Mean and Median as Centers of Data

The mean (arithmetic average) of a data set is calculated by summing all the values and dividing by the number of values. The median is the middle value when the data is sorted. Both are measures of central tendency, representing the 'center' of the data.

For two numbers, the mean is exactly halfway between them. For example, for numbers 4 and 6, the mean is $\frac{4+6}{2} = 5$, which lies midway on the number line.

For three or more numbers, the mean balances the data such that the total distances of values on either side of the mean are equal. This balance point is unique for any data set.

The median divides the data into two equal halves, with an equal number of values less than and greater than it. When the number of data points is even, the median is the average of the two middle values.

Formula Derivation

Given data values $x_1, x_2, \ldots, x_n$, the mean a is:

$$a = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

The median is the middle value after sorting the data. If n is odd, median is $x_{\frac{n+1}{2}}$. If n is even, median is $\frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$.

Worked Illustration

Consider the data set: 3, 7, 8, 9, 10.

Mean: $\frac{3+7+8+9+10}{5} = \frac{37}{5} = 7.4$

Median: Sorted data is 3, 7, 8, 9, 10. Middle value (3rd) is 8.

Solved Example

Find the mean and median of the data: 2, 6, 10.

Step 1: Calculate mean:

$$\text{Mean} = \frac{2+6+10}{3} = \frac{18}{3} = 6$$

Step 2: Find median:

Sorted data: 2, 6, 10. Middle value is 6.

Thus, mean = 6 and median = 6.

Practice Set

- Level 1 – Easy
  - Find the mean and median of 5, 10, 15.
  - Calculate the mean of 8 and 12.
- Level 2 – Moderate
  - Find the median of 3, 7, 9, 12, 15, 18.
  - Calculate the mean of 4, 6, 8, 10, 12.
- Level 3 – Challenging
  - Given data 10, 10, 11, 17, find the mean and verify the uniqueness of the mean as the balance point.
  - Explain how the mean changes when a new value greater than the mean is added to the data set 4, 6, 8.

Answer Key

- Level 1
  - Mean = $\frac{5+10+15}{3} = 10$, Median = 10
  - Mean = $\frac{8+12}{2} = 10$
- Level 2
  - Sorted data: 3, 7, 9, 12, 15, 18; Median = $\frac{9+12}{2} = 10.5$
  - Mean = $\frac{4+6+8+10+12}{5} = 8$
- Level 3
  - Mean = $\frac{10+10+11+17}{4} = 12$. The mean balances the sum of distances on both sides uniquely.
  - Adding a value greater than the mean increases the mean to maintain balance.

Quick Reference

- Mean: $\frac{\text{Sum of all values}}{\text{Number of values}}$
- Median: Middle value in sorted data; average of two middle values if even number of data points

- Mean is the unique balance point of data

Glossary

- **Mean:** The arithmetic average of a set of numbers.
- **Median:** The middle value in an ordered data set.
- **Central Tendency:** A measure that represents the center of a data set.
- **Balance Point:** The point where the sum of distances on either side is equal.

Effects of Including or Removing Values on Mean and Median

Including or removing values from a data set affects the mean and median depending on the value's relation to the current mean or median.

If a new value greater than the mean is added, the mean increases; if less, the mean decreases. Similarly, adding a value greater than the median increases the median, and adding a value less than the median decreases it.

Removing values equal to the mean or median may leave these measures unchanged, but removing values greater or less than them affects the measures accordingly.

Formula Derivation

Let the original data have n values with mean a. The sum of all values is $S = n \times a$.

When a new value x is added, the new mean a' is:

$$a' = \frac{S + x}{n + 1}$$

If $x > a$, then $a' > a$; if $x < a$, then $a' < a$.

Worked Illustration

Original data: 4, 6, 8 with mean $\frac{4+6+8}{3} = 6$.

Add value 10 (greater than mean):

New mean = $\frac{18+10}{4} = \frac{28}{4} = 7$ (increased).

Add value 2 (less than mean):

New mean = $\frac{18+2}{4} = 5$ (decreased).

Solved Example

Data: 42, 40, 39, 33, 48, 38, 42, 35, 32 with mean 39.2. Find the missing value w if the total number of values is 10.

Sum of known values = 42 + 40 + 39 + 33 + 48 + 38 + 42 + 35 + 32 = 349.

Mean formula:

$$\frac{349 + w}{10} = 39.2$$

Multiply both sides by 10:

$$349 + w = 392$$

$$w = 392 - 349 = 43$$

The missing value is 43.

Practice Set

- Level 1 – Easy
  - Find the new mean when 5 is added to data 2, 4, 6.
  - Find the missing value if the mean of 4 numbers 3, 7, 9, and x is 6.
- Level 2 – Moderate
  - Data has mean 10 with 5 values. If a value 15 is added, find the new mean.
  - Explain how the median changes when a value less than the median is added.
- Level 3 – Challenging
  - Show algebraically that adding a fixed number to all data values increases the mean by that number.
  - Find two values to add to data 3, 5, 7 so that the mean remains unchanged.

Answer Key

- Level 1
  - New mean = $\frac{2+4+6+5}{4} = \frac{17}{4} = 4.25$
  - $\frac{3+7+9+x}{4} = 6 \Rightarrow 19 + x = 24 \Rightarrow x = 5$
- Level 2
  - Sum = 10 × 5 = 50; New sum = 50 + 15 = 65; New mean = $\frac{65}{6} \approx 10.83$
  - Adding a value less than median decreases the median.
- Level 3

- See formula derivation for proof.
- Adding 2 and 4 to data 3, 5, 7: New mean = $\frac{3+5+7+2+4}{5} = \frac{21}{5} = 4.2$ which is different; need to find values that balance sum to keep mean same.

Quick Reference

- Adding a value greater than mean increases mean.
- Adding a value less than mean decreases mean.
- Adding a fixed number to all values increases mean by that number.

Glossary

- **Mean:** Average of data values.
- **Median:** Middle value in sorted data.
- **Sum of distances:** Total absolute differences from a point.

Mean and Median with Frequencies

When data values have frequencies, the mean and median are calculated by considering the frequency of each value.

The mean is the weighted average:

$$\text{Mean} = \frac{\sum(x_i \times f_i)}{\sum f_i}$$

where x_i are the values and f_i their frequencies.

The median is the value at the middle position when data is expanded according to frequencies.

Worked Illustration

Data values and frequencies:

Value (x_i)	Frequency (f_i)
3	3
4	11
5	9
6	7
7	3
8	1
9	1
10	1

Calculate mean:

$$\frac{3\times3+4\times11+5\times9+6\times7+7\times3+8\times1+9\times1+10\times1}{3+11+9+7+3+1+1+1}=\frac{188}{36}\approx5.22$$

Median position: $\frac{36+1}{2}=18.5$th value.

Cumulative frequencies: 3, 14, 23, ... The 18th and 19th values are 5, so median = 5.

Solved Example

Find the mean family size from the above data.

Mean = 5.22 (as calculated above).

Practice Set

- Level 1 – Easy
  - Calculate mean for data: values 1, 2, 3 with frequencies 2, 3, 5.
- Level 2 – Moderate
  - Find median for data: values 4, 5, 6, 7 with frequencies 1, 2, 3, 4.
- Level 3 – Challenging
  - Given frequency data, find missing frequency if mean is known.
  - Explain how median changes when frequencies change.

Answer Key

- Level 1
  - Mean = $\frac{1\times2+2\times3+3\times5}{2+3+5} = \frac{2+6+15}{10} = 2.3$
- Level 2
  - Sorted data expanded: 4(1), 5(2), 6(3), 7(4) total 10 values; median is average of 5th and 6th values, both 6, so median = 6.
- Level 3
  - Use mean formula to solve for missing frequency.
  - Median shifts depending on cumulative frequencies.

Quick Reference

- Mean with frequencies: weighted average.
- Median with frequencies: middle value in ordered expanded data.

Glossary

- **Frequency:** Number of times a value occurs.
- **Weighted Average:** Average accounting for frequencies.

Visualising Data Using Line Graphs

Line graphs display data points connected by line segments, useful for showing trends over time.

They are effective for continuous data such as temperature, rainfall, or counts over months or years.

Worked Illustration

Monthly maximum temperatures of Kerala and Punjab are plotted over a year. Kerala's temperature remains steady, while Punjab's varies significantly.

Solved Example

Given monthly rainfall data for three cities, plot line graphs to compare rainfall patterns.

Practice Set

- Level 1 – Easy
  - Plot monthly sales data for a shop over 6 months.
- Level 2 – Moderate
  - Interpret a given line graph showing temperature changes.
- Level 3 – Challenging
  - Compare two line graphs and infer differences in trends.

Answer Key

- Level 1
  - Plot points and connect with lines to show sales trend.
- Level 2
  - Identify peaks, troughs, and seasonal patterns.
- Level 3
  - Analyze differences in slopes and fluctuations.

Quick Reference

- Line graphs connect data points to show trends.
- Useful for time series and continuous data.

Glossary

- **Line Graph:** Graph connecting data points with lines.
- **Trend:** General direction of data over time.