KIN 610: Quantitative Methods in Kinesiology

Chapter 4: Measures of Central Tendency

Ovande Furtado Jr., PhD.

2026-01-10

1 FYI

This presentation is based on the the following books. The references are coming from these books unless otherwise specified.

Main sources:

- Weir, J. P., Vincent, W. J. (2021). Statistics in Kinesiology. Human Kinetics.
- Furtado, O., Jr. (2026). *Statistics for movement science: A hands-on guide with SPSS* (1st ed.). https://drfurtado.github.io/sms

ClassShare App

You may be asked in class to go to the ClassShare App to answer questions.

- https://classsharedrfurtado.netlify.app/

2 Learning Objectives

- Define central tendency and identify the three primary measures: mode, median, and mean.
- Calculate the mode for raw data and frequency distributions.
- Calculate the median for raw data and frequency distributions.
- Calculate the mean for raw data and frequency distributions.
- Identify the appropriate measure of central tendency for different types of data.
- Explain the advantages and disadvantages of each measure of central tendency.
- Describe the relationship between the mean, median, and mode in different types of distributions.

3 Distribution Shapes

Understanding distribution shapes is crucial for selecting the appropriate measure of central tendency[1].

3.1 Normal (Symmetric)

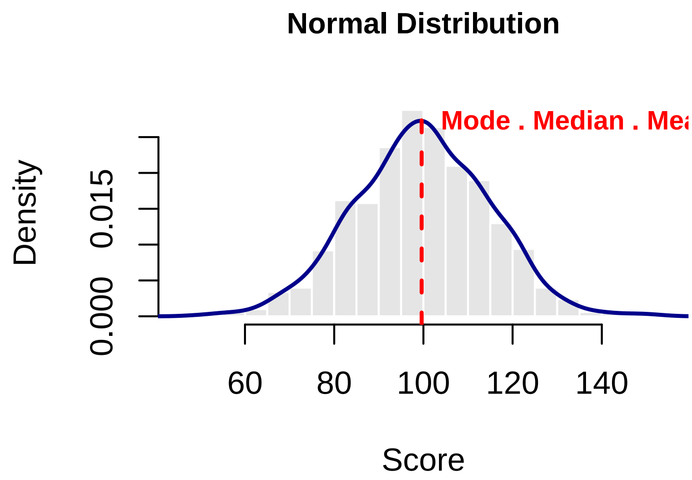


Figure 1: Normal Distribution

Characteristics: - Symmetric bell shape - Mean = Median = Mode - **Best measure**: Mean

3.2 Positively Skewed

Characteristics: - Tail extends right - Mode < Median < Mean - **Best measure**: Median

3.3 Negatively Skewed

Characteristics: - Tail extends left - Mean < Median < Mode - **Best measure**: Median

❗ Important

The **shape of the distribution** determines which measure of central tendency is most appropriate. In skewed distributions, the mean is pulled toward the tail, while the median

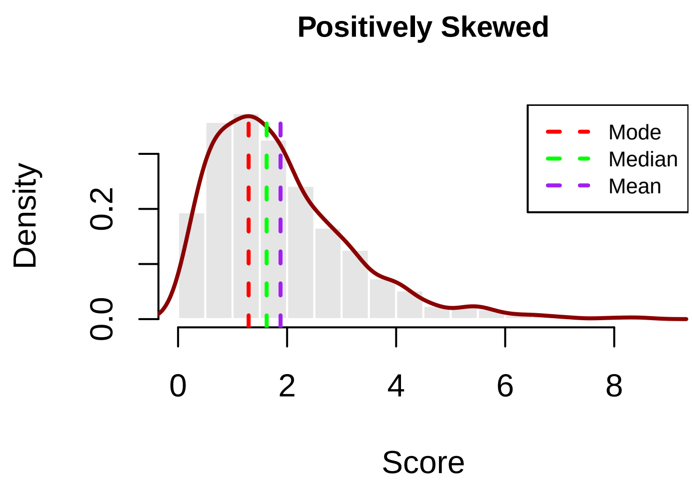


Figure 2: Positively Skewed

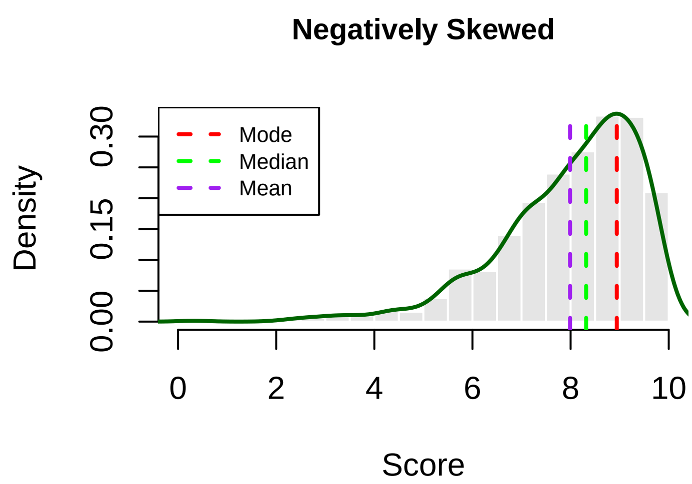


Figure 3: Negatively Skewed

remains more representative of the "typical" value.

- Key teaching point: "Distribution shape tells us which measure to trust."
- Normal: All three measures agree — use the mean for further calculations.
- Skewed: Mean is pulled by outliers — use the median for a better "typical" value.
- Example: "Income data is usually positively skewed (few very high earners), so median income is more meaningful than mean income."

4 Introduction: Describing the Center

This chapter focuses on **measures of central tendency** — statistical values that describe the middle or central characteristics of a data set[1].

- We will explore three primary measures:
  - **Mode**: The most frequent score
  - **Median**: The middle score
  - **Mean**: The arithmetic average

> **!** Important
>
> These measures describe how scores tend to cluster in a distribution, which is fundamentally different from measures of variability (how scores spread out), covered in Chapter 5.

- Opening: Emphasize that central tendency answers the question "What is typical?" in a data set.
- Example: If you measure vertical jump heights of 30 athletes, what single value best represents the group?
- Key distinction: Central tendency (where scores cluster) vs. variability (how spread out they are).
- Teaching tip: Ask students to think of a scenario where knowing the "average" matters (e.g., typical recovery time, average sprint speed).

5 Key Terms

Understanding the terminology is essential for selecting and interpreting the appropriate measure[1]:

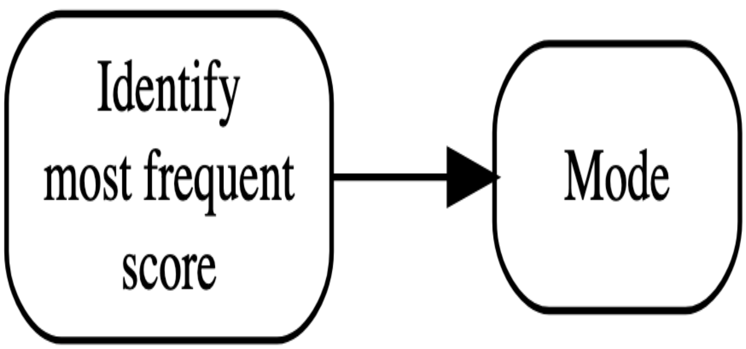
- **Measures of Central Tendency**: Values that describe the middle or central characteristics of a data set

- **Mode**: The most frequent or common score in a distribution
- **Median**: The score above and below which half (50%) of all scores fall; the "typical" score
- **Mean**: The sum of all scores divided by the total number of scores; the arithmetic average

> **Note**
>
> Each measure has specific use cases, strengths, and limitations. The choice depends on the `data type` and `distribution shape`. Recall we discussed data types in Chapter 2 - nominal, ordinal, interval, and ratio. Distribution shapes will be discussed in Chapter 6.

- Emphasize: These are not interchangeable — each serves a different purpose.
- Example: Mode for categorical data, median for skewed data, mean for normal distributions.
- Quick check: "Which measure would you use for jersey numbers?" (Mode — nominal data)

6 The Mode: Most Frequent Score



- **Definition**: The score that occurs most frequently in a distribution
- **Calculation**: No formula required — identified by inspection
- **Methods**:
  - **Rank order listing**: Look for the score that appears most often
  - **Frequency distributions**: Find the highest value in the frequency (f) column

> **Note**
>
> Distributions can be **unimodal** (one mode), **bimodal** (two modes), **trimodal** (three modes), or **multimodal** (more than three modes). A distribution can technically have any number of modes.

- Key point: The mode is the simplest measure — just count frequencies.
- Example: In a class of 25 students, if 8 scored 85 on a test and no other score appeared more than 5 times, the mode is 85.
- Teaching prompt: "What does it mean if a distribution has two modes?" (Bimodal — suggests two distinct groups)

7 Mode Example: Pull-Up Scores

Table 4.1: Rank Order Distribution of Pull-Up Scores

Score	Frequency
18	1
17	2
16	3
15	5
14	4
13	3
12	2
11	1
10	1

> **Tip**
>
> **Question**: What is the mode in this distribution?
> Answer
> The mode is **15** because it has the highest frequency (5 occurrences).

- Walk through the table: "We scan the frequency column and find that 15 appears 5 times — more than any other score."
- Emphasize: This is purely descriptive — it tells us the most common performance level.

8 Disadvantages of the Mode

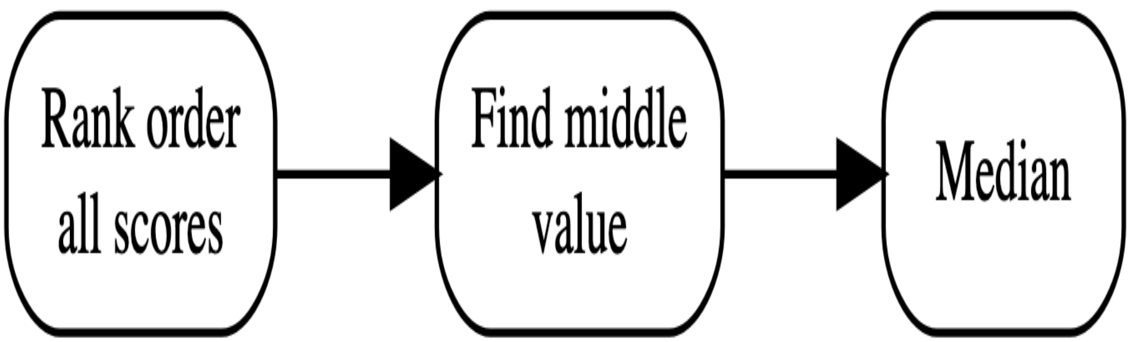
While simple, the mode has significant limitations:

1. **Unstable**: Can change dramatically based on how data is grouped or categorized
2. **Terminal statistic**: Not useful for further mathematical or statistical calculations
3. **Ignores information**: Does not account for the values of extreme scores or the overall distribution shape

> ⚠ Warning
>
> The mode is best used as a rough estimate when `data are near normal`, or for `nominal data` where it is the only appropriate measure.

- Example of instability: Changing bin width in a histogram can shift the mode.
- Terminal statistic: You can't use the mode to calculate standard deviation or other advanced statistics.
- Teaching prompt: "When would the mode be misleading?" (When the distribution is flat or has multiple peaks)

9 The Median: The Middle Score



> ℹ Note
>
> The median is the "typical" score and is particularly useful for `ordinal data` or `highly skewed distributions`[1].

- **Definition**: The 50th percentile — the score that divides the data set into two equal halves
- **Characteristics**:

  - Half of all scores fall above the median
  - Half of all scores fall below the median

- **Calculation (Odd N)**: The middle score in the ordered list
- **Calculation (Even N)**: The average of the two middle scores

- Key point: The median is positional — it depends on rank, not actual values.
- Example: For scores [5, 7, 9, 11, 13], the median is 9 (middle position).
- For [5, 7, 9, 11], the median is $(7+9)/2 = 8$.
- Teaching prompt: "Why is the median better than the mean for income data?" (Income is often skewed by very high earners)

10 Median: Strengths and Applications

Key Strengths:

- **Unaffected by extreme scores (outliers)**: A few very high or very low values do not distort the median[2]
- **Appropriate for ordinal data**: Works well when data has rank order but unequal intervals
- **Robust measure**: Provides a stable center even when distributions are skewed[1]

When to Use the Median:

- Data is ordinal (e.g., Likert scales, race finish positions)
- You need the middle or "typical" score
- The distribution is badly skewed by outliers

> Tip
>
> Example: If measuring recovery time and one participant takes 10x longer than others due to injury, the median recovery time is more representative than the mean.

- Emphasize robustness: "If one athlete's VO2max is measured incorrectly as 200 (instead of 50), the median won't change much, but the mean will be severely distorted."
- Teaching prompt: "Give an example from sports where the median is more useful than the mean."

11 The Mean: Arithmetic Average

- **Definition**: The arithmetic average — the sum of all scores divided by the number of scores[1]
- **Status**: The most widely used index of central tendency in statistics
- **Logic**: Balances all values in the distribution (the "balance point")

> **!** Important
>
> The mean is **sensitive to every score** in the distribution, including extreme values (outliers)[3].

- Key point: The mean uses all information in the data set — every value contributes.
- Example: "If you change even one score, the mean will change."
- Contrast with median: "The median only cares about position, the mean cares about actual values."

12 Formula for the Mean

The mean is calculated using the following formula:

$$\bar{X} = \frac{\sum X}{N}$$

Components:

- $\bar{X}$ = Mean (pronounced "X-bar")
- $\sum$ = Summation sign (add up all values)
- X = Individual raw score
- N = Total number of scores

Note

This formula is fundamental to many subsequent statistical calculations, including standard deviation, z-scores, and inferential tests.

- Walk through the formula: "Sum all the scores, then divide by how many scores you have."
- Emphasize: This is the population mean formula; sample mean uses the same logic.
- Teaching prompt: "Why do we use $\bar{X}$ for sample mean and μ for population mean?"

13 Sample Calculation of the Mean

VO max data (mL/kg/min) for 7 participants:

30.6, 29.5, 28.2, 27.8, 26.5, 26.1, 25.7

Calculation:

$$\bar{X} = \frac{\sum X}{N} = \frac{30.6 + 29.5 + 28.2 + 27.8 + 26.5 + 26.1 + 25.7}{7} = \frac{194.4}{7} = 27.77 \text{ mL/kg/min}$$

Important

Note that the mean is expressed in the same units as the raw data (mL/kg/min in this case).

- Walk through step-by-step: "First, add all seven values. Then divide by 7."
- Emphasize units: "Always include units in your final answer."
- Teaching activity: Have students calculate the mean for a small data set.

14 Mean: Strengths and Limitations

Strengths:

- Uses all information in the data set (every value matters)[1]
- Essential for subsequent statistical inference and calculations
- Most appropriate for `interval` and `ratio` data
- Algebraically defined, allowing for advanced mathematical operations

Limitations:

- **Sensitive to outliers**: Extreme scores have a disproportionate impact[3]
- **Arguably inappropriate for ordinal data**: Assumes equal intervals between values
- Can be misleading when distributions are highly skewed

> ⚠ Warning
>
> When extreme scores are present, the mean may not represent the "typical" value well. Consider using the median instead[2].

- Example of outlier effect: "If 9 athletes earn $50,000 and 1 earns $500,000, the mean is $95,000 — not representative of most athletes."
- Teaching prompt: "When would you choose the mean over the median?"

15 Test Your Knowledge: Calculating the Mean

You measure the maximum bench press (in kg) for 5 athletes:

Data: 80, 85, 90, 95, 100

Calculate the mean bench press using the calculator on your device.

Answer

$$\bar{X} = \frac{80 + 85 + 90 + 95 + 100}{5} = \frac{450}{5} = 90 \text{ kg}$$

The mean bench press is **90 kg**.

Checking your calculation

Open SPSS using the Canvas link and enter the data into the Data View. Then, go to Analyze > Descriptive Statistics > Frequencies. Select the variable you entered and click OK. The mean will be displayed in the output.

16 Normal Distribution: Mode = Median = Mean

Characteristics:

- All three measures **coincide** at or near the same value
- Mode = Median = Mean
- This is the ideal scenario for using the mean
- Symmetric bell-shaped distribution

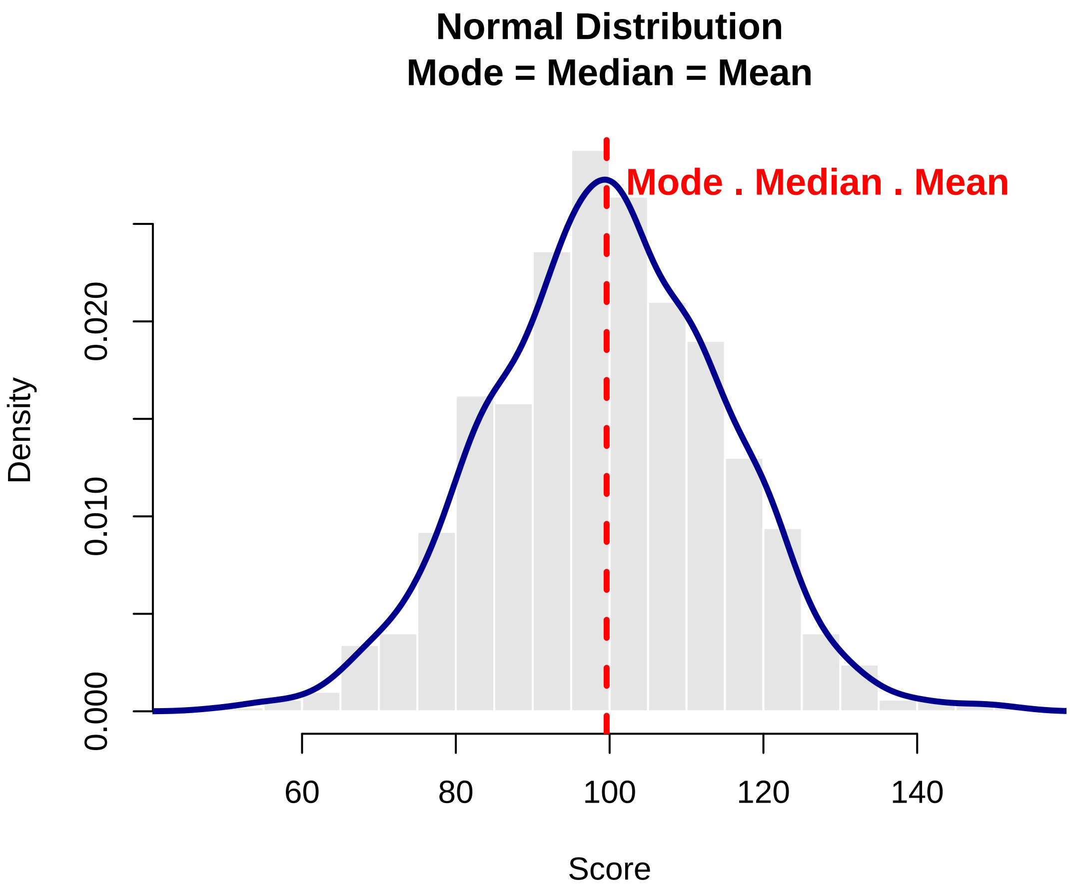


Figure 4: Normal distribution: all measures coincide

Best measure: Mean (can use for further calculations)

- Key teaching point: “In a normal distribution, all three measures tell the same story.”
- This is why the mean is preferred for normal distributions - it uses all the data and coincides with the median.

17 Positively Skewed: Mode $<$ Median $<$ Mean

Characteristics:

- The measures are **pulled apart**
- The mean is pulled furthest toward the tail (right)
- The median sits between the mode and mean
- Tail extends to the right

Best measure: Median (resistant to outliers)

- Example: “In a positively skewed income distribution, the mean is pulled up by high earners, but the median better represents the typical person.”
- Common in real data: income, reaction time, housing prices

18 Negatively Skewed: Mean $<$ Median $<$ Mode

Characteristics:

- The measures are **pulled apart**
- The mean is pulled furthest toward the tail (left)
- The median sits between the mean and mode
- Tail extends to the left

Best measure: Median (resistant to outliers)

- Example: “Test scores with a ceiling effect - most students score high, but a few struggle.”
- Less common than positive skew, but important to recognize

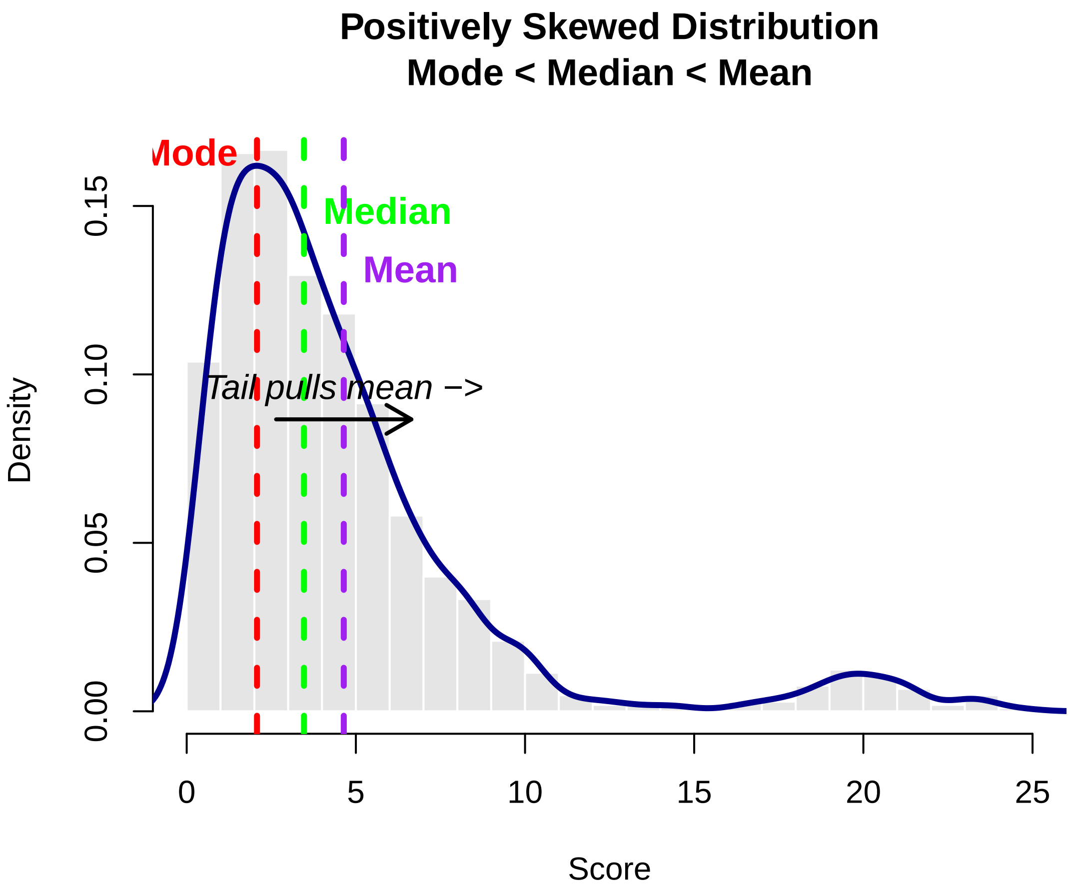


Figure 5: Positively skewed distribution showing Mode, Median, and Mean

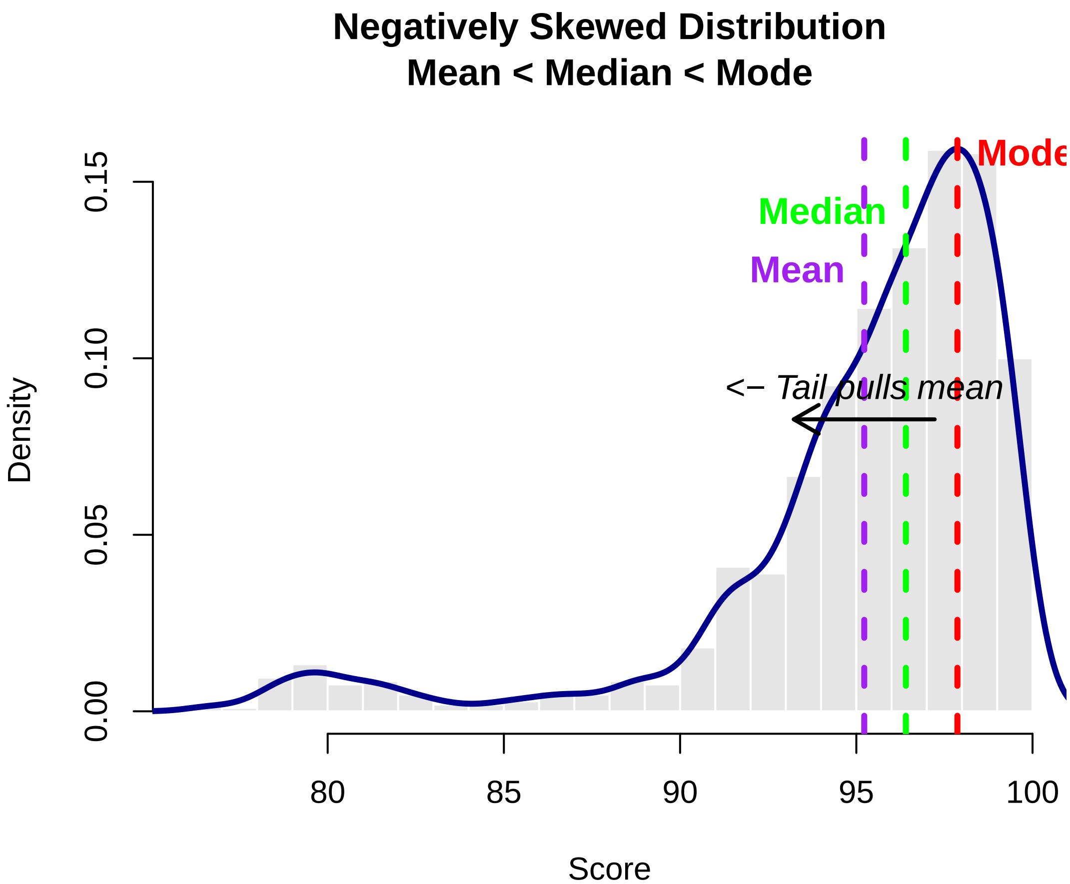
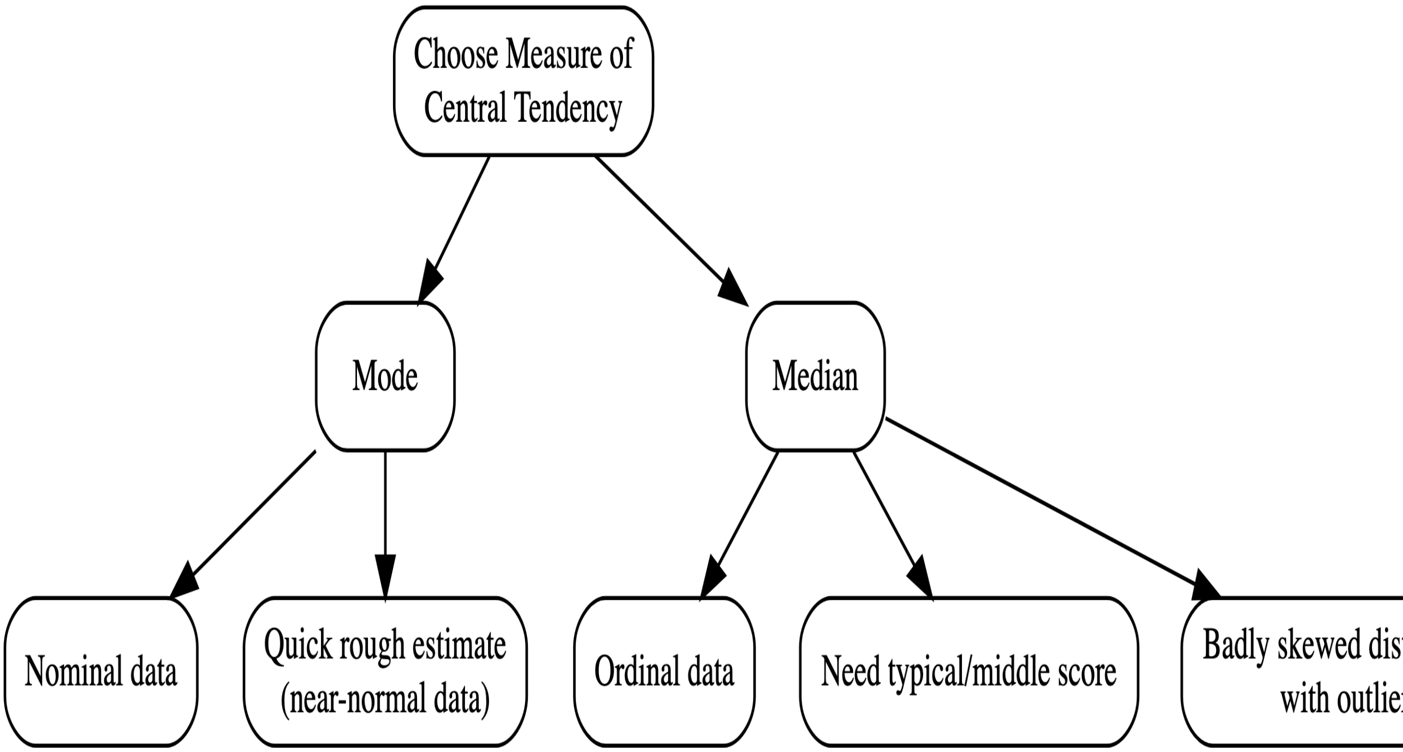


Figure 6: Negatively skewed distribution showing Mode, Median, and Mean

19 Summary: When to Use Mode and Median

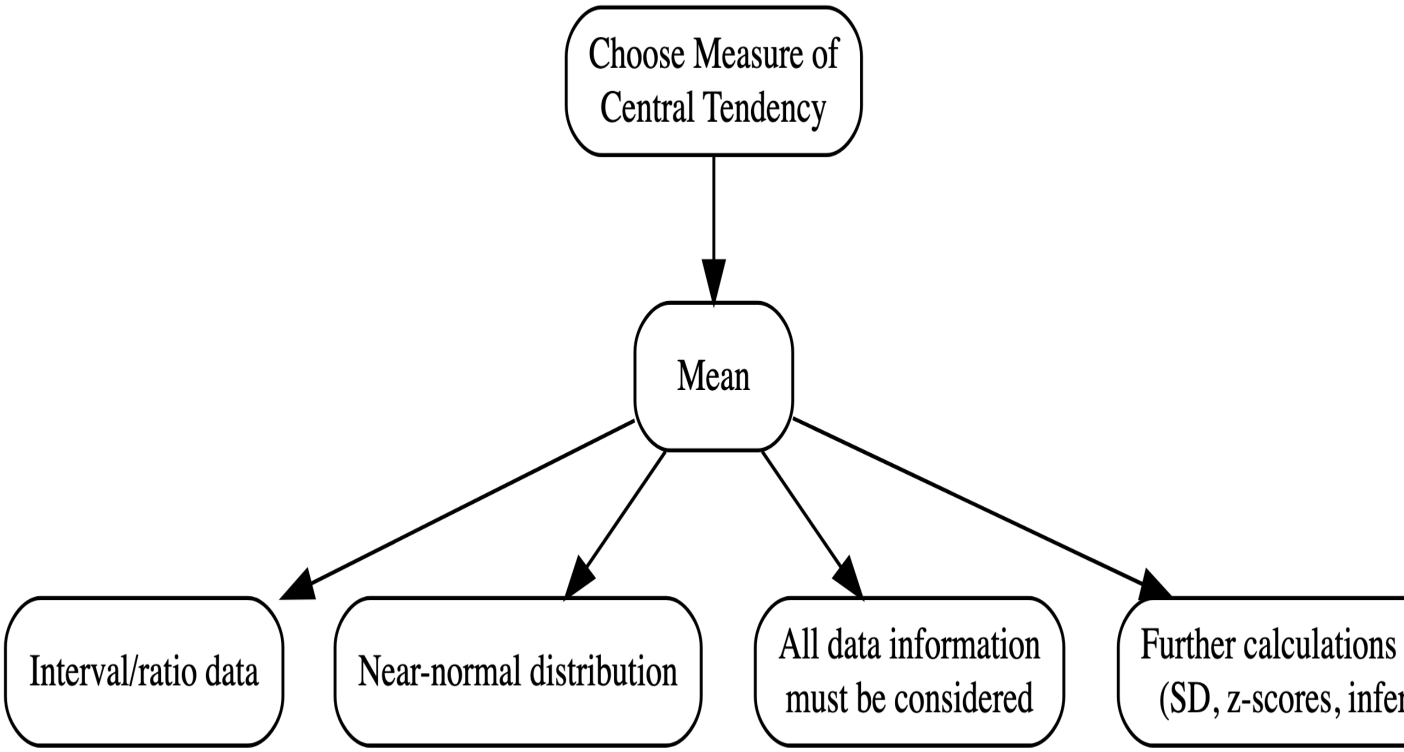


Tip

Mode: Use for a rough estimate if data are near normal, or when working with nominal (categorical) data.
Median: Use if data is ordinal, you need the middle/typical score, or the distribution is badly skewed by outliers.

- Emphasize decision-making: "The choice of measure depends on your data type and distribution shape."
- Teaching prompt: "Give an example of when you would use the mode vs. the median."

20 Summary: When to Use the Mean

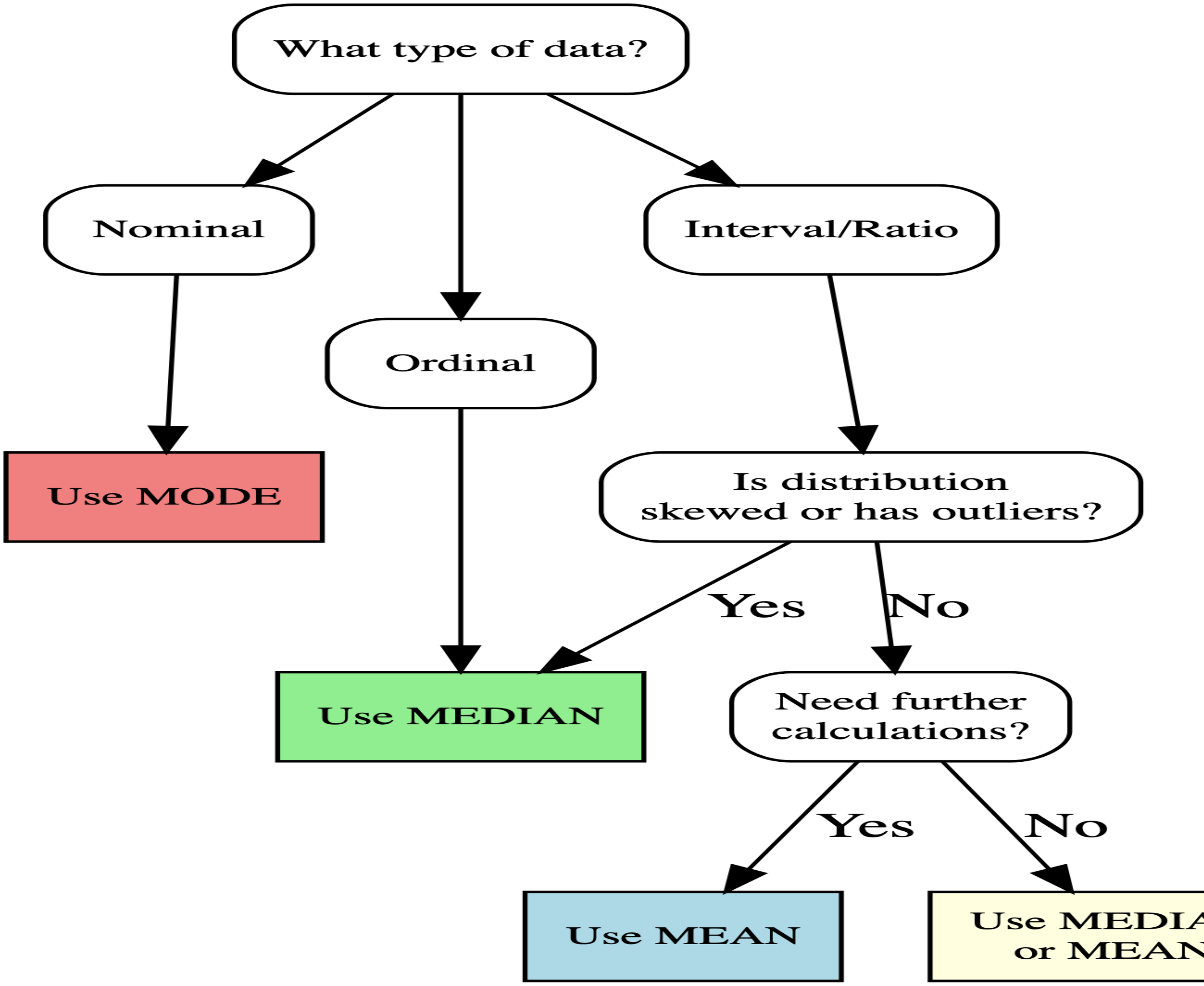


Tip

Mean: Use if data is interval/ratio and the distribution is near normal, all data information (values and order) must be considered, or further calculations (e.g., standard deviation, standard scores, inferential tests) are required.

- Key point: "The mean is the workhorse of statistics — most advanced analyses require it."
- But: "Don't use it blindly — check for outliers and skewness first."

21 Decision Tree: Selecting the Right Measure



> **! Important**
>
> This decision tree provides a practical guide, but always consider the research question and context when selecting a measure of central tendency.

22 Test Your Knowledge: Choosing the Right Measure

For each scenario, identify which measure of central tendency (mode, median, or mean) is most appropriate:

1. **Sport type** (soccer, basketball, swimming) for 100 athletes
2. **Finish position** in a marathon (1st, 2nd, 3rd, …)
3. **Reaction time** (in milliseconds) with one extreme outlier
4. **Body mass** (in kg) for a normally distributed sample, and you need to calculate standard deviation

Using the ClassShare App, submit your answers.

Answers

1. **Mode** — nominal data (categories only)
2. **Median** — ordinal data (rank order)
3. **Median** — interval/ratio data with outliers
4. **Mean** — interval/ratio data, normal distribution, further calculations needed

23 Practical Application: Real Data Example

Scenario: You measure sprint times (in seconds) for 10 athletes:

Data: 10.2, 10.5, 10.3, 10.8, 10.2, 10.6, 10.4, 10.2, 10.7, 12.5

Questions:

1. What is the mode?
2. What is the median?
3. What is the mean?
4. Which measure best represents "typical" performance? Why?

Click to reveal answers

1. **Mode**: 10.2 seconds (appears 3 times)
2. **Median**:

    - Ordered data: 10.2, 10.2, 10.2, 10.3, 10.4, 10.5, 10.6, 10.7, 10.8, 12.5
    - Middle two values: 10.4 and 10.5
    - Median = (10.4 + 10.5) / 2 = **10.45 seconds**

3. **Mean**:

    - Sum = 106.4
    - Mean = 106.4 / 10 = **10.64 seconds**

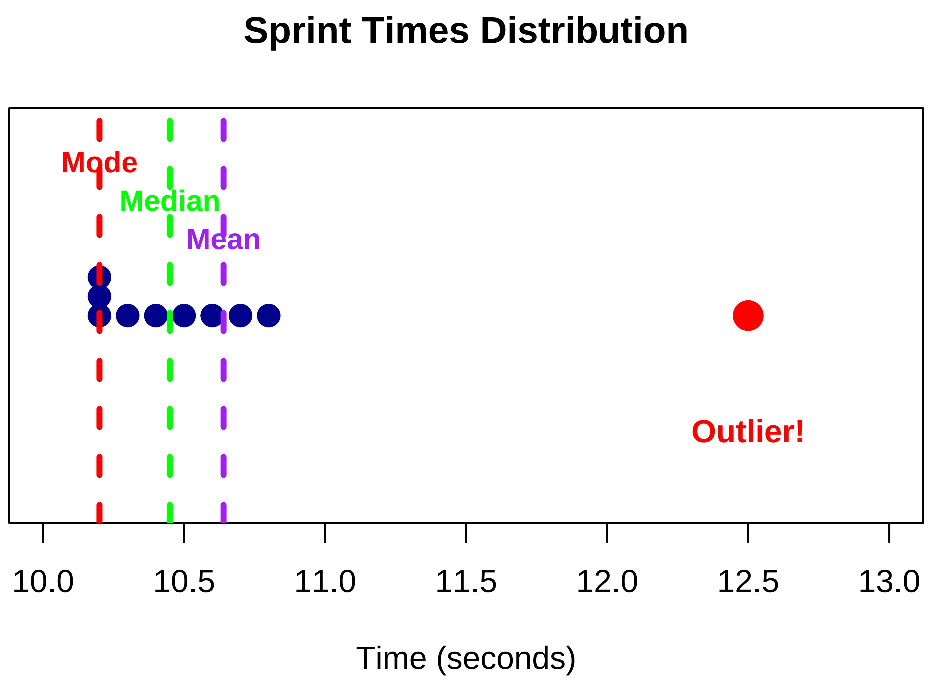


Figure 7: Distribution of sprint times showing outlier effect

4. **Best measure**: The **median (10.45 s)** best represents typical performance because the mean is pulled upward by the outlier (12.5 s). The median is more robust to this extreme value.

> **Note**
>
> This example demonstrates why understanding distribution shape and outliers is crucial for selecting the appropriate measure of central tendency. Notice how the outlier (12.5 s) pulls the mean to the right, while the median stays closer to the bulk of the data.

> **Tip**
>
> The graph above is a dot plot. To learn how to create dot plots, click here.

24 Key Takeaways

> **Remember These Core Concepts:**
>
> 1. **Mode**: Most frequent score — simple but limited; best for nominal data[1]
> 2. **Median**: Middle score (50th percentile) — robust to outliers; best for ordinal or skewed data[1,2]
> 3. **Mean**: Arithmetic average — uses all information; best for interval/ratio data with normal distributions[1]
> 4. **Distribution shape matters**: In normal distributions, all three measures converge; in skewed distributions, they diverge[3]
> 5. **Choose wisely**: Match the measure to your data type, distribution shape, and research goals[4]

- Closing: "These three measures are the foundation for understanding any data set. Master them, and you'll be able to describe and interpret data with confidence."
- Final prompt: "What questions do you have about when to use each measure?"

25 Next Steps: Chapter 5

In the next chapter, we will explore **measures of variability**:

- How do scores spread out around the center?
- What is the range, variance, and standard deviation?
- How do we interpret variability in the context of research?

Note

Understanding both central tendency (where scores cluster) and variability (how spread out they are) gives you a complete picture of your data distribution.

- Preview: "Central tendency tells us the 'typical' value, but variability tells us how much scores differ from that typical value."
- Example: "Two groups might have the same mean, but very different spreads — variability helps us see that difference."

References

1. Moore, D. S., McCabe, G. P., & Craig, B. A. (2021). *Introduction to the practice of statistics* (10th ed.). W. H. Freeman; Company.
2. Wilcox, R. R. (2017). *Introduction to robust estimation and hypothesis testing* (4th ed.). Academic Press.
3. Hippel, P. T. von. (2005). Mean, median, and skew: Correcting a textbook rule. *Journal of Statistics Education*, *13*(2). https://doi.org/10.1080/10691898.2005.11910570
4. Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.
5. Furtado, O., Jr. (2026). *Statistics for movement science: A hands-on guide with SPSS* (1st ed.). https://drfurtado.github.io/sms/