# Biostatistics

Lecture 2

Prepared by Dr. Ibrahim AL-Jaafari

# Organizing and Displaying Data
# Qualitative Data

- **Raw Data:**

Data recorded in the sequence in which they are collected and before they a processed or ranked are called ***raw data***.

# Example of Raw Data

**Scores of 50 students**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 21 | 19 | 24 | 25 | 29 | 34 | 26 | 27 | 37 | 33 |
| 18 | 20 | 19 | 22 | 19 | 19 | 25 | 22 | 25 | 23 |
| 25 | 19 | 31 | 19 | 23 | 18 | 23 | 19 | 23 | 26 |
| 22 | 28 | 21 | 20 | 22 | 22 | 21 | 20 | 19 | 21 |
| 25 | 23 | 18 | 37 | 27 | 23 | 21 | 25 | 21 | 24 |

# ORGANIZING AND GRAPHING QUALITATIVE DATA

- Frequency Distributions
- Relative Frequency and Percentage Distributions
- **Graphical Presentation of Qualitative Data**
  - **Bar Graphs**
  - **Pie Charts**

# Frequency Distributions

## Definition

A ***frequency distribution*** for qualitative data lists all categories and the number of elements that belong to each of the categories.

# Frequency Table

Variable → Type of College

Number of Students ← Frequency Column

Category → Public Health — 200

Nursing — 300 ← Frequency

Pharmacy — 400

Sum = 900

# Example for Frequency Table

- A sample of 30 employees from Al-Baha city was selected, and these employees were asked for their monthly salary. The responses of these employees are recorded next where ***high income***, ***middle income***, and ***low income***.

# Construct a frequency distribution table for these data ?

| | | | | | |
|---|---|---|---|---|---|
| High | Low | High | Low | Low | Middle |
| Low | Middle | High | Middle | Middle | High |
| Middle | High | Middle | Middle | Middle | Middle |
| High | Low | High | Middle | Low | Middle |
| High | Low | Low | High | Middle | High |

# Solution for Frequency Table

| Income | Tally | Frequency ($f$) |
|---|---|---|
| High | 𝍸 𝍸 | 10 |
| Middle | 𝍸 𝍸 \|\| | 12 |
| Low | 𝍸 \|\|\| | 8 |
| | | Sum = 30 |

# Relative Frequency and Percentage Distributions

## Calculating Relative Frequency of a Category

$$\text{Relative frequency of a category} = \frac{\text{Frequency of that category}}{\text{Sum of all frequencies}}$$

# Relative Frequency and Percentage Distributions

## Calculating Percentage

**Percentage = (Relative frequency) * 100**

# Determine the Relative Frequency and Percentage for the Data in the Table

| Income | Tally | Frequency ($f$) |
|---|---|---|
| **High** | 𝍸 𝍸 | 10 |
| **Middle** | 𝍸 𝍸 \|\| | 12 |
| **Low** | 𝍸 \|\|\| | 8 |
| | | **Sum = 30** |

# Solution

| Income | Relative Frequency | Percentage |
|---|---|---|
| **High** | 10/30 = 0.333 | 0.333(100) = 33.3 |
| **Middle** | 12/30 = 0.40 | 0.40(100) = 40.0 |
| **Low** | 8/30 = 0.267 | 0.267(100) = 26.7 |
| | **Sum = 1.00** | **Sum = 100** |

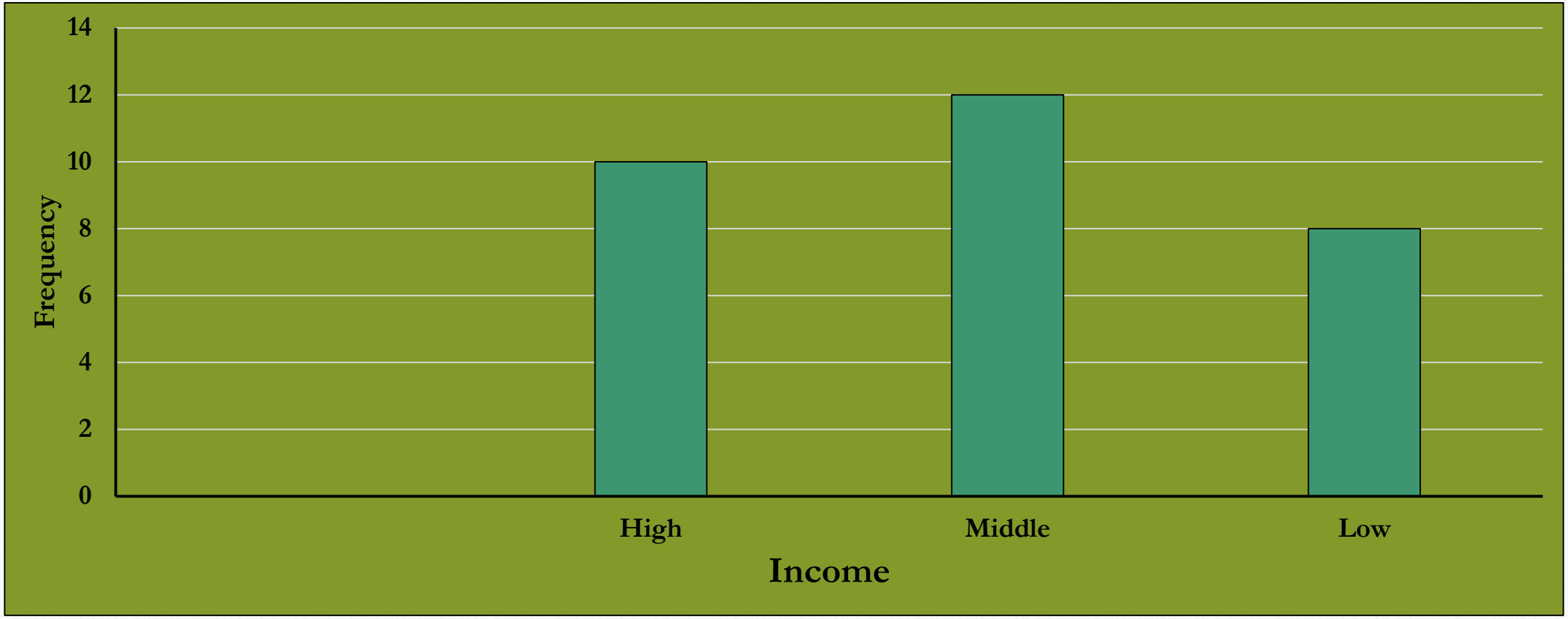
# Graphical Presentation of Qualitative Data

## Bar Graph – Bar Chart

### Definition

A graph made of bars whose heights represent the frequencies of respective categories is called a ***bar graph***.

# Bar Graph – Bar Chart



# Graphical Presentation of Qualitative Data

## Pie Graph - Pie Chart

### Definition

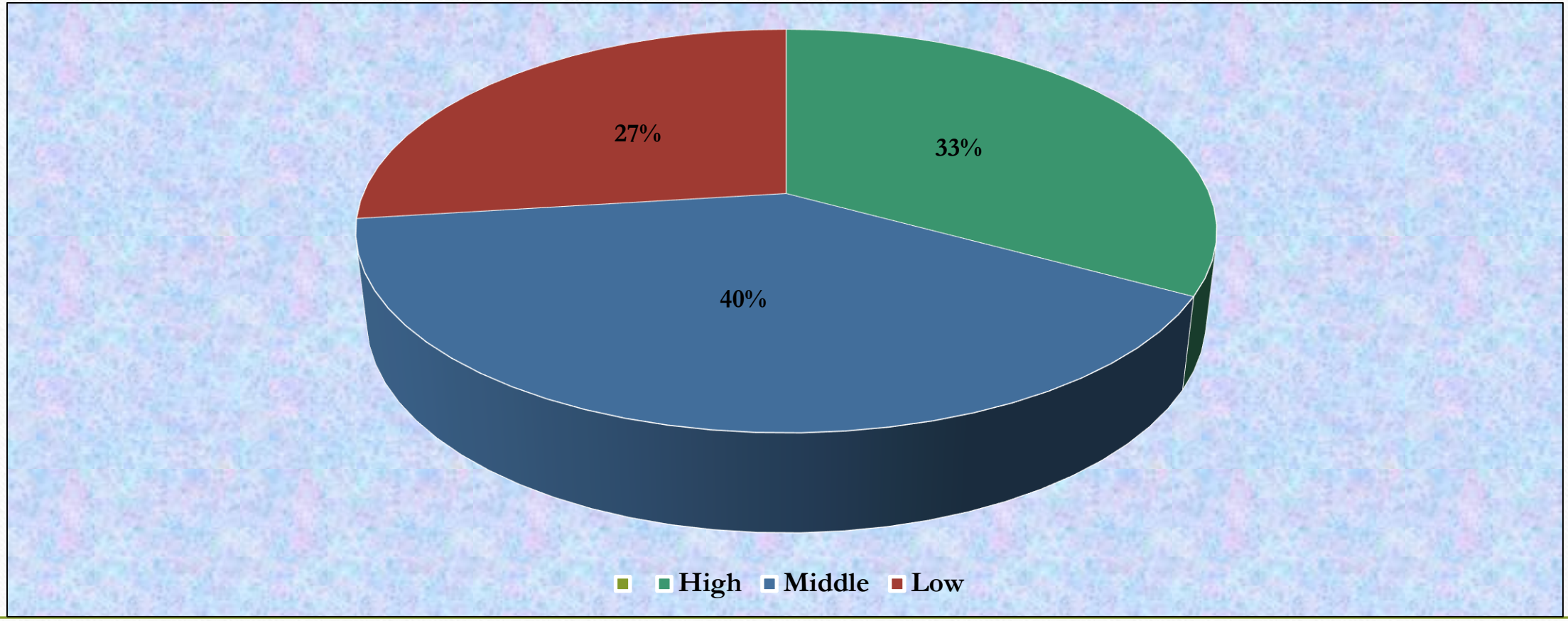
A circle divided into portions that represent the relative frequencies or percentages of a population or a sample belonging to different categories is called a ***pie chart.***

# Calculating Angle Sizes for the Pie Chart

## Pie Graph - Pie Chart

| Income | Relative Frequency | Angle Size |
|---|---|---|
| **High** | 10/30 = 0.333 | 360 (0.333) = 119.9 |
| **Middle** | 12/30 = 0.40 | 360 (0.40) = 144.0 |
| **Low** | 8/30 = 0.267 | 360 (0.267) = 96.1 |
| | **Sum = 1.00** | **Sum = 360** |

# Pie Graph - Pie Chart



# Organizing and Displaying Data
# Quantitative Data

- Graphing Grouped Data
  - Histograms
  - Polygons

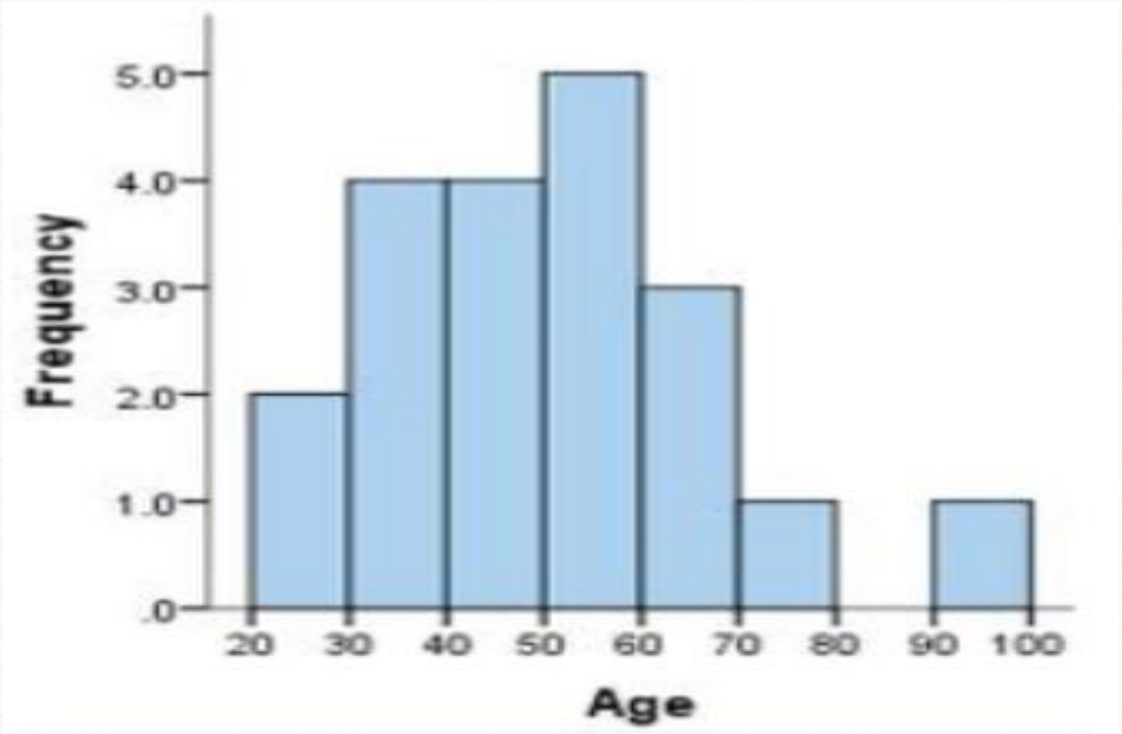
# GRAPHING QUANTITATIVE DATA

**How to present quantitative data ??**

**1- Histograms /** a graph in which classes are marked on horizontal axis and the frequencies are marked on the vertical axis which represent the height of bars.

In a histogram, the bars are attached to each other.

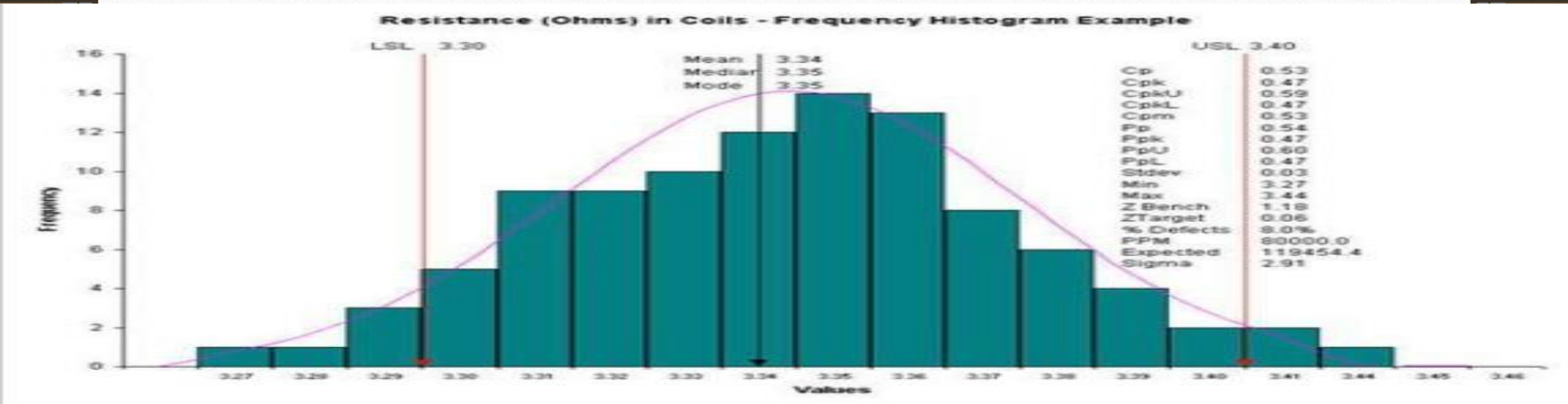
# Histograms



# Histograms

**What are the shapes of histograms ??**

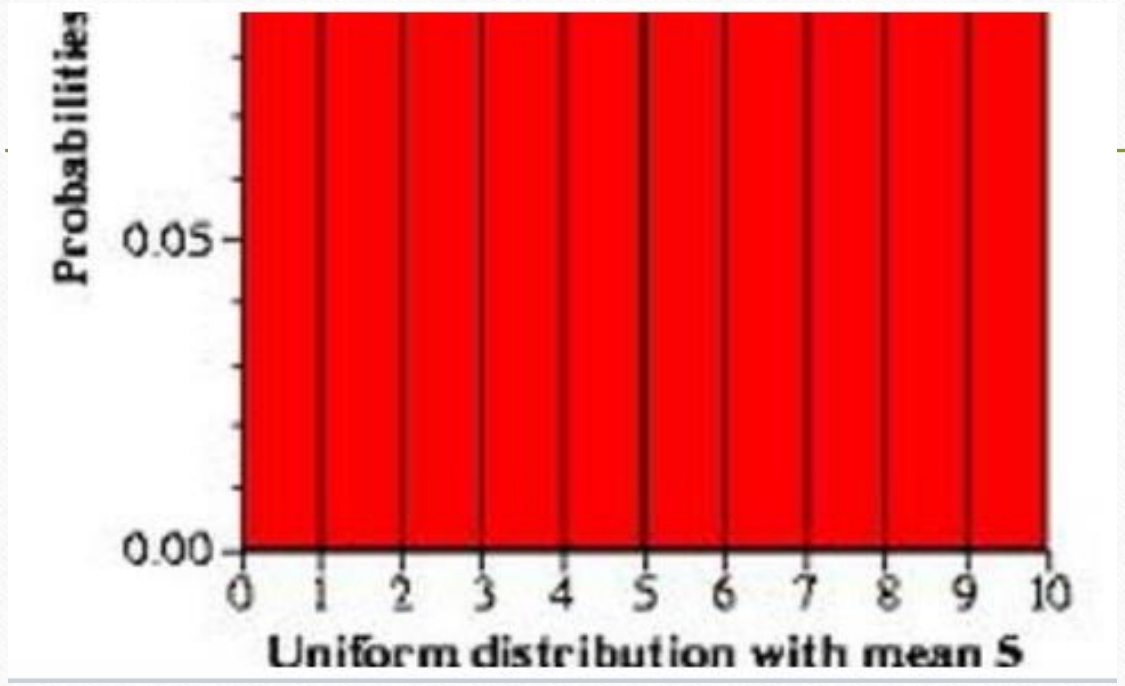
**A-) Symmetric histogram /** is identical on both sides of its central points.



# Histograms

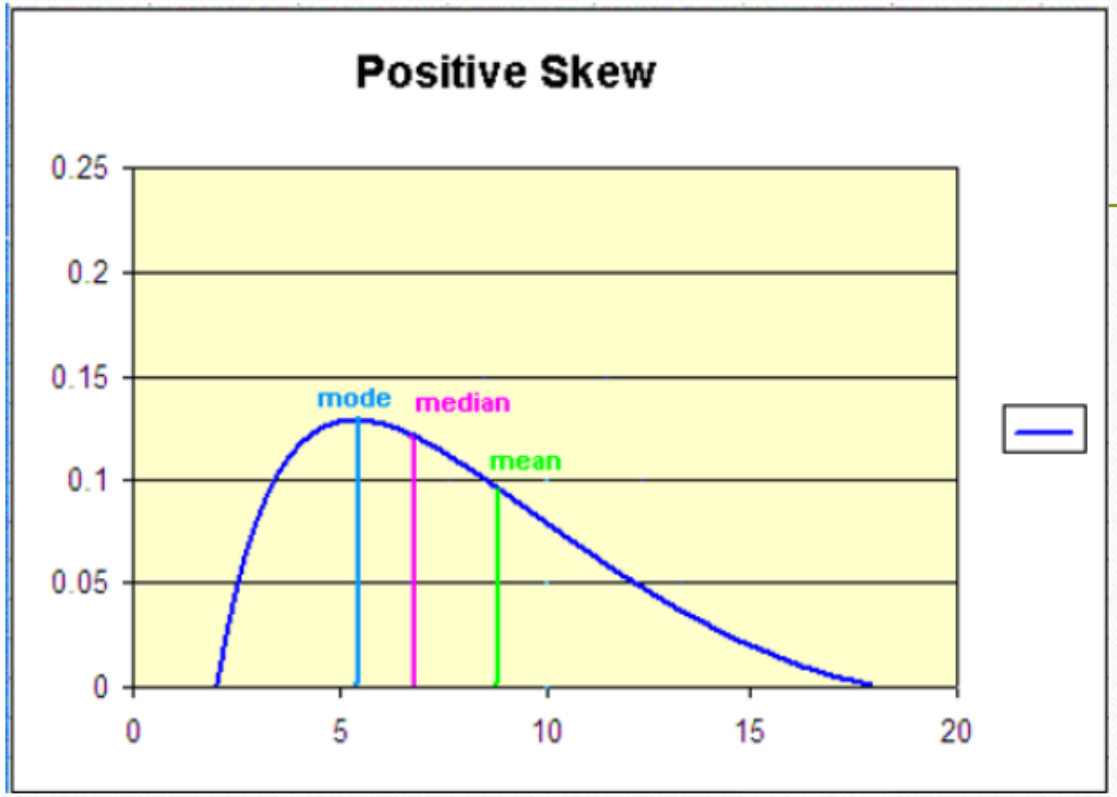
**B-) Bimodal symmetric histogram /** is identical on both sides of its central points with two modes

# Histograms



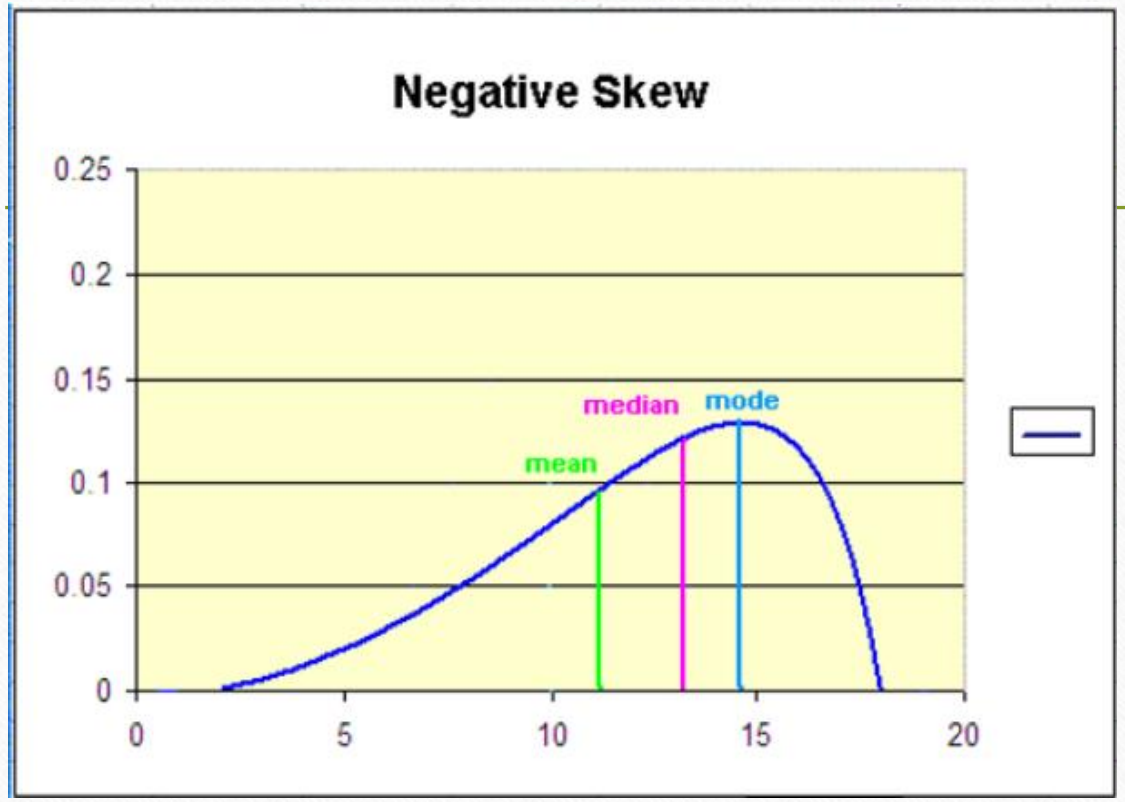

**C-) Uniform or rectangular histogram ( symmetric ) /** the frequencies of each class are the same or equal to each other.

# Histograms




**D-) Skewed to the right histogram ( positive skewed ) /** Most of data is shown in the left side of histogram and the tail on the other right side.

# Histograms




**E-) Skewed to the left histogram ( negative skewed ) /** Most of data is shown in the right side of histogram and the tail on the other left side.

# Polygons

**2- Polygons /** A graph formed by joining the midpoints of the tops of bars in a histogram with straight lines.

# Polygons


No. of Week
20
18
16
14
12
10
8
6
4
2
0
140-150
150-160
160-170
170-180
180-190
190-200
Cost of living index

# Stem and Leaf display

**3- Stem and Leaf display /** each value is divided into two portions -- a stem and a leaf. Then the leaves for each stem are shown separately in a display.

# Stem and Leaf display

**Construct a stem-and-leaf display for these data**

**Example / 22 , 26 , 27 , 31 , 33 , 35 , 42 , 44 , 46 , 57 , 58 , 59 , 61 , 63, 64 , 65 ,67**

# Solution

| Stem | Leaf |
|---|---|
| 2 | 2 6 7 |
| 3 | 1 3 5 |
| 4 | 2 4 6 |
| 5 | 7 8 9 |
| 6 | 1 3 4 5 7 |

# Reference


- **Prem S. Mann 1998, Introductory Statistics, 7th edn, New York, USA.**

# Good Luck for All Students

- Please do not hesitate to contact me if you have any questions.
- Dr. Ibrahim AL-Jaafari
- www.Alghamdi-Biostatistics.com
- Email. Bio-stat@Hotmail.com
- Mobile Number : 0553777925