Q.1. How can we figure out what the interquartile range is?

→ Order the data from least to greatest
→ Find the median
→ Calculate the median of both the lower and upper half of the data.
→ The IQR is the difference between the upper and lower medians.

⇒ Order the data from least to greatest :-

In order to calculate the IQR, we need to begin by ordering the values of the data set from the least to the greatest. Likewise, in order to calculate the median, we need to arrange the numbers in ascending order.

Quarter 1 (Q_1) - 25%.
Quarter 2 (Q_2) - 50%. - median
Quarter 3 (Q_3) - 75%.

Dataset :- 1, 2, 2, 2, 3, 3, 4, 5, 5, 5, 6, 6, 6, 6, 7, 8, 8, 9, 27

ii) calculate the median

Next, we need to calculate the media
The median is the center of the data
if the data set has an odd number
of data points, then the mean is
centermost number. On the other hand,
if the data set has an even number
of values, then we will need to take
the arithmetic average of the two
centermost values. We will calculate the
average by adding the two numbers
together and then dividing that number
by two.

$$Q_2 \text{ (median)} - 10th \text{ position value}$$
$$- 5$$

iii) calculate upper & lower limit :-
$$Q_1 (25\%) = \frac{\text{Percentile}}{100} \times (n+1)$$

$$= \frac{25}{100} \times (19+1)$$

$$= \frac{25}{100} \times 20 = 5th \text{ index}$$
$$= 3 \text{ (value)}$$

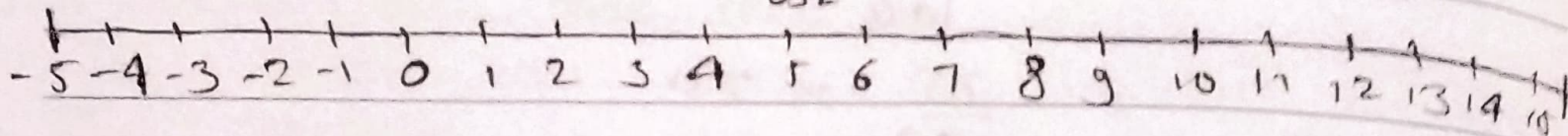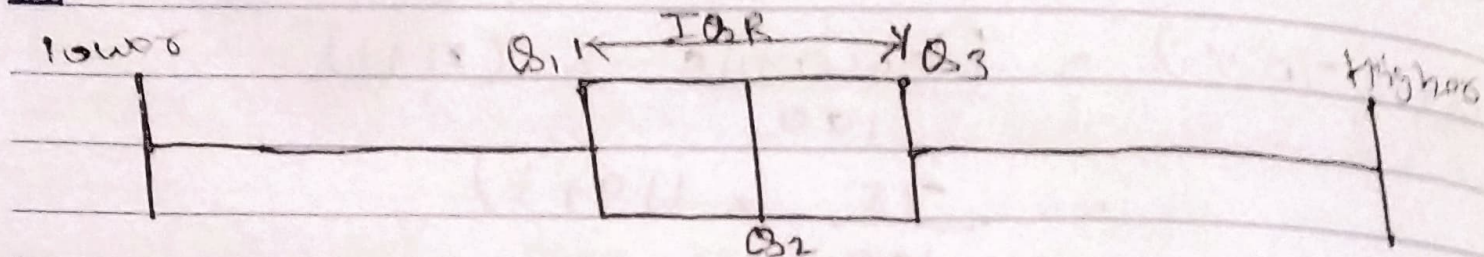$$Q_3 \ (70\%) = \frac{Percentile}{100} \times (n+1)$$

$$= \frac{75}{100} \times (19+1)$$

$$= \frac{75}{100} \times 20$$

$$= 15^{th} \ index$$

$$\Rightarrow 7$$

iv) Calculate the difference:-

$$IQR = Q_3 - Q_1$$

$$= 7 - 3$$
$$= 4$$

Lower fence:- $Q_1 - 1.5 \times IQR$

$$= 3 - 1.5 \times 4$$
$$= 3 - 6$$
$$= -3$$

Higher fence $= Q_3 + 1.5 \times IQR$

$$= 7 + 1.5 \times 4$$
$$= 7 + 6 \ = 13$$

lower

Q_1 |← IQR →| Q_3

higher

Q_2

-5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

2022 FEBRUARY
M T W T F S S M T W T F S S
1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20 21 22 23 24 25 26 27
28

013-352 • WK-03

JANUARY • THURSDAY

13

2022

82. What enactly is the value of the 5-number Summary?

⇒ There are a variety of descriptive statistics. Numbers such as mean, median, mode, standard deviation, first quartile, third quartile etc. rather than looking at these descriptive statistics individually. Sometimes combining them helps us to give us a complete picture. This is the reason we use five number summary. Which includes :–

i) Lower fence:-
The Smallest value in our data set.

ii) First Quartile:-
The number is denoted as Q_1 and 25% of our data falls under it.

iii) Median:- This is the midway point of the data (exactly in between)

iv) Third Quaotile:-.

This number is denoted by Q_3 and 75% of data falk below it.

v) Maximum:- This is the maximum value in our data set.

2022   FEBRUARY

M T W T F S S M T W T F S S
1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 26 21 22 23 24 25 26 27
28

015-350 • WK-03

JANUARY • SATURDAY

15

2022

Q3. What is the relationship between standard deviation and variance?

→ The relationship between the variance and the standard deviation for a sample data set is given below.

i) variance represents the average squared deviations from the mean value of data, while standard deviation represents the square root of that number.

ii) Both, the variance and the standard deviation measures variability in a distribution.

iii) Both have different units like the standard deviation has the same units as the original values like minutes or meters while the variance has much large units like meter squared.

iv) The variance is equal to the square of the standard deviation or the standard deviation is square root of variance.

Q1. What does the difference between variance and standard deviation mean?

→ Variance is average squared difference from mean, whereas Standard deviation is the squared root of the variance.

: variance measures Dispersion within the Data set, Standard deviation spread around the mean.

: variance is not sub-additive, Standard deviation is a measure of spread for symmetrical distributions with no outlier

: variance measures the volatile nature of data of a population. whereas Standard deviation, in finance, is often called volatility.

: variance measures how far the outcome varies from the mean. Standard deviation measures how far the normal standard deviation is from the expected value.

2022 FEBRUARY

M T W T F S S M T W T F S S
1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 20 21 22 23 24 25 26 27
28

018-347 • WK-04

JANUARY • TUESDAY

18

2022

: In finance, it helps to measure the actual deviation to performance from the standard, Standard deviation is a useful tool to take a decision regarding the investment in stocks, mutual fund etc.
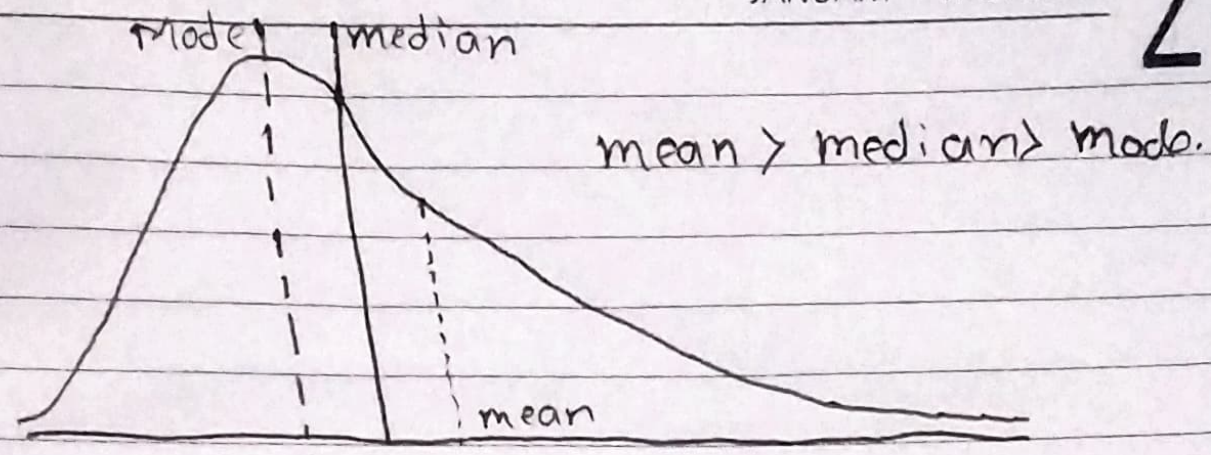
019-346 • WK-04

2022

19 JANUARY • WEDNESDAY

2022

JANUARY
M T W T F S S M T W T F S S
1 2 3 4 5 6 7 8 9
10 11 12 13 14 15 16 17 18 19 20 21 22 23
24 25 26 27 28 29 30 31

Q.5: when is it appropriate to refer to a skewed data distribution?

→ A distribution is said to be skewed when the data points cluster more towards one side of the scale than the others, creating a curve that is not symmetrical. In other words, the right and the left side of the distribution are shaped differently from each other. There are two types of skewed distribution.

A distribution is positively skewed if the scores fall towards the lower side of the scale and there are very few higher scores. Positively data skewed is also reffered to as skewed to the right because that is the direction of the long tail end of the chart.
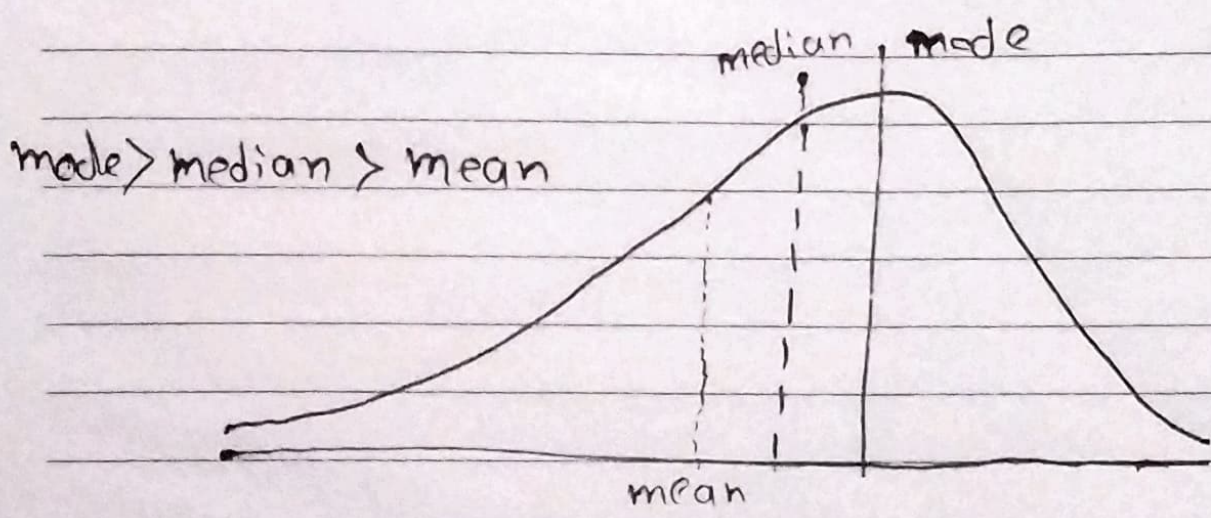
A distribution is negatively skewed it the scores fall towards the higher side of the scale and there are few low scores.

2022
FEBRUARY

M T W T F S S M T W T F S S
1 2 3 4 5 6 7 8 9 10 11 12 13
14 15 16 17 18 19 26 21 22 23 24 25 26 27
28

020-345 • WK-04

2022

JANUARY • THURSDAY

20

Mode   median

mean > median > mode.

mean

Right   Skewed   Distribution

Eg:- wealth Distribution
      Length of Comments etc

median, mode

mode > median > mean

mean

left skewed   Distribution

Eg:- Lifespan of humans.