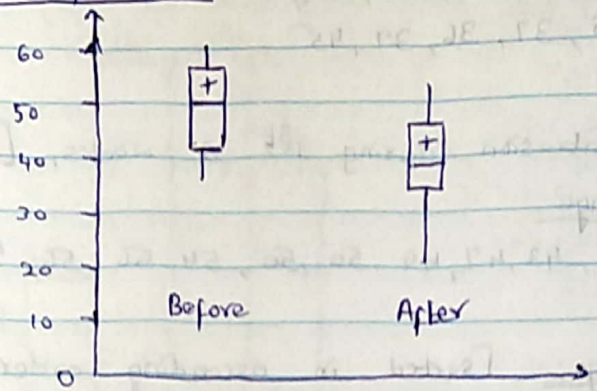
Before change: $\min X_i = 37$, $\max X_i = 60$, $\hat{Q}_1 = 43$, $\hat{Q}_3 = 56$, $\hat{M} = 50$

$\therefore (37, 43, 50, 56, 60)$

After change: $\min X_i = 21$, $\max X_i = 53$, $\hat{Q}_1 = 35$, $\hat{Q}_3 = 46$, $\hat{M} = 39$

$\therefore (21, 35, 39, 46, 53)$

Parallel Box-plots:



(c) Findings: The Mean value is decreased after changing the Firewall settings, since the Mean and Median values tend to be almost equal, thus reducing the number of Hacked Intrusions.

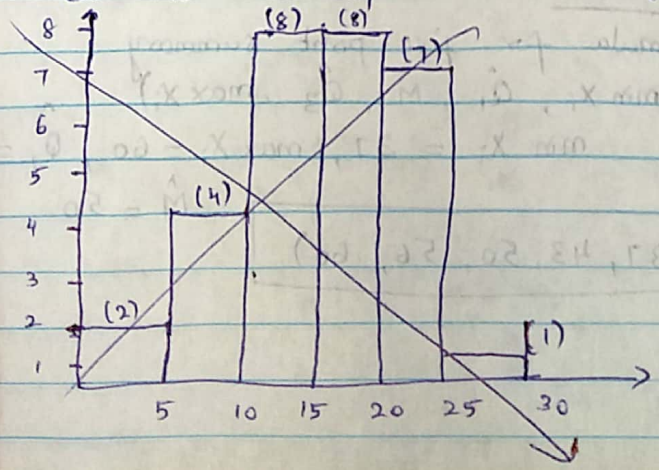
8.8. 3 Datasets:

Data-Set 1:

(1) 19, 24, 12, 19, 18, 24, 8, 5, 9, 20, 13, 11, 1, 12, 11, 10, 22, 21, 7, 16, 15, 15, 26, 16, 1, 13, 21, 21, 20, 19

In the dataset, we have (1-5) → 3 samples)
(6-10) → 3+1 = 4
(11-15) → 8
(16-20) → 8
(21-25) → 6
(26-30) → 1

30 samples



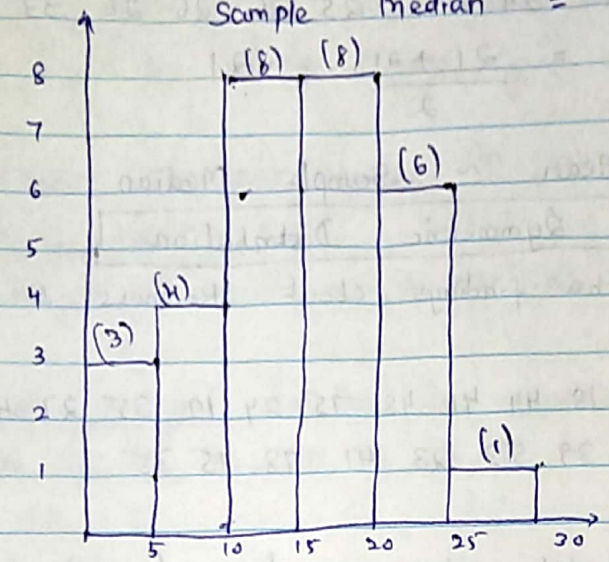
(b) Here, Sample Mean $= \frac{449}{30} = 14.96667 = \boxed{14.97}$

Sample Median (Sort the samples and compute)

1 1 5 7 8 9 10 11 11 12 12 13 13 15 15 16 16 18 19 19 19 20 20 21 21 21 22 24 24 26

Sample median $= \frac{15+16}{2} = \boxed{15.50}$

(a)



Sample Mean < Sample Median

Left - Skewed

Yes they support the findings about Skewness and Symmetry.

(2) (Data-Set 2)

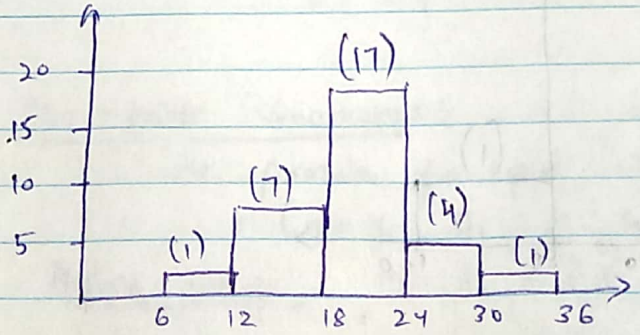
17, 24, 21, 22, 26, 22, 19, 21, 23, 11, 19, 14, 23, 25, 26, 15, 17, 26, 21, 18, 19, 21, 24, 18, 16, 20, 21, 20, 23, 33

In the above dataset, we have

1-6 => 0 samples
7-12 => 1
13-18 => 7
19-24 => (17)
25-30 => 4
31-36 => 1

30 samples

(a)



The curve is Symmetric

(b) Sample Mean $= \frac{625}{30} = 20.83$

Sample Median → sort the samples & calculate

11 14 15 16 17 17 18 18 19 19 19 20 20 21 | 21 21 | 21 21
22 22 23 23 23 24 24 25 26 26 26 33

$$\text{Median} = \frac{21 + 21}{2} = 21$$

Sample Mean ≃ Sample Median

∴ Symmetric Distribution

Yes they support the findings about Skewness & Symmetry

(3) Data - Set 3:

56 52 13 34 33 18 44 41 48 75 24 19 35 27 46 62 71 24
66 94 40 18 15 39 53 23 41 78 15 35

10,

In the above data-set, we have (1-20) → 6 samples

21-40 → 10

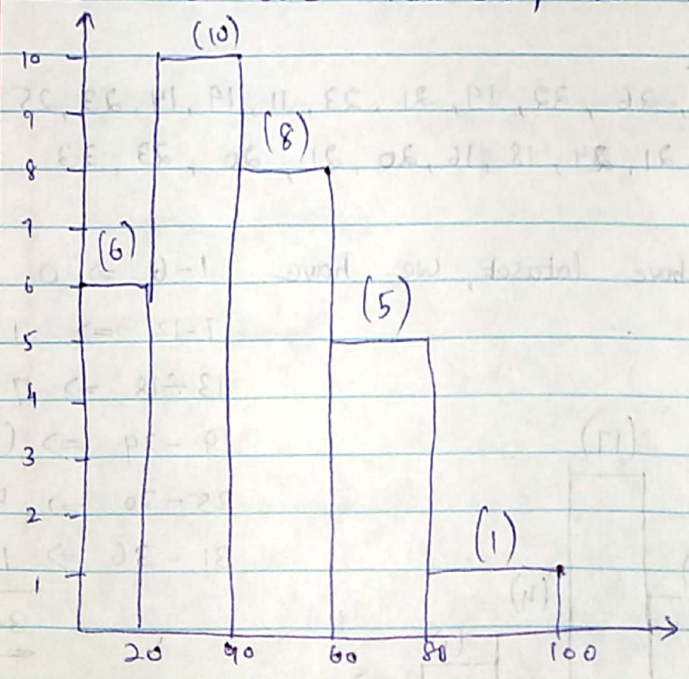
41-60 → 8

61-80 → 5

81-100 → 1

30 samples

(a)



(b) Sample Mean $= \frac{1239}{30} = 41.3$

Sample Median ⇒ sort the samples & calculate

13 15 15 18 18 19 23 24 24 27 33 34 35 38 | 39 40 |
41 41 44 46 48 52 53 56 62 66 71 75 78 94

Sample Median $= \frac{39+40}{2} = 39.5$

Sample Mean > Sample Median

∴ Right-Skewed

Yes they support the findings about Skewness & Symmetry.

Q3. Input Samples:

69, 47, 175, 70, 53, 64, 74, 52, 58, 53, 64, 49, 70, 65, 70, 16, 67, 55, 42, 72, 61, 65, 77, 70, 60, 39

(i) Sample Mean $= \frac{1657}{26} = 63.73077 \Rightarrow 63.73$

Sample Variance $= (x - \bar{x})^2$

* $(x - \bar{x})^2$

$= 691.484$

Standard deviation $= \sqrt{\text{Variance}}$

$= \sqrt{691.484}$

$= 26.29609$ (or) 26.2961

(ii) Parameters of Normal Distribution (μ and σ)

In case of Normal Distribution, Sample Mean is equal to Population Mean

Sample Mean = Population Mean

$$\bar{X} = \mu$$

$$\boxed{\mu = 63.73}$$

Standard deviation (for Normal distribution) $= \frac{\sigma}{\sqrt{n}}$

$$S.D. = \frac{\sigma}{\sqrt{n}}$$

$$\sigma = S.D \times \sqrt{n}$$

$$= 26.2961 \times \sqrt{26}$$

$$\boxed{\sigma = 134.0843}$$

(iii) Eliminate Outliers:

In the given sample, outliers are 16 and 175

Eliminating 16 and 175 from the sample data we calculate sample mean, variance and standard deviation

(a) Sample Mean $= \frac{1466}{24} = 61.083$

Sample Variance $= (X - \bar{X})^2$

$= 106.94$

Standard deviation $= \sqrt{\text{Variance}} = \sqrt{106.94}$

$= 10.3416$

(b) Parameters of Normal Distribution (μ and σ)

Incase of Normal distribution,

Sample Mean = Population Mean

$$\bar{x} = \mu$$

$$\boxed{\mu = 61.083}$$

Standard deviation $= \frac{\sigma}{\sqrt{n}}$

$$S.D = \frac{\sigma}{\sqrt{n}}$$

$$\sigma = S.D. \times \sqrt{n}$$

$$= 10.3416 \times \sqrt{24}$$

$$\boxed{\sigma = 50.6632}$$

(c) Since the population follows Normal distribution, eliminating the Outliers did improve the accuracy of the estimate.