Measure of Variability

Measure of variability is also known as measure of dispersion and used to describe variability in a sample or population. In statistics, there are three common measures of variability.

(a) ~~Range~~ Range

It is given measure of how to spread apart values in sample (set), or data set.

Range = max value - min value

```
data = np.array([4,5,5,5,6,6,6,6,7,7,7,8])
range = data.max() - data.min()
print(rang) → 4.
```

How to find range in any file.

~~range = y.column~~

```
range = (dataframe.columnname).max() - (dataframe.
                                        columnname).min())
```

(b) Variance

It simply describes how much a random variable defers from expected value and it is also computed as square of deviation.

$$S^2 = \sum_{j=1}^{n} [(x_i - \bar{x})^2 \div n]$$

In these formula, n represent total data points $\bar{x}$ represent mean of data points and x_i represent individual data points.

(c) Dispersion

It is measure of dispersion of set of data from its mean

$$\sigma = \sqrt{(1/n) \sum_{i=1}^{n} (x_i - \mu)^2}$$

μ = population mean.
$\bar{x}$ = sample mean

Dataset A

⊙ Calculate the population mean (μ) of Dataset A.

$(4+5+5+5+6+6+6+6+7+7+7+8)/12$

mean $(\mu) = 6$

⊙ Calculate the deviation of the individual values from the mean by subtracting the mean from each value in the dataset.

$-2, -1, -1, -1, 0, 0, 0, 0, 1, 1, 1, 2$

⊙ Square each individual deviation value.

$4, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 4$

⊙ Calculate the mean of the squared deviation values.

$(4+1+1+1+0+0+0+0+1+1+1+4)/12$

$= \frac{14}{12} = 1.1666$

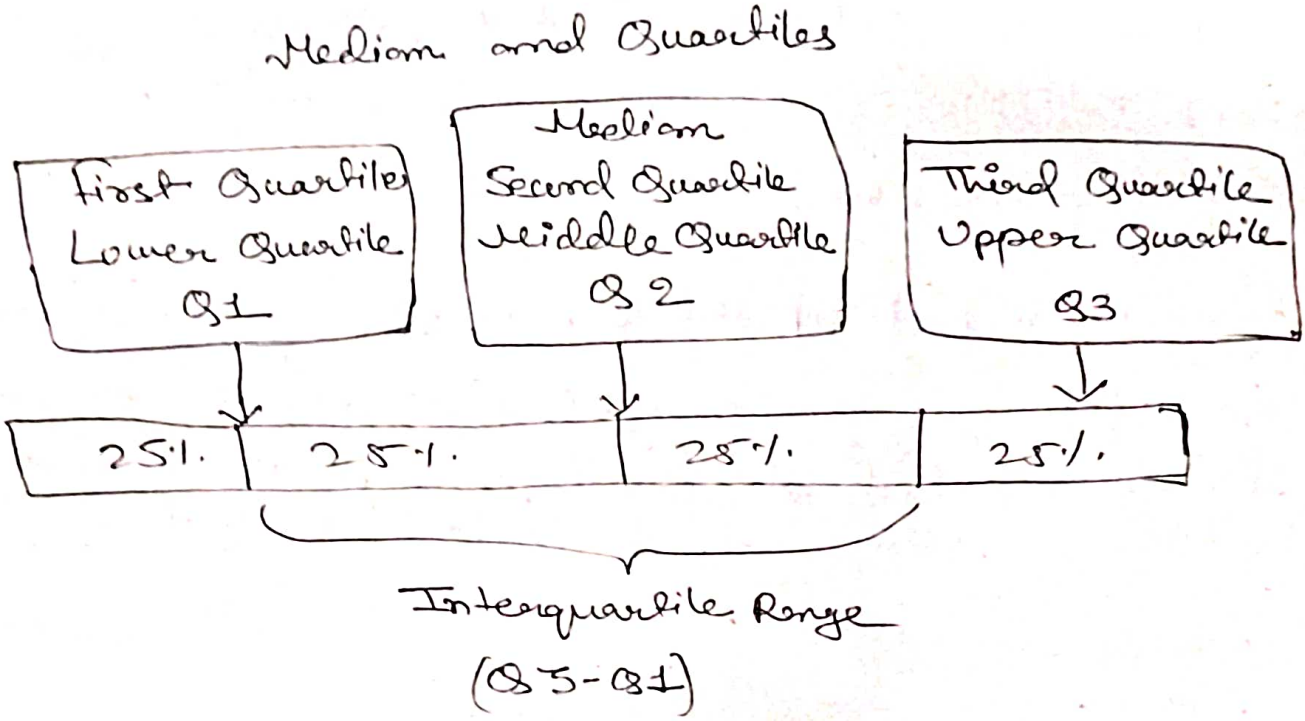
$= 1.17$

Variance $\sigma^2 = 1.17$.

⊙ Calculate the square root of the variance.

Standard deviation $\sigma = \sqrt{1.17} = 1.08$

Quartile ..



Interquartile range..

The interquartile range (IQR) is the difference between the upper (Q3) and lower (Q1) quartiles, and describes the middle 50% of values when ordered from lowest to highest.

The IQR is often seen as a better measure of spread than the range.

The IQR for Dataset A is = 2

$$IQR = Q3 - Q1$$
$$= 7 - 5$$
$$= 2$$

```
data = np.array([4,5,5,5,6,6,6,6,7,7,7,8])

q1 = np.percentile(data, 25)
q2 = np.percentile(data,50)
q3 = np.percentile(data, 75)

IQR = q3-q1
print(IQR)
2.0

IQR= q3 - q2
   = 1.0
```