Table of Contents

# 1.4 Data Types, Central Tendency & Dispersion

## Data types

* **Qualitative (categorical)** – Non‑numerical categories such as gender or product name.
* **Quantitative** – Numeric values that can be discrete (counts) or continuous (measurements).

## Measures of central tendency

* **Mean (average)** – Sum of all values divided by the number of observations【65710891574784†L24-L36】. The mean is sensitive to extreme values【65710891574784†L47-L64】.
* **Median** – The middle value when data are ordered from smallest to largest. When there is an even number of observations, it is the average of the two middle values.
* **Mode** – The most frequently occurring value.

## Measures of dispersion

* **Range** – Difference between the maximum and minimum values【972379405084779†L182-L201】.
* **Variance** – Average of the squared deviations from the mean; quantifies variability【972379405084779†L203-L236】.
* **Standard deviation** – Square root of the variance; expresses spread in the same units as the data【972379405084779†L203-L236】.

## Example

Suppose the ages of five students are 19, 20, 21, 21 and 22.

* The **mean** age is (19 + 20 + 21 + 21 + 22)/5 = 20.6 years.
* The **median** age (middle value) is 21 years.
* The **mode** is 21, since it appears most frequently.
* The **range** is 22 − 19 = 3 years.
* The **variance** is the average of squared deviations from the mean (≈1.04), and the standard deviation is the square root (≈1.02).

## Summary

Measures of central tendency describe a “typical” value in a dataset, while measures of dispersion describe how spread out the data are. Together they characterise the distribution and aid comparisons across datasets.

## Reflection questions

1. When might the median be more appropriate than the mean?
2. Explain how variance and standard deviation differ.
3. Give an example of a discrete and a continuous quantitative variable.

## References