

Chapter 1

1. Basic concepts, methods of data collection and presentation

1. 1. Introduction

1.1.1 Definition and Classification of Statistics

Definition: Statistics is the science of conducting studies to collect, organize, analyzed and draw conclusions from the data.

- In general , statistics can be defined into **two** senses.
 1. **In singular sense:** It is defined as the science that deals with the methods of collection, organization, analysis of data and interpretation of the results.
 2. **In plural sense:** It is defined as a set (aggregate) of numerical data or a quantitative aspects of facts.

1.1.2 Classification of Statistics

- Statistics can be classified into **two** broad areas.
 1. **Descriptive Statistics:** It is a part of statistics which can be used to organize and summarize masses of data.
 - ✓ The frequency distribution, measure of central tendencies such as mean and median, and measure of variation such as range and standard deviation belong to this category of statistics.
 - Example: The average age of students in this class is 21.
 2. **Inferential Statistics:** It is a major part of statistics which concerned with making decisions, inferences (conclusions) and forecasting about the population based on sample results.
 - ✓ It includes estimation and test of hypothesis about the population.
 - Example: Drinking decaffeinated coffee can raise cholesterol levels by 7%.

Exercise: Describe the following sentences whether inferential statistics or descriptive statistics. Suppose that the height of 6 randomly selected students from section 2 are the following: 160cm,165cm,175cm,170cm,180cm and 185cm.

1. The average height of six students is 172.5cm.
2. The average height of students in these section is not less than 172.5cm.
3. About half of the six students have the height more than 170cm.
4. The average height of students in section 2 is greater than that of section 1.

1.1.3 Stages in Statistical Investigation

According to the definition of statistics (in singular sense), there are **5 stages** in statistical investigation.

Stage 1: Collection of Data: It is a process of obtaining data.

Stage 2: Organization of Data: This includes

- ✓ Editing: measurement of how important it is .
- ✓ Classification: similar and differences.
- ✓ Tabulation: organization of data in row and column.

Stage 3: Presentation of Data: It is a process of showing our data in understandable way.

Example: charts, graphs and tables.

Stage 4: Analysis of Data: It is a process of extracting a useful characteristics associated with data.

Stage 5: Interpretation of Data (Inference): It is a process of making interpretations or conclusions from sample data for the totality of the population.

- ✓ It is the most difficult and risk stage. It needs professionals in statistics.

1.1.4 Definition of Some Basic Terms

Data: is any recordable interrelated observations.

Population: is the totality of all individuals of the phenomena under study.

Sample: It is a part of population selected in statistical manner to study the population.

Parameter: It is statistical value which refers to the *population characteristics*.

or it is a result obtained from the population.

Statistic: It is statistical value which refers to the *sample characteristics*.

or it is a result obtained from the sample.

Census: It is a process of studying a population at large.

Example: a researcher wants to study the academic performance of first year student in MTU. But for several constraints he cannot enumerate the whole students. So he took randomly 500 students and obtained the average GPA to be 2.58.

- a. Identify the population? b. Identify the sample? c. Identify the statistic?

1.1.5 Uses, Applications and Limitation of Statistics

Uses of Statistics

- a. It represents the facts in the form of numerical data.
- b. It condenses and summarizes mass of data into a few presentable, understandable and precise figures.
- c. It facilitates comparison of data.
- d. It helps in predicting future trends.
- e. It helps in formulating policies.

Applications of Statistics

- Statistics can be applied in almost all fields of study. Some of these are:
1. In health
 2. In education
 3. In agriculture etc

Limitations of Statistics

- ✓ It is not suited to the study of qualitative phenomena.
- ✓ It's results are true on the average. (It does not show the exact fact) like law of physics.
- ✓ It deals with a set (aggregate) of individuals not a single individual.
- ✓ It can be easily misused.
- ✓ Statistical interpretations requires a high degree of skill and understanding of the subject.

1.1.6 Types of Variables and Level of Measurements

Types of variables: There are **two** types of variables.

1. **Qualitative (Categorical) Variables:** are variables that can be placed into **distinct category** according to some characteristics. They are not numeric. They cannot be counted or measured.

- ✓ Example: gender, religion, color etc

2. **Quantitative Variables:** are variables which are **numerical in nature** and can be measured and counted.

- ✓ Example: height, weight, no of students, GPA etc.

✚ **Quantitative variables** can also divided into *discrete and continuous* variables.

- **Discrete variables:** are variables whose values are determined **by counting**.

Example: no of students in the class.

- **Continuous Variables:** are variables whose values are determined **by measuring** rather than counting.

Example: height of a person.

Exercise: are the following variables discrete or continuous?

- a. The no of correct answers on true false test.
- b. The duration of effectiveness of a pain medication.
- c. The weight of Sunday newspapers.

Measurement Scales (Levels)

- ✓ There are **4 types** of measurement scales. These are:

- | | |
|------------------|-------------------|
| 1. Nominal Scale | 3. Interval Scale |
| 2. Ordinal Scale | 4. Ratio Scale |

1. Nominal Scale: When the possible **categories** of a variable **have no a natural order** then the measurement is called nominal scale.

- ✓ we cannot apply any mathematical operations and inequalities.

Example: Blood type (A,B,AB,O) , sex (f,m), no's given to region (1,2,3,...)

2. Ordinal Scale: When the possible **categories** of a variable have **a natural order** then the measurement is called ordinal scale.

- ✓ we can apply any mathematical inequalities but we can not apply any mathematical operations.

Example: Economic status (low, medium, high), Education level (diploma, degree, master).

3. Interval Scale: It is a scale with **arbitrary zero point**, and zero does not shows a **total absence** of the quantity being measured.

- ✓ we can apply any mathematical inequalities.
✓ we can also apply addition and subtraction but we cannot form multiplication and division.

Example: a) The temperature of a certain area may be 0_C^0 . But this does not mean that there is no heat at all. It simply indicates that it is too cool.

b) The temperature of a certain areas may be 63_F^0 , 68_F^0 , 110_F^0 , 126_F^0 & 131_F^0 .

→ we can say that $68_F^0 > 63_F^0 \Rightarrow 68_F^0$ is warmer than 63_F^0 .

→ $68_F^0 - 63_F^0 = 131_F^0 - 126_F^0 \Rightarrow$ since equal temprature differences are equal.

- ✓ But we cannot say that 126_F^0 is twice as hot as 63_F^0 . Eventhogh $\frac{126}{63} = 2$.

- ✓ To show this change the scale to degree Celsius.

$$126_F^0 \Rightarrow \frac{5}{9}(126_F^0 - 32) = 52.2_C^0$$

$$63_F^0 \Rightarrow \frac{5}{9}(63_F^0 - 32) = 17.2_C^0$$

$$\Rightarrow 52.2_C^0 \text{ is more than 3 times } 17.2_C^0$$

4. Ratio Scale: It is a scale with **true zero point** and zero shows **a total absence** of the quantity being measured.

- ✓ we can apply any mathematical operation and inequalities.

Example: weight $x = 40kg, y = 80kg$.

$\Rightarrow y$ is twice heavy as x .

1.2. METHOD OF DATA COLLECTION AND PRESENTATION

1.2.1 Source and Types of Data

There are two types of data:

a) Primary Data

- Data collected by the investigator directly from the source.

Example: observe signs, measure characteristics, record symptoms and interview respondents, etc.

- Two activities involved: planning and measuring.

- ✓ Identify source and elements of the data.
- ✓ Decide whether to consider sample or census.
- ✓ If sampling is preferred, decide on sample size, selection method,... etc
- ✓ Decide measurement procedure.
- ✓ Set up the necessary organizational structure.

b) Secondary Data

- Data gathered or compiled from published and unpublished sources or files.

Example: Hospital records, vital statistics and registers, etc.

- When our source is secondary data check that:

- ✓ The type and objective of the situations.
- ✓ The purpose for which the data are collected and compatible with the present problem.
- ✓ The nature and classification of data is appropriate to our problem.
- ✓ There are no biases and misreporting in the published data.

Note: Data which are primary for one may be secondary for the other.

1.2.2 Methods of Data Collection

There are **three** major methods of data collection.

1. Observational or measurement.
2. Interview with questionnaires.
 - a. Face to face interview.
 - b. Telephone interview.
 - c. Self administered questionnaires returned by mail (mailed questionnaire).
3. The use of documentary sources

1. **Observational or measurement** (direct personal observation)

In this case data can be obtained through direct observation or measurement. This requires training and monitoring of the measurer to ensure the use of standard procedure.

- Provides accurate information but it is expensive and inconvenient.

Example: physical examination, clinical measurements, laboratory tests etc.

2. **Interview with questionnaires:** Hear one drafts a detailed questionnaire. These questionnaires can either be mailed to the respondent for filling and returning, or can put in charge of the enumerators who go around and fill them after obtaining the desired information.

✚ **Questionnaires:** are written documents which instruct the reader or listener to answer the questions written on it.

✚ **Respondents (Interviewees):** are individuals those who are answered the questions on the questionnaire.

✚ **Interviewers:** are individuals those who are recorded the responses given by the respondents.

a) **Face to Face Interviews** (questionnaires in charge of enumerators)

The interviewer knows exactly who is responding to the questionnaire.

Advantages

- The interviewer can help the respondent if he/she has difficulty in understanding the questions. The difficulty could be due to language, concentration or limited intellectual capacity.
- There is more flexibility in presenting the items; they can range from closed to open.
- There is the ability to use the method of skip patterns.
 - ✓ Skip patterns means skipping a questions or a group of questions which are not applicable.

Disadvantages

- ❖ Untrained interviewer may distort the meaning of the questions.
- ❖ Attribute of the interviewer may affect the responses due to:
 - a) Bias of the interviewer and b) his/her social or ethnic characteristics.
- ❖ It costs much in terms of time and money.

b) Telephone Interviews

Advantages

- It is less expensive in time and money compared with face to face interviews.
- The interviewer is able to help the respondent if he/she doesn't understand the question (as seen with face to face interview)
- Broad representative samples can be obtained for those who have telephone lines.

Disadvantage

- ✓ Under representation of those groups which do not have telephones.
- ✓ Problem with unlisted telephone number in the directory.
- ✓ Respondent may be substituted by another.

c) Self administered questionnaires returned by mail (mailed questionnaire)

- ✚ Here the questionnaire is mailed to the respondents to be filled. Sometimes it is known as self enumeration.

Advantages

- These are the cheapest.
- There is no need for trained interviewer.
- There is no interviewer bias.

Disadvantage

- ✓ Low response rate
- ✓ Uncompleted questionnaires due to omission or invalid responses.
- ✓ No assurance that the questionnaire was answered by the right person
- ✓ Needs intense follow up to get a high response rate.

3. The use of documentary sources

Extracting information from existing sources (e.g. Hospital records) is much less expensive than the other two methods. It can be an important source of data.

Limitation: It is difficult to get information needed, when records are compiled in unstandardized manner.

1.2.3 METHODS OF DATA PRESENTATION

✚ After having the collected and edited data, the next important step is to organize it. That is to present it in a readily comprehensible condensed form that aids to draw inferences from it. It is also necessary that the like be separated from the unlike ones.

✓ The presentation of data is broadly classified in to the following three categories:

- Tabular presentation (frequency distribution).
- Diagrammatical presentation and
- Graphical presentation.

✚ The process of arranging data in to classes or categories according to similarities technically is called *classification*.

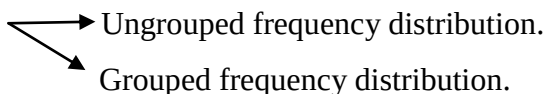
✓ Classification is a preliminary and it prepares the ground for proper presentation of data.

1.2.3.1 Tabular Presentation of Data (Frequency Distribution)

Definitions:

- **Raw data:** is a data which is collected in original form (survey), whether it may be counts or measurements.
- **Frequency (f):** is the number of observations (values) in a specific class of a distribution.
- **Frequency distribution(FD):** is the organization of raw data in table form, using classes and frequencies.

✓ Depending on the type of data, there are **two** basic types of frequency distributions:

- Qualitative (Categorical) frequency distribution and
- Quantitative frequency distribution 

✚ **NB:** The main purpose of grouping is now summarization and condensation of a masses of data.

1). Categorical (Qualitative) frequency Distribution:

It is often constructed for some data sets that can be placed in *a specific categories* such as nominal, or ordinal data's.

Example: A social worker collected the following data on marital status for 25 persons. (*M = married, S = single, W = widowed, D = divorced*). Construct a frequency distribution for the following data.

M	S	D	W	D
S	S	M	M	M
W	D	S	M	M
W	D	D	S	S
S	W	W	D	D

Solution: Since the data are qualitative (categorical), discrete classes can be used. There are four types of marital status M, S, D, and W. These types will be used as the classes for the distribution.

Classes	Frequency (f)
M	6
S	7
D	7
W	5

2). Quantitative frequency Distribution:

a). Ungrouped frequency Distribution:

It is often constructed for some data sets in which the number of "*distinct values*" are small. And also it is constructed for small set or data on discrete variable.

Steps for constructing ungrouped frequency distribution:

- Arrange the data in order of magnitude and then count the frequency.

Example: A survey taken in a restaurant shows that the following number of cups of coffee consumed with each meal. Construct an ungrouped frequency distribution for the following data.

0	2	2	1	1	2
3	5	3	2	2	2
1	0	1	2	4	2
0	1	0	1	4	4
2	2	0	1	1	5

Solution: First arrange the data in order of magnitude (in ascending order) and then count the frequency. The distinct values for these data are: 0,1,2,3,4 & 5. $\Rightarrow$ small.

No of cups	Frequency (f)
0	5
1	8
2	10
3	2
4	3
5	2
Total	30

- ✓ Each individual value is presented separately, that is why it is named ungrouped frequency distribution.

b). Grouped frequency Distribution:

When the number of "**distinct values**" of the data is too large, the data must be grouped in to classes. So, we divide the values into groups or class intervals, and then count the number of data values falling in each class interval.

Class intervals (CI): are a non-overlapping intervals such that each value in the set of observations can be placed in one, and only one, of the intervals.

Steps for constructing Grouped frequency Distribution

1. First arrange the data in ascending order.
2. Find the range (R) : $R = \text{Maximum} - \text{Minimum}$
3. Find the number of class intervals (k): It should be between 5 and 20. i.e. $5 \leq k \leq 20$ or **use Sturge's formula: $k = 1 + 3.322 \times \log_{10} n$.**

where: k is the number of class intervals desired and n is the total number of observations.

✚ **NB:** k must be rounded up/down to the nearest whole number.

4. Find the class width (w): It is the gap between two consecutive class intervals.

$$w = \frac{R}{k} \quad \text{and it is always rounded up.}$$

✓ When the data is given as

- Whole number "w" always rounded up to the next whole number. e.g. $w = 4.13 \approx 5$
- Tenth digit "w" always rounded up to the next tenth digit. For e.g. $w = 0.325 \approx 0.4$.
- Hundredth digit "w" always rounded up to the next hundredth digit. For e.g.
 $w = 2.532 \approx 2.54$; $w = 0.981 \approx 0.99$.

5. Find the class limits (CL): These are extreme values for each class. They are called lower and upper class limits.

- **Lower class limit (LCL):** The LCL of the first class interval should be equal to or smaller than the smallest observation in the data. i.e.

$$lcl_1 \leq \text{the smallest observation} \Rightarrow lcl_1 = \text{the smallest observation.}$$

✚ Continue to add the class width to this lower limit to get the rest of the lower limits. i.e. $lcl_{i+1} = lcl_i + w$, $i = 1, 2, \dots, k - 1$.

- **Upper class limit (UCL):** To find the upper class limit of the **first class**, subtract " u " from the lower limit of the **second class**. i.e. $ucl_1 = lcl_2 - u$.

✚ Then continue to add the class width to this upper limit to get the rest of the upper class limits. i.e. $ucl_{i+1} = ucl_i + w$, $i = 1, 2, \dots, k - 1$.

- ✓ where "**u**" is a unit measurement or the **smallest difference** between the two **nearest** observations in the data. It is usually taken as **1, 0.1, 0.01,...** as the data is given as **whole numbers, tenth digit, hundredth digit, ...** respectively.

6. Find the frequencies.

❖ **Class boundaries (CB):** are the set of exact limits or true limits. They are called lower and upper class boundaries.

- **Lower class boundary (LCB):** The **lcb** is obtained by subtracting half the unit of measurements from the **lcl** of the class. i.e.

$$lcb_i = lcl_i - \frac{u}{2} \quad \text{Note:} \quad lcb_{i+1} = lcb_i + w$$

- **Upper class boundary (UCB):** The **ucb** is obtained by adding half the unit of measurements from the **ucl** of the class. i.e.

$$ucb_i = ucl_i + \frac{u}{2} \quad \text{Note:} \quad ucb_{i+1} = ucb_i + w$$

❖ **Class marks (mid points) (m):** It is the average of **lcl** and **ucl** or **lcb** and **ucb**.

$$m_i = \frac{lcl_i + ucl_i}{2} \quad \text{or} \quad m_i = \frac{lcb_i + ucb_i}{2} \quad \text{Note:} \quad m_{i+1} = m_i + w$$

Modified frequency distribution

- ✓ Relative frequency (rf): $rf = \frac{f_i}{n}$
- ✓ Percentage relative frequency (%rf): $\%rf = \frac{f_i}{n} \times 100\%$
- ✓ **Cumulative frequency:** is the number of observations less than/more than or equal to a specific value.
- ✓ **Less than cumulative frequency (lcf):** it is the total frequency of all values less than or equal to the upper class boundary of a given class.
- ✓ **More than cumulative frequency (mcf):** it is the total frequency of all values greater than or equal to the lower class boundary of a given class.
- ✓ **Relative cumulative frequency (rcf):** it is the cumulative frequency divided by the total frequency.

Example: Construct a grouped frequency distribution for the following data.

11	29	6	33	14	31	22	27	19	20
18	17	22	38	23	21	26	34	39	27

Solutions:

Step 1: Arrange the data in ascending order.

Step 2: Find the range (R) : $R = \text{Max} - \text{Min} = 39 - 6 = 33$.

Step 3: Select the number of classes desired using Sturge's formula;

$$k = 1 + 3.322 \times \log n = k = 1 + 3.322 \times \log(20) = 5.32 \approx 5 \text{ (rounding down).}$$

Step 4: Find the class width; $w = \frac{R}{k} = w = \frac{33}{5} = 6.6 \approx 7$ (rounding up).

Step 5: Find the lower and the upper class limits.

Select the starting point, let it be the smallest observation.

- 6, 13, 20, 27, 34 are the lower class limits.

Find the upper class limits; e.g. the first upper class limit (ucl_1) = $13 - U = 13 - 1 = 12$.

$u = 1$ since the data is given as a whole number.

- 12, 19, 26, 33, 40 are the upper class limits.

✚ So combining **lcl and ucl**, one can construct the following classes.

Class limits
6 – 12
13 – 19
20 – 26
27 – 33
34 – 40

Step 6: Find the class boundaries;

$$E.g. \text{ for class 1; } lcb_1 = 6 - \frac{u}{2} = 6 - \frac{1}{2} = 5.5$$

$$ucb_1 = 12 + \frac{u}{2} = 12 + \frac{1}{2} = 12.5$$

- Then continue adding **w** on both boundaries to obtain the rest boundaries. By doing so one can obtain the following classes.

Class boundary
5.5 – 12.5
12.5 – 19.5
19.5 – 26.5
26.5 – 33.5
33.5 – 39.5

Step 7: Find the frequencies.

- ✓ The complete frequency distribution is given as follows:

Class limit	Class boundary	Class Mark	f	Lcf	Mcf	rf.	%rf	%rcf
6 – 12	5.5 – 12.5	9	2	$\leq 12.5 (\leq 12) = 2$	$\geq 5.5 (\geq 6) = 20$	0.10	10%	10%
13 – 19	12.5 – 19.5	16	4	$\leq 19.5 (\leq 19) = 6$	$\geq 12.5 (\geq 13) = 18$	0.20	20%	30%
20 – 26	19.5 – 26.5	23	6	$\leq 26.5 (\leq 26) = 12$	$\geq 19.5 (\geq 20) = 14$	0.30	30%	60%
27 – 33	26.5 – 33.5	30	5	$\leq 33.5 (\leq 33) = 17$	$\geq 26.5 (\geq 27) = 8$	0.25	25%	85%
34 – 40	33.5 – 39.5	37	3	$\leq 39.5 (\leq 39) = 20$	$\geq 33.5 (\geq 34) = 3$	0.15	15%	100%

1.2.3.2 DIAGRAMMATICAL PRESENTATION OF DATA

These are techniques for presenting data in visual displays using geometric and pictures.

Importance:

- They have greater attraction.
 - They facilitate comparison.
 - They are easy to understand.
- ✓ Diagrams are appropriate for presenting discrete data.
- The two most commonly used diagrammatic presentation for discrete as well as qualitative data are:
- Bar charts and
 - Pie charts

1. Bar chart

There are **three** types of bar charts. These are:

- I) Simple bar chart II) Component bar chart III) Multiple bar chart

a). Simple Bar chart:

It is a chart which is used to present data that has **only one** variable. It shows changes in the totals of different categories.

Example: Construct a simple bar chart for the following table showing annual cases of HIV patients reported in Ethiopia as of July 31, 1993.

Year of report	1986	1987	1988	1989	1990	1991	1992	1993
Cases	2	17	87	190	448	885	3256	2814

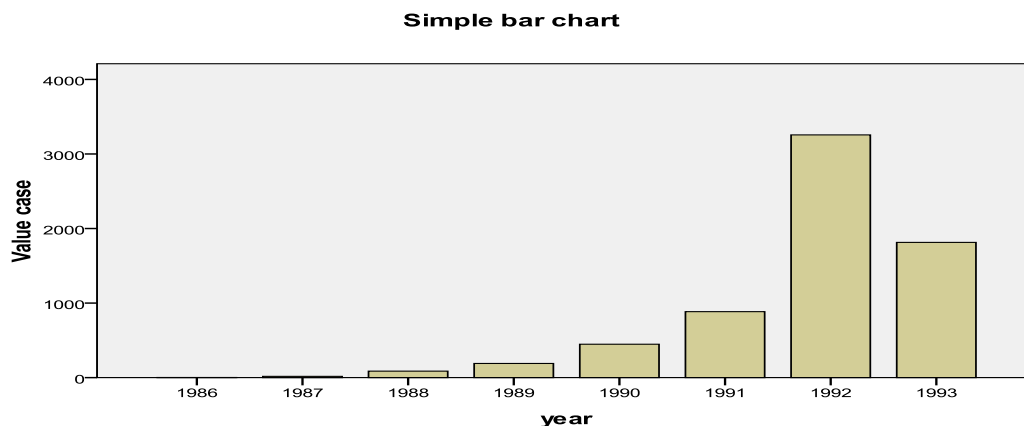


Figure 1 : Annual HIV cases reported in Ethiopia as of 31 July 1993

Source : Addis Zemene Nehasse 27,1985

b). Component Bar chart

It is used to present data which have *more than one* variable. For each category the bars are subdivided in to components to allow comparison between parts. The bars represent the total value of a variable with each total broken in to its component parts and different colors or designs are used for identifications.

Example

Construct component bar chart for the number of children who were vaccinated with DPT, POLIO and BCG antigens in Mizan-Aman General Hospital in 1979 E.C.

Antigen	Sex		Total
	Male	Female	
DPT	250	300	550
Polio	300	320	620
BCG	200	210	410

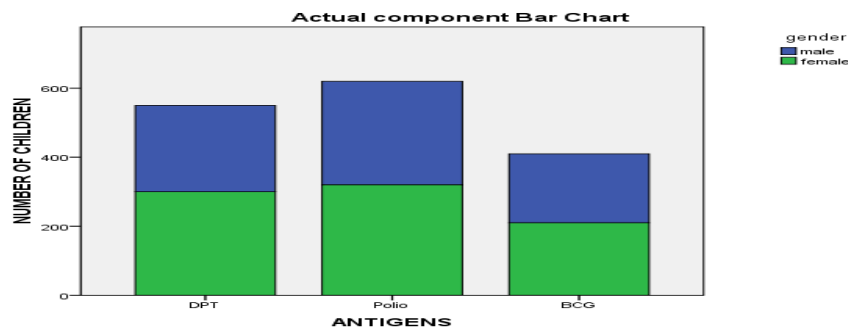


Figure 2 Number of children who were vaccinated against DPT, POLIO and BCG, Jimma Hospital 1979 E.C.

c). Multiple Bar chart

- ✓ These are used to display data on *more than one* variable.
- ✓ They are used for comparing different variables at the same time.

Example: draw a multiple bar chart for the above vaccination data.

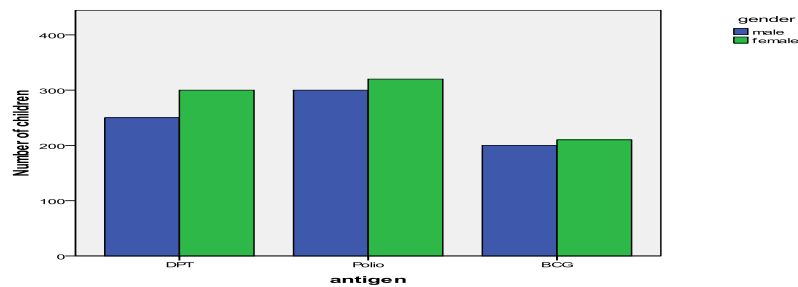


Figure 4 Number of Children who received the DPT, POLIO and BCG antigen Jimma Hospital, 1979 E.C.

2. Pie-Chart

It is used to show the partitioning of a total data into its component parts using circles. The circles should be divided into sectors proportional to the frequencies of the categories they represent.

Steps to draw a pie chart

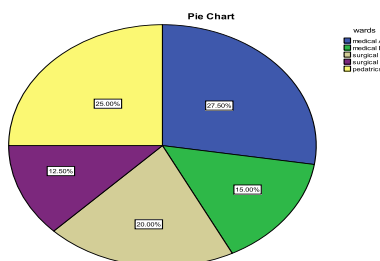
1. Convert frequencies into percentage relative frequency.
2. Draw a circle of any radius.
3. Convert percentage relative frequencies into degree measures.

$$\text{angle of a sector} = \frac{360^\circ \times \%rf}{100\%}$$

Example

Draw the pie chart for the following data. First construct a table providing the central angles.

Wards	Frequency	Percentage rf	Central angle
Medical A	55	27.5%	99
Medical B	30	15%	54
Surgical A	40	20%	72
Surgical B	25	12.5%	45
Pediatrics	50	25%	90
Total	200	100%	360



1.2.3.3 Graphical presentation of data

a) Histogram

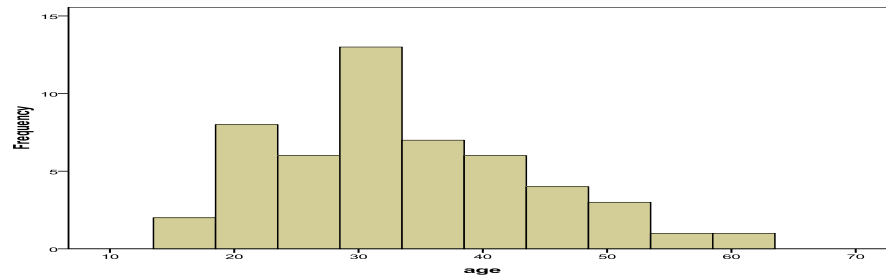
It presents a grouped frequency distribution of a continuous type. It is drawn by making class boundaries in the x-axis and frequencies in the y-axis.

Example: Draw a histogram for the following grouped age data.

Class limit	Class boundaries	Mid point	Frequency
15-19	14.5-19.5	17	2
20-24	19.5-24.5	22	8
25-29	24.5-29.5	27	6
30-34	29.5-34.5	32	12
35-39	34.5-39.5	37	7
40-44	39.5-44.5	42	6
45-49	44.5-49.5	47	4

50-54	49.5-54.5	52	3
55-59	54.5-59.5	57	1
60-64	59.5-64.5	62	1

Histogram

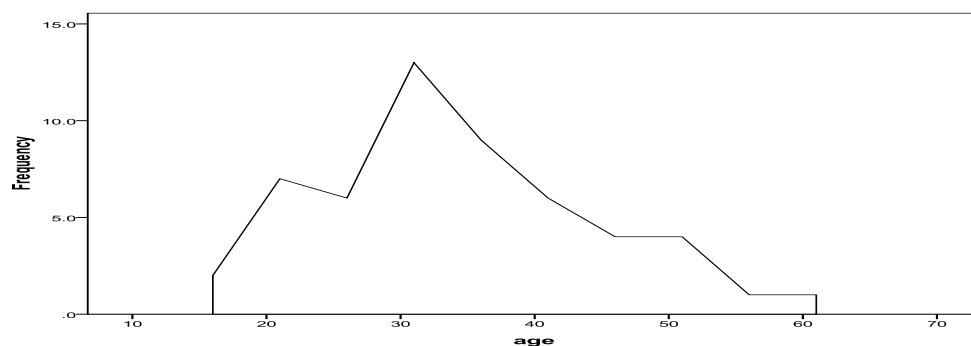


b) Frequency polygon

It is a multi-sided figure which is drawn by plotting the class marks (midpoints) in the x-axis and the frequencies in the y-axis. Then connect the points with straight lines and extend these lines on both ends so that it reaches the horizontal axis at the class mid points. This allows the total area to be enclosed.

Example: draw the frequency polygon for the following age data.

Class limit	Mid point	Frequency
15-19	17	2
20-24	22	8
25-29	27	6
30-34	32	12
35-39	37	7
40-44	42	6
45-49	47	4
50-54	52	3
55-59	57	1
60-64	62	1



Note: The total area under the frequency polygon is equal to the area under the histogram.

c) Ogives or cumulative frequency polygon (curve)

It plotted in association with the class boundaries on the x- axis and the cumulative frequencies on the y- axis. Then connect the points with straight lines.

- ✓ The curves obtained are called the *“less than”* and *“more than”* ogives (curves).
- **Less than ogive:** It is plotted by *“UCB”* in the x-axis against the *“lcf”* in the y-axis.
- **More than ogive:** It is plotted by *“LCB”* in the x-axis against the *“mcf”* in the y-axis.

Example: draw the less than and more than ogives for the following age data.

Class limit	Frequency	LCF	More than
23-26	3	$\leq 26.5 (\leq 26) = 3$	$\geq 22.5 (\geq 23) = 20$
27-30	4	$\leq 30.5 (\leq 30) = 7$	$\geq 26.5 (\geq 27) = 17$
31-34	3	$\leq 34.5 (\leq 34) = 10$	$\geq 30.5 (\geq 31) = 13$
35-38	5	$\leq 38.5 (\leq 38) = 15$	$\geq 34.5 (\geq 35) = 10$
39-42	5	$\leq 42.5 (\leq 42) = 20$	$\geq 38.5 (\geq 39) = 5$

