Question: CGPA of 30 students are given below:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 3.60 | 3.17 | 3.90 | 2.53 | 3.23 | 3.11 | 3.34 | 2.96 |
| 2.82 | 3.39 | 3.01 | 2.71 | 2.80 | 2.73 | 2.79 | 2.82 |
| 3.06 | 2.53 | 3.21 | 3.24 | 3.31 | 2.64 | 2.65 | 2.51 |
| 2.79 | 3.32 | 3.08 | 2.67 | 3.35 | 2.75 | | |

Now answere the following question.

a) What is the mean of CGPA of students?

b) Find the variance of CGPA of students.

c) Find the 5-number summary for this data.

d) Construct a boxplot for the given data.

e) Construct a stem & leaf plot for this data.

f) Determine if there is any outlier CGPA and which CGPA is outlier?

Solution:

a) There are three types of mean. such as

i) Sample or Arithmetic mean

ii) Weighted mean

iii) Geometric mean

i) Sample or Arithmetic mean:

summation of CGPA of students, $\Sigma X$

$= 3.60 + 3.17 + 3.90 + --- + 2.67 + 3.35 + 2.75$

$= 90.02$

number of students, $n = 30$

$\therefore$ Sample or Arithmetic mean, $\bar{X} = \frac{\Sigma X}{n}$

$= \frac{90.02}{30}$

$= 3.00067$

ii) Weighted mean :

This ~~is~~ data doesn't have weight for each element. So, weighted mean can not be determined for this kind of data.

iii) Geometric mean:

Geometric mean. $= \sqrt[30]{3.6 \times 3.17 \times --- \times 3.35 \times 2.75}$

$= 2.9822$

b) Varciance :

Arithmetic mean, $\bar{x} = 3.00067$

summation of square of deviation from mean,

$$\Sigma (x-\bar{x})^2 = (3.60-3.00067)^2 + (3.17-3.00067)^2 + \cdots + (3.35-3.00067)^2 + (2.75-3.00067)^2$$

$$= 3.42102$$

| $x$ | $x-\bar{x}$ | $(x-\bar{x})^2$ | $x$ | $x-\bar{x}$ | $(x-\bar{x})^2$ |
|---|---|---|---|---|---|
| 3.60 | 0.59933 | 0.35919 | 2.82 | -0.18067 | 0.03264 |
| 3.17 | 0.16933 | 0.02867 | 3.06 | 0.05933 | $3.52\times10^{-3}$ |
| 3.90 | 0.89933 | 0.80879 | 2.53 | -0.47067 | 0.22153 |
| 2.53 | -0.47067 | 0.22153 | 3.21 | 0.20933 | 0.04382 |
| 3.23 | 0.22933 | 0.05259 | 3.24 | 0.23933 | 0.05727 |
| 3.11 | 0.10933 | 0.01195 | 3.31 | 0.30933 | 0.09567 |
| 3.34 | 0.33933 | 0.11514 | 2.64 | -0.36067 | 0.13008 |
| 2.96 | -0.04067 | $1.654\times10^{-3}$ | 2.65 | -0.35067 | 0.12297 |
| 2.82 | -0.18067 | 0.03264 | 2.51 | -0.49067 | 0.24076 |
| 3.39 | 0.38933 | 0.15158 | 2.79 | -0.21067 | 0.04438 |
| 3.01 | $9.33\times10^{-3}$ | $8.705\times10^{-5}$ | 3.32 | 0.31933 | 0.10197 |
| 2.71 | -0.29067 | 0.08449 | 3.08 | 0.07933 | $6.29\times10^{-3}$ |
| 2.80 | -0.20067 | 0.04027 | 2.67 | -0.33067 | 0.10934 |
| 2.73 | -0.27067 | 0.07326 | 3.35 | 0.34933 | 0.12203 |
| 2.79 | -0.21067 | 0.04438 | 2.75 | -0.25067 | 0.06283 |
| | | | | $\Sigma(x-\bar{x})^2 = 3.42102$ | |

$\therefore$ variance, $s^2 = \frac{1}{n-1} \sum (x-\bar{x})^2$

$= \frac{3.42102}{30-1}$

$= 0.117966$

c) 5-number summary:

The 5-number summary is composed of:

1. L, the smallest value in the data set
2. $Q_1$, the first quartile
3. $\tilde{x}$, the median ($2^{nd}$ quartile)
4. $Q_3$, the third quartile
5. H, the largest value in the data set

Sorted Data:

| | | | | | | |
|---|---|---|---|---|---|---|
| 2.51 | 2.53 | 2.53 | 2.64 | 2.65 | 2.67 | 2.71 |
| 2.73 | 2.75 | 2.79 | 2.79 | 2.8 | 2.82 | 2.82 |
| 2.96 | 3.01 | 3.06 | 3.08 | 3.11 | 3.17 | 3.21 |
| 3.23 | 3.24 | 3.31 | 3.32 | 3.34 | 3.35 | 3.39 |
| 3.60 | 3.90 | | | | | |

1. The smallest value in the data set, $L = 2.51$

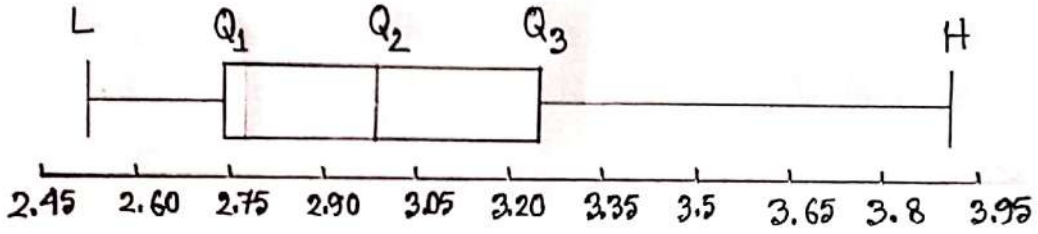
2. The first quartile, $Q_1 = 2.73$

3. The median ($2^{nd}$ quartile), $\tilde{x} = 2.985$

4. The third quartile, $Q_3 = 3.24$

5. The largest value in the data set, $H = 3.90$

d) Boxplot:




Boxplot for weight data

e) Stem & leaf display:

~~Use~~ All the speeds are in the 2.5s to 3.9s. Use first digit, decimal point and digit after decimal point of each speeds as the stem and the second digit after decimal point as the leaf. Draw a vertical line and list the stems in order to the left of the line. Place each leaf on its stem: place the trailing digit on the right side of the vertical line opposite its corresponding leading digits.

30 speeds

| | |
|---|---|
| 2.5 | 1 3 3 |
| 2.6 | 4 5 7 |
| 2.7 | 1 3 5 9 9 |
| 2.8 | 0 2 2 |
| 2.9 | 6 |
| 3.0 | 1 6 8 |
| 3.1 | 1 7 |
| 3.2 | 1 3 4 |
| 3.3 | 1 2 4 5 9 |
| 3.4 | |
| 3.5 | |
| 3.6 | 0 |
| 3.7 | |
| 3.8 | |
| 3.9 | 0 |

f) Checking for outliers:

- First, we need to arrange the data set in order. The ordered data set is in c.
- Next, determine the first and third quartiles.

  $Q_1 = 2.73$ and $Q_3 = 3.24$

- Thus the $IQR = 3.24 - 2.73 = 0.51$
- Now, $Q_1 - (1.5 \times IQR) = 2.73 - (1.5 \times 0.51) = 1.965$
- And, $Q_3 + (1.5 \times IQR) = 3.24 + (1.5 \times 0.51) = 4.005$

As there is no value less than 1.965 or greater than 4.005. in the data set, so there is no outliers in the data set.