**Chapter 11 - Biostatistics and Data Handling**

**All Lectures Uploaded on YouTube:**

[**https://tinyurl.com/fkm10-biology**](https://tinyurl.com/fkm10-biology)

## 

# **11.1. Biostatistics and Its Uses**

## **11.1.1. Biostatistics**

**Biostatistics** is the application of statistical methods to analyze, interpret, and present data related to **biology, medicine, health sciences, agriculture, and other life sciences**.

It involves:

* **Collection** of biological data
* **Analysis** of data
* **Interpretation** of results
* **Presentation** of findings to derive solutions, policies, or scientific conclusions



Biostatistics helps understand:

* Disease patterns
* Health trends
* Effectiveness of treatments
* Productivity in agriculture, poultry, and dairy sectors

### **Components of Biostatistics Studies**

1. **Identification of the problem**
   * The first step in studying any biological issue, especially those related to humans or living organisms.
2. **Designing experiments**
   * Planning procedures to investigate the biological problem.
3. **Collecting and analyzing data**
   * Systematic data collection followed by thorough statistical analysis.
4. **Interpreting results**
   * Determining causes, outcomes, and future implications based on the analyzed data.
5. **Developing new tools**
   * Using interpretations to create new methods, tools, or plans (e.g., healthcare solutions).

Biologists are encouraged to have basic knowledge of biostatistics.

## **11.1.2. Uses of Biostatistics**

### **1. Agriculture and Cattle Farming**

* Helps estimate food, dairy, and livestock demands of a growing population.
* Guides governments about import/export needs of food items.

### **2. Medical and Pharmacological Research**

* Designs clinical trials and controlled studies.
* Assesses safety, effectiveness, and optimal dosage of drugs.
* Compares treatment benefits vs. side effects.



### **3. Epidemiology and Policy Development**

* Monitors spread and rate of epidemic diseases.
* Identifies risk factors and disease patterns.
* Helps prevent future outbreaks (Polio, COVID-19, Hepatitis).
* Provides statistical evidence for public health policies.

### **4. Public Health Management**

* Studies population data: birth rate, death rate, disease prevalence.
* Assists in planning health strategies and resources.
* Helps improve hospital performance (patients, doctors, medicines).



### **5. Genetic Diseases**

* Analyzes inheritance patterns of disorders (thalassemia, muscular dystrophy).
* Estimates risk factors and predicts genetic disease behavior.

### **6. Environmental Protection**

* Measures pollution levels and related health risks.
* Helps plan solutions (e.g., smog control, identifying affected areas).
* Supports environmental policies (plantation drives in Lahore).

### **7. Survival Analysis**

* Predicts survival rates after treatments.
* Estimates life expectancy and treatment success (e.g., 5-year survival in cancer).

# **11.2. Definition and Calculation of Mean, Median and Mode**

Average-type statements (e.g., heart rate = 72 BPM, rainfall = 1500 mm/year) represent **central tendency**, not exact values.

**Average:** A single value representing an entire dataset.

**Types of averages:**

1. **Arithmetic Mean (Mean)**
2. **Median**
3. **Mode**
4. Geometric Mean
5. Harmonic Mean

## **11.2.1. Arithmetic Mean (Mean)**

### **Definition**

Symbol: **x̄ (x-bar)**

### **Example 1 (Un-grouped Data)**

Monthly polio cases in Pakistan for 2022 & 2023.

**2022 Total:** 192  
**2023 Total:** 144

Mean number of Polio patients per month in 2022:

Mean number of Polio patients per month in 2023:

**Difference:** 4 fewer cases/month in 2023 → **Prevalence decreased**.

### **Example 2 (Grouped Data)**

Mean mass of 100 students.

Using:

Final answer:

Mean mass = 49.5 Lbs

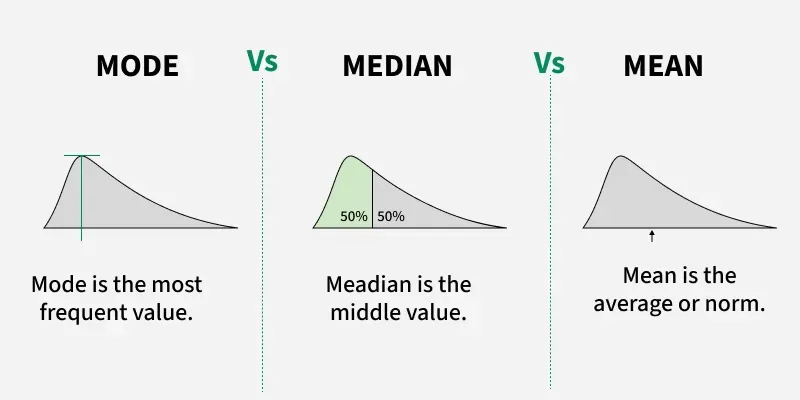
## **11.2.2. Median**

### **Definition**

Median is the **middle value** of a dataset arranged in ascending or descending order.

Symbol: **x̃ (x-tilde)**

* **Odd number of values:**
* **Even number of values:** Median = Mean of the two middle values.



### **Example (Odd number of values)**

Plant heights (13 values): 46,49, 55, 58, 62, 63, 64, 65, 66, 67, 68, 70,72

Median = (7th value)

Median = 64

### **Example (Even number of values)**

Fruit count in 16 plants: 3,5,18,21,15,10,8,12,13,7, 11, 14,9, 16, 20, 24

Median = (Average of 8th and 9th values)

Median =

Median = **12.5 fruits per plant**

### **Advantages of Median**

1. Easy to calculate.
2. Not affected by very large or very small values.
3. Useful for quantitative data.

### **Disadvantages**

1. Cannot combine medians of multiple datasets.
2. May not represent the actual central value.
3. Not useful where exact values matter (age, weight).

## **11.2.3. Mode**

### **Definition**

Mode is the value that occurs **most frequently** in a dataset.

Symbol: **X̂ (x-hat)**

A dataset may be:

* **Unimodal** – one mode
* **Bimodal** – two modes
* **Multimodal** – many modes
* **Amodal** – no mode

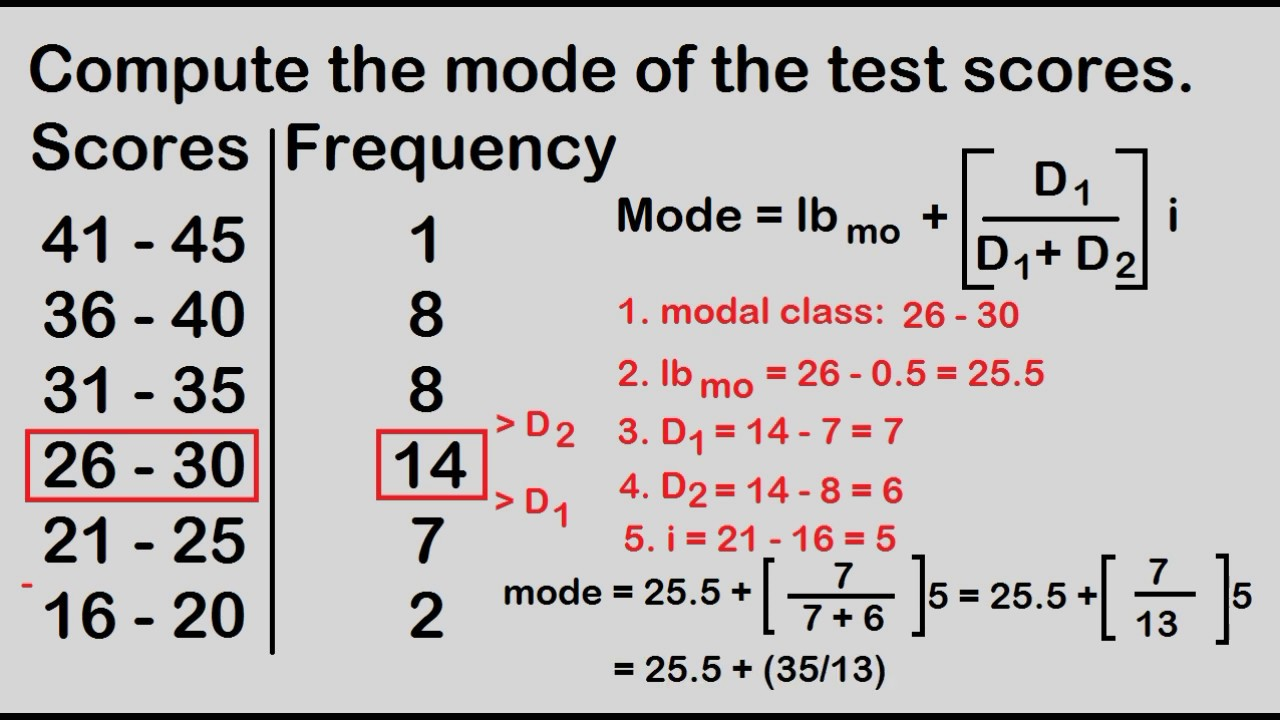
### **Examples**

1. Word **BIOLOGY** → Mode = **O**
2. Values: 3, 5, 6, 6, 9, 10, 9, 6, 6 → Mode = **6**
3. Values: 3, 5, 5, 6, 9, 9, 10 → Mode = **5 & 9** (bimodal)

### **Mode for Grouped Data**

Where:

* **l** = lower boundary of modal class
* **fm** = maximum frequency
* **f1** = frequency before modal class
* **f2** = frequency after modal class
* **h** = class interval



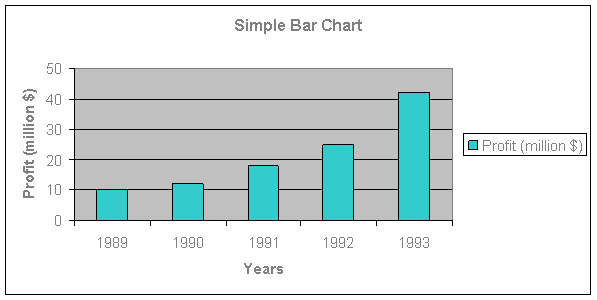
# **11.3. Sketching a Bar Chart for a Given Set of Biological Data**

## **11.3.1. Graphic Presentation: Need and Importance**

* Numerical tables are difficult to understand quickly.
* Diagrams and graphs provide:  
  + Clear representation
  + Better understanding
  + Stronger visual impact
* Types include: bar charts, rectangles, pie charts.

## **11.3.2. Simple Bar Chart**

* Represents **one characteristic** of data only.



* Bars have:  
  + **Equal width**
  + **Variable length** (representing magnitude)
* One bar = one value
* Bars may be vertical or horizontal (usually vertical).

## **11.3.3. Steps for Constructing a Bar Graph**

1. Choose a suitable graph paper.
2. Draw two perpendicular axes meeting at zero.
3. Label horizontal axis (X-axis) and vertical axis (Y-axis).
4. Choose equal widths for bars and equal spacing.
5. Label X-axis with categories.
6. Select a scale for Y-axis based on data range.
7. Mark Y-axis units.
8. Plot heights of bars.
9. Draw bars neatly and check all values.



