

Submitted by: Adithya Acharya

USN : 01SU24CS011

Branch: CSE

Section : A section

Subject: Fundamentals of AI and ML

Topic : Assignment 3

Submitted to : Prof. Mahesh Kumar

ASSIGNMENT (Question – 3)

If $\mu = 55$, $\sigma_a = 4$, $\sigma_b = 10$, $\sigma_c = 15$, In this which is better?

1. Introduction

In statistics, the mean (μ) and standard deviation (σ) are important measures used to describe a dataset.

The mean represents the average value of the data.

The standard deviation measures how much the data values deviate from the mean.

It shows the spread or variability in the dataset.

When comparing multiple datasets with the same mean but different standard deviations, the dataset with the lower standard deviation is considered more consistent and stable.

In this problem:

$$\mu = 55$$

$$\sigma_a = 4$$

$$\sigma_b = 10$$

$$\sigma_c = 15$$

We need to determine which one is better.

2. Understanding Mean (μ)

The mean is the average value of all observations in a dataset.

Here, all three cases have the same mean:

$$\mu = 55$$

This means the average value of all three datasets is 55.

Since the mean is equal in all cases, the comparison depends entirely on the standard deviation.

3. Understanding Standard Deviation (σ)

Standard deviation measures the dispersion of data around the mean.

- Small standard deviation → Data values are close to the mean.
- Large standard deviation → Data values are widely spread from the mean.

A lower standard deviation indicates:

- Less variation
- More consistency
- More reliability
- Stable performance

A higher standard deviation indicates:

- More variation
- Less consistency
- Greater fluctuations
- Unstable performance

4. Comparison of the Three Cases

Case A: $\sigma_a = 4$

- Very small deviation from the mean
- Data values are tightly clustered around 55
- High consistency

- Less variation
- More stability

This indicates that most values are close to 55.

Case B: $\sigma_B = 10$

- Moderate deviation from the mean
- Data values are somewhat spread out
- Medium consistency
- Moderate variation

Values are less consistent compared to Case A.

Case C: $\sigma_C = 15$

- Large deviation from the mean
- Data values are widely spread
- Low consistency
- High variation
- Less stability

Values fluctuate significantly around 55.

5. Which is Better?

Since all three datasets have the same mean (55), the best dataset is determined by the smallest standard deviation.

$$\sigma_a = 4$$

$$\sigma_b = 10$$

$$\sigma_c = 15$$

Among these:

$$4 < 10 < 15$$

Therefore,

$\sigma_a = 4$ is the best.

6. Reason

The dataset with $\sigma = 4$ is better because:

- It has the least variability.
- Values are closest to the mean.
- It shows high consistency.
- It is more reliable.
- It has less uncertainty.

In most real-world applications such as manufacturing, education results, finance, and healthcare, lower standard deviation is preferred because it indicates stable and predictable outcomes.

7. Practical Interpretation

If this represents:

Marks of students

Production output

Machine performance

Test scores

Then:

- $\sigma = 4$ means results are consistent.
- $\sigma = 10$ means results vary moderately.
- $\sigma = 15$ means results vary greatly.

Organizations usually prefer consistency over fluctuation.

8. Conclusion

When comparing datasets with the same mean, the standard deviation determines which one is better.

Since $\sigma_a = 4$ is the smallest standard deviation, it represents the most consistent and stable dataset.

Therefore, the better option is:

$$\sigma_a = 4$$

It provides higher reliability, lower variation, and better performance compared to $\sigma_b = 10$ and $\sigma_c = 15$.