MEASURES OF VARIABILITY

DESCRIPTIVE STATISTICS

prepared by:

Gyro A. Madrona

Electronics Engineer

TOPIC OUTLINE

Measures of Variability

Range and Interquartile Range

Variance and Standard Deviation

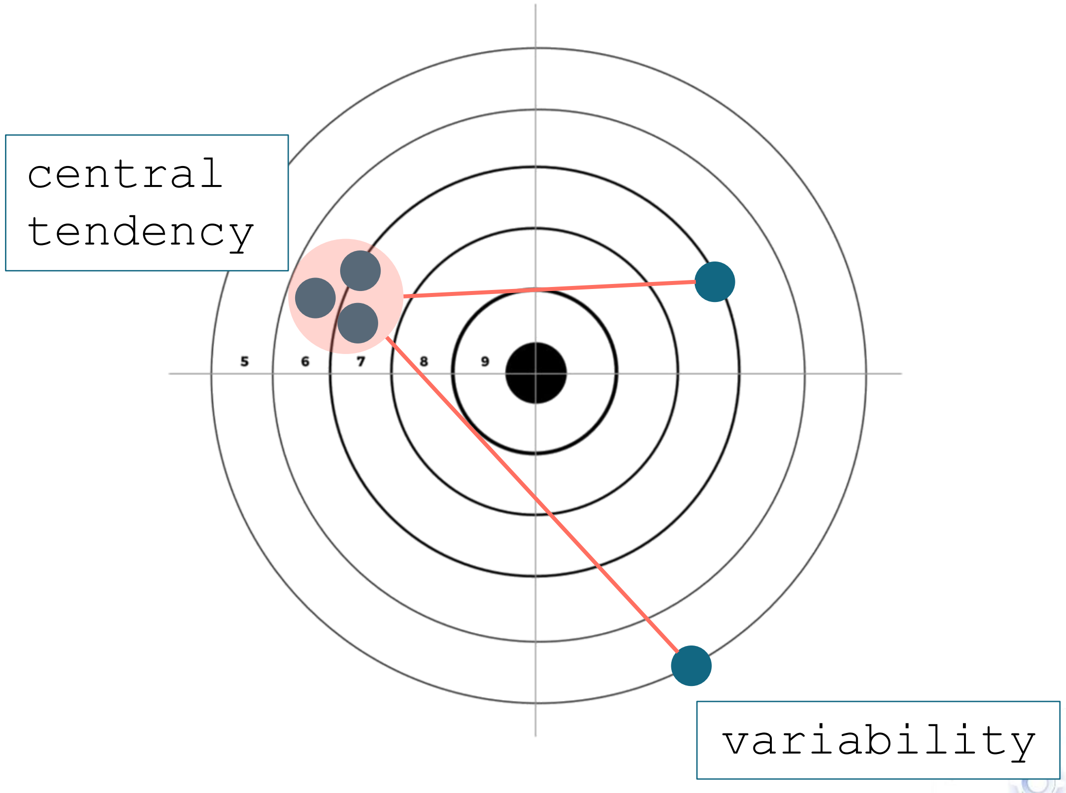
Coefficient of Variation

MEASURES OF VARIABILITY

MEASURES OF VARIABILITY

Measures of variability (or dispersion) describe how **spread out** or scattered a dataset is. These measures provide insights into the consistency of data points relative to the central tendency (mean, median, or mode).

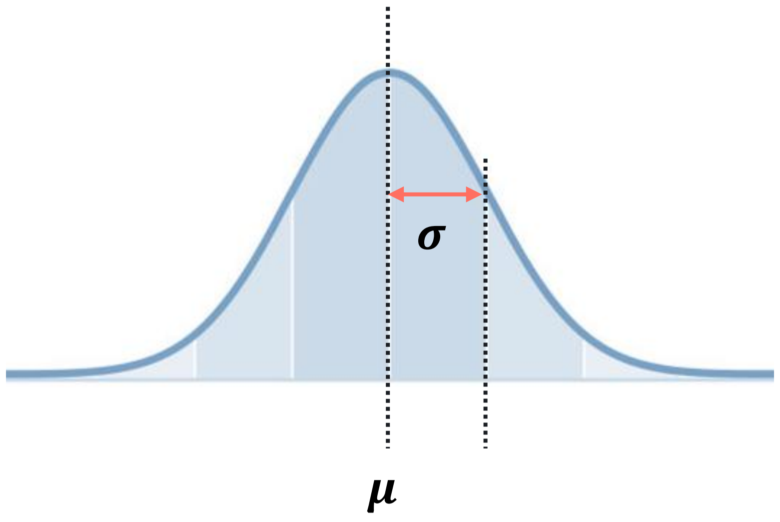
Dartboard Analogy:



MEASURES OF VARIABILITY

Measures of variability (or dispersion) describe how **spread out** or scattered a dataset is. These measures provide insights into the consistency of data points relative to the central tendency (mean, median, or mode).

Normal Distribution:



RANGE AND INTERQUARTILE RANGE

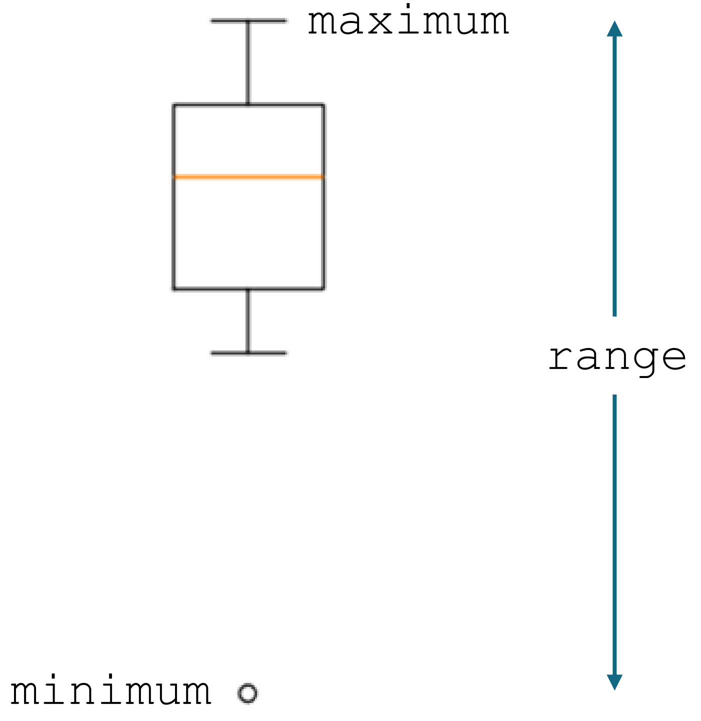
RANGE

The **range** is the simplest measure of variability and is calculated as the **difference** between the maximum and minimum values in a dataset.

Formula:

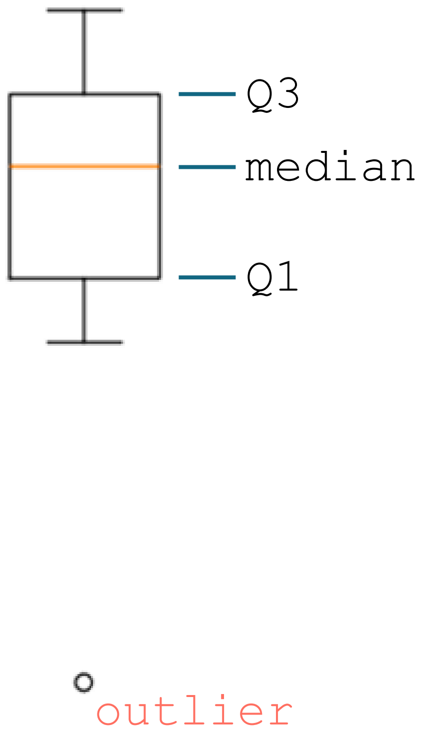
$$\boldsymbol{range = maximum\ value - minimum\ value}$$

Boxplot:



INTERQUARTILE RANGE

Boxplot:



The **interquartile range (IQR)** measures the spread of the middle 50% of the data, reducing the influence of outliers.

Formula:

$$IQR = Q_3 - Q_1$$

where:

Q_1(first quartile) is the median of the lower half of the data (25%).

Q_3(third quartile) is the median of the upper half of the data (75%).

EXERCISE

The dataset provided contains the exam grades of 12 students. Calculate the **range** and **interquartile range** (IQR) to analyze the spread and variability of the grades.

Exam Performance

Student	Grade
1	3.5
2	6.7
3	7
4	7.4
5	7.8
6	8.2
7	8.5
8	8.8
9	9
10	9.1
11	9.4
12	9.8

VARIANCE AND STANDARD DEVIATION

VARIANCE

Population Variance:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}$$

Variance measures the **average squared deviation** of each data point from the mean.

Sample Variance:

$$s^2 = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{n - 1}$$

STANDARD DEVIATION

The **standard deviation** is the **square root** of variance.

Population Variance:

$$\sigma = \sqrt{\sigma^2}$$

Sample Variance:

$$s = \sqrt{s^2}$$

EXERCISE

The dataset provided contains the sugar content (in grams) per serving for 10 popular breakfast cereals. Calculate the **variance** and **standard deviation** to measure the spread or variability in the sugar content across these cereals.

Breakfast Cereal

Brand	Sugar
A	12
B	9
C	15
D	8
E	10
F	11
G	13
H	7
I	14
J	6

POOLED STANDARD DEVIATION

Variances add:

$$\sigma_{total}^2 = \sigma_1^2 + \sigma_2^2 + \cdots \sigma_n^2$$

Pooled standard deviation is a weighted **average** of the standard deviations from two or more groups.

Standard deviations do not:

$$\sigma_{total} \neq \sigma_1 + \sigma_2 + \cdots \sigma_n$$

Formula:

$$\bar{\sigma}_{pooled} = \sqrt{\overline{\sigma^2}}$$

where:

$$\overline{\sigma^2} = \frac{\sum_{i=1}^{n} \sigma_i^2}{n}$$

EXERCISE

The dataset provided contains the battery life (in hours) for smartphones from different models. Calculate the **pooled standard deviation** to measure the combined variability in battery life across these models.

Battery Life

Model	Hours
A	12.5
A	12.8
A	12.7
A	13.3
A	12.6
B	13.5
B	14.1
B	13.9
B	14.3
B	13.7
C	11.8
C	11.9
C	12.1
C	12.2
C	11.6

COEFFICIENT OF VARIATION

COEFFICIENT OF VARIATION

Coefficient of variation (c_v) is a relative measure of variability, expressed as the ratio of the standard deviation to the mean.

Population Coefficient of Variation:

$$c_v = \frac{\sigma}{\mu}$$

Sample Coefficient of Variation:

$$\widehat{c_v} = \frac{\sigma}{\bar{x}}$$

EXERCISE

The provided dataset includes ice cream prices listed in both USD and PHP. Calculate the **standard deviation** and **coefficient of variation** for each currency to analyze the variability in prices.

Ice Cream Price List

Ice Cream	Price (USD)	Price (PHP)
Brand A	3.5	203
Brand B	4	232
Brand C	3.75	217.5
Brand D	4.25	246.5
Brand E	3.9	226.2
Brand F	4.1	237.8
Brand G	3.6	208.8
Brand H	4.5	261
Brand I	3.8	220.4
Brand J	4.15	240.7

LABORATORY

LABORATORY

The given dataset consists of test results from two machines, the Jaguar and Panther models, which produce 10 Ω resistors with ±5% tolerance. Determine which machine performs better based on its **measures of variability** for resistance values.

Resistance Test

Test No.	Jaguar	Panther
1	10.6	10.1
2	9.1	11
3	9.3	9.1
4	9.8	20
5	10.5	9.2
6	10.4	10.8
7	9.5	9.9
8	11	9.2
9	10.4	9.1
10	3	9.1
11	9.8	