Introduction to Statistics and Probability I

Dr. E.N.N. Nortey
&
Dr. C.Chapman-Wardy

Department of Statistics & Actuarial Science
University of Ghana

Outline

UNIVERSITY OF GHANA

Introduction

- We introduce two important concepts: Statistics and Probability
- There are two types of variables used in many fields, referred to as **non-stochastic** or deterministic and **stochastic** or random variables.
- Stochastic variables are random variables that have an associated probability structure.
- eg. tossing a die- we don't know in advance what number will come up.
- Non-stochastic variables are deterministic in nature without a probability attachment.
- Interest and annuity calculations based on fixed time periods.
- Number of females in STAT 111 class.

Purpose and Learning outcomes

- The purpose of this presentation is to equip students with basic ideas of statistics, its use and applications.
- By the end of this lesson, students should be able to;
  1. Define and explain the meaning of Statistics and its usefulness
  2. Identify the types of data and how to use them
  3. identify three key areas where statistics can be applied
  4. Explain the levels of measurement in statistics and its usefulness
  5. Distinquish between samples and population, census and sample survey
  6. Explain the difference between discrete and continuous variables
- The evaluation of stochastic variables requires the use of basic probability and statistical tools.
- We now focus on statistical tools and later on turn our attention to concepts in probability

Definition of Statistics I

- Decision makers make better decisions when they use all available information in an effective and meaningful way.
- The primary role of statistics is to provide decision makers with methods for obtaining and analyzing information to help make these decisions.
- Statistics is used to answer long-range planning questions.
- The word **Statistics** has two meanings. In the real common, usage, statistics refers to numerical facts and figures.
- For example: the numbers that represent the diastolic blood pressure of a patient, the heart rate of a student, and the starting salary of a typical university graduate are examples of Statistics in this sense of the word.

UNIVERSITY OF GHANA

Definition of Statistics II

- The second meaning of statistics refers to the field or discipline of study. Statistics in this sense of the word, is a field of study concerned with:
  1. The collection, organization, and summarization of data, and
  2. The drawing of inferences about a body of data (numerical facts) when only a part of the data is observed.
- In general, statistics consist of a set of methods and rules for organizing and interpreting observations.
- These statistical procedures help ensure that observations (data) are presented and interpreted in an accurate and informative way.

UNIVERSITY OF GHANA

Definition of Statistics III

- Statistics, like almost all fields of study has two aspects: theoretical and applied statistics.
  1. Theoretical or mathematical statistics deals with the development, derivation, and proof of statistical theorems, formulas, rules, and laws.
  2. Applied statistics involves the application of those theorems, formulas, rules and laws to solve real-world problems.
- **Summary:** The role of statistics is to help you collect, organize, summarize, analyze and communicate quantitative information.

UNIVERSITY OF GHANA

Uses of Statistics I

- Statistical techniques are used extensively by managers in marketing, accounting, quality control, consumers, professional, sports people, hospital administrators, educators, and by politicians, physicians, etc.
- Statistics as a subject is used for description, comparison, projection, prediction, decision making etc.
- **Statistics and Government:** Governments need to correctly collect process and analyze statistical data on national output, earnings, expenditure, imports and exports, employment, population growth and decline, health in order to make right decisions.
- **Statistics and the District Council:** District councils need statistics on education welfare, transport, infrastructure, and recreational need etc. to plan the development of the area.

Uses of Statistics II

- **Statistics and Business:** Firms need statistics on production sales, payrolls, capital expenditure, and depreciation etc to effectively make correct decisions and projections.
- **Statistics and Other Professions:** Statistical analysis is used in practically every profession. It is used for example in
  1. testing the efficiency of an alternative production technique in industries
  2. testing the effectiveness of a new drug in medicine
  3. testing the effectiveness of fertilizer on a particular crop in agriculture
  4. analyzing the results of a drug rehabilitation programme in sociology
  5. testing the production design or package that maximizes sales in business
  6. forecasting voting patterns in politics

UNIVERSITY OF GHANA

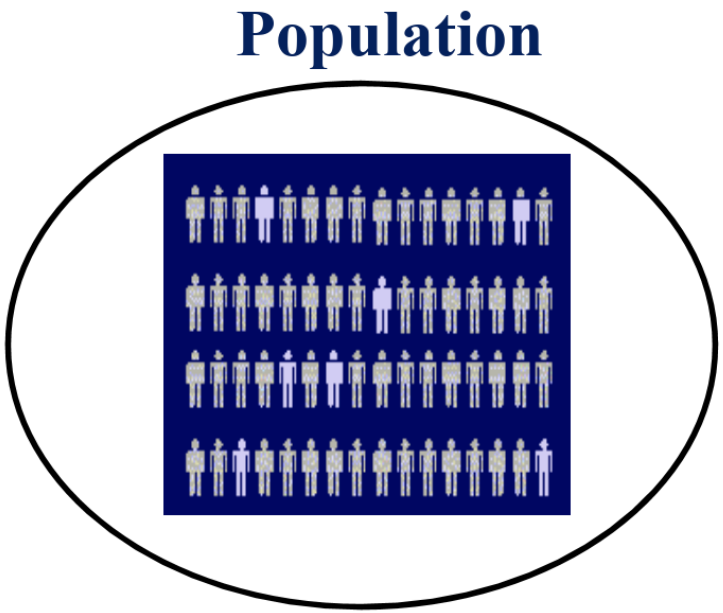
Definition of Terminologies in Statistics I

- Broadly speaking applied statistics can be divided into two areas:
  1. Descriptive statistics
  2. Inferential statistics.
- **Descriptive statistics** comprises of methods for organizing, presenting, and summarizing data using tables, graphs and summary measures. It is about presenting data in an informative way.
- **Inferential statistics** comprise methods that use sample results to help make decisions or predictions about a population.
- **Variables:** A variable is a characteristic or condition that can change or take on different values for different elements. Examples are diastolic blood pressure, heart rate, or STAT 111 test score.

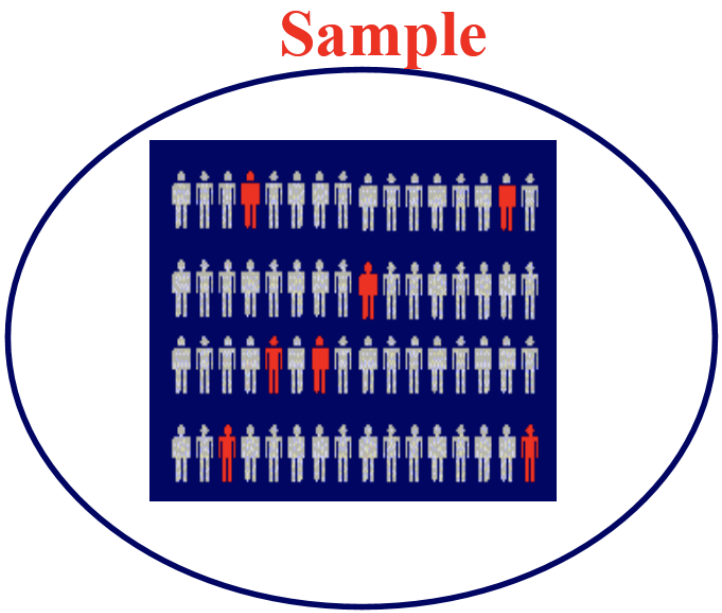
Definition of Terminologies in Statistics II

- A **population** is the collection of all persons, places, or things of interest in a particular study.
- It is convenient to refer to the individual person, places, or things of which the population is composed as the elements of the population. Frequently, especially when the elements of a population are human beings, we refer to an individual member of the population as a subject.
- A **sample** is a set of individuals, places, or things selected from a population, usually intended to represent the population in a study.
- A descriptive value for a population is called a **parameter** and a descriptive value for a sample is called a **statistic**.
- The discrepancy between a sample statistic and its population parameter is called **sampling error**.

Population vs. Sample I



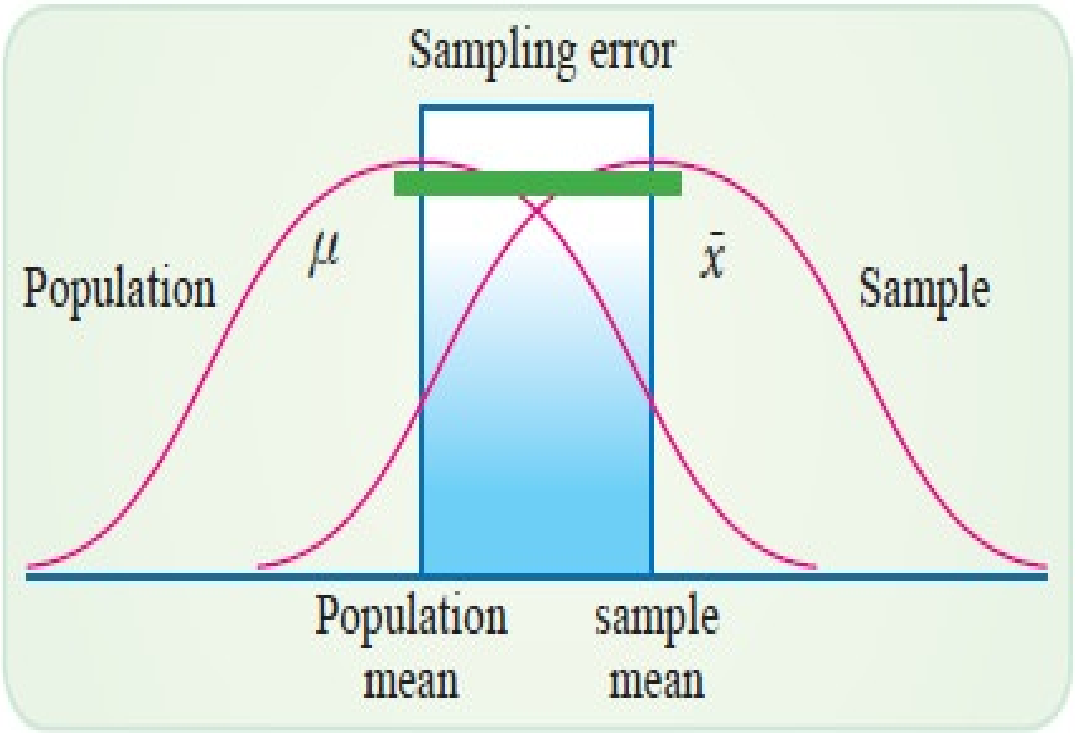
Measures used to describe the population are called **parameters**



Measures computed from sample data are called **statistics**

Population vs. Sample II

- A descriptive value for a population is called a **parameter** and a descriptive value for a sample is called a **statistic**.
- The discrepancy between a sample statistic and its population parameter is called **sampling error**.



Census and Sample Survey

- A survey which includes every member of the population is called a **census** or a complete enumeration.
- While a survey which collect information from only a portion of the population is called a **sample survey**.
- **Methods of Collecting Sample from the Population**
- **simple random sample** (each sample of the same size has an equal chance of being selected)
- **systematic sample** (randomly select a starting point and take every n-th piece of data from a listing of the population)
- **stratified sample** (divide the population into groups called strata and then take a sample from each stratum)
- **cluster sample** (divide the population into strata and then randomly select some of the strata.
  All the members from these strata are in the cluster sample.)

UNIVERSITY OF GHANA

Types of Data

- Data are the values (observations or measurements) that the variable can assumed. For example, a set of scores on a statistics test.
- The sources of statistical data can be put under two main categories; primary and secondary sources.
- When data used in a statistical study are collected under the control and supervision of the person or organization making the particular study, they are termed **Primary data**.
- **Secondary data** are data originally collected not under the supervision of the person or organization using such data.

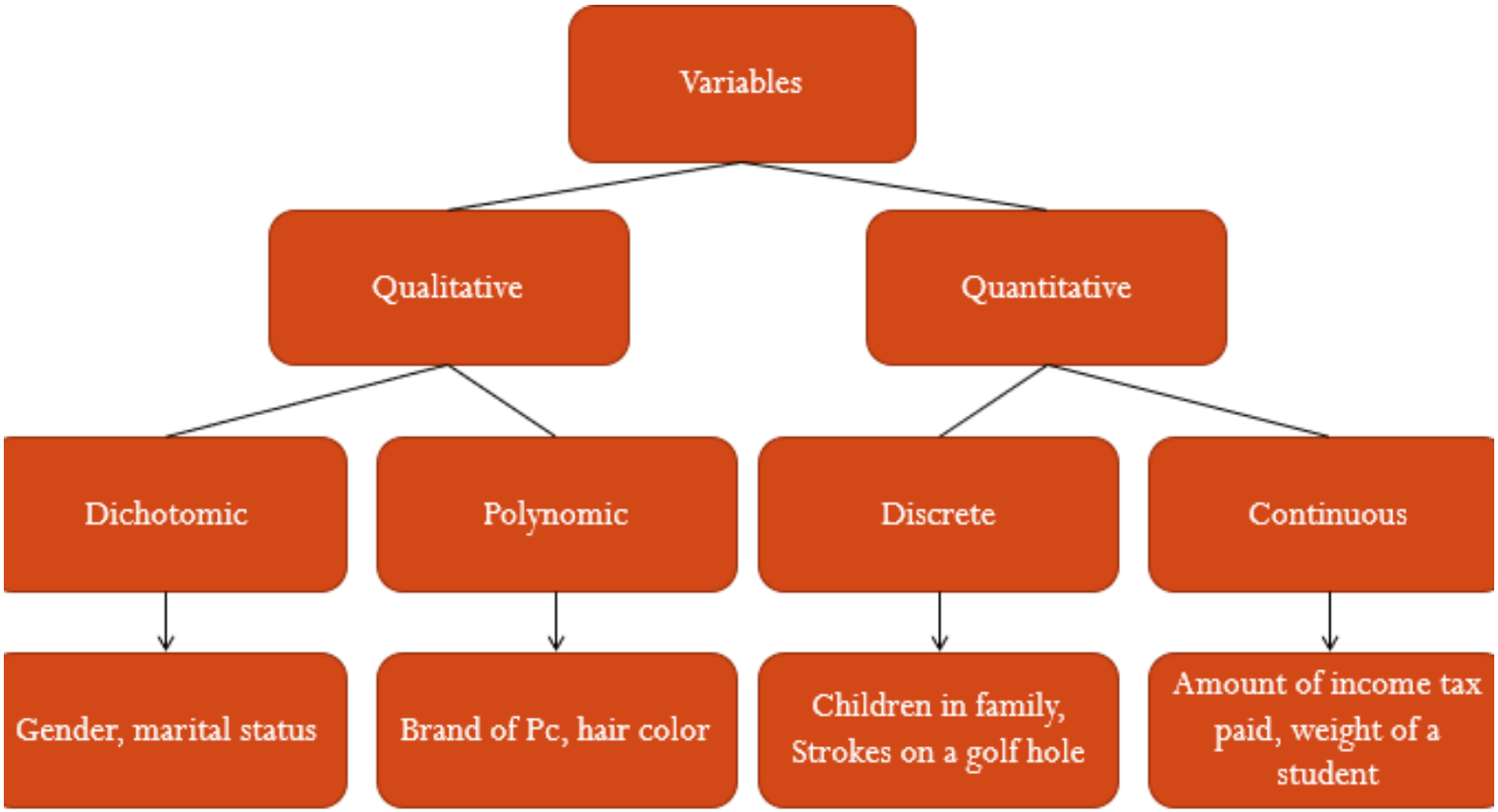
UNIVERSITY OF GHANA

Sources of Data

- **Primary Sources**: The data collector is the one using the data for analysis
  - Data from a political survey
  - Data collected from an experiment
  - Observed data
- **Secondary Sources**: The person performing data analysis is not the data collector
  - Analyzing census data
  - Examining data from print journals or data published on the internet.

UNIVERSITY OF GHANA

Types of Variables I



UNIVERSITY OF GHANA

Types of Variables II

- When the values of a variable obtained arise as a result of chance, the variable is called a random variable.
- **Qualitative variables** are variables that can be placed into categories or groups, according to some characteristic or attribute. For example, if subjects are classified according to gender (male or female), then the variable "gender" is qualitative.
- **Quantitative variables** are numerical in nature and can be ordered and ranked. For example, the variable "height" is numerical, and people can be ranked according to the value of the heights. Other examples of quantitative variables are age, weight, income, and body temperature.

Types of Variables III

- Random quantitative variables may be either discrete or continuous.
- Random variables that can assume a countable number of values are called **discrete**.
- Random variables that can assume values corresponding to any points contained in one or more intervals are called **continuous**.
- In other words, a continuous random variable has infinitely many values, and those values can be associated with measurements on a continuous scale without gaps or interruptions.

Example 1

Identify each of the following examples as attribute (qualitative) or numerical (quantitative) variables.

1. The residence hall for each student in a statistics class.
2. The amount of gasoline pumped by the next 10 customers at the local Unimart.
3. The amount of radon in the basement of each of 25 homes in a new development.
4. The color of the baseball cap worn by each of 20 students.
5. The length of time to complete a mathematics homework assignment.
6. The state in which each truck is registered when stopped and inspected at a weigh station.

Example 2

Identify each of the following as examples of qualitative or numerical variables:

1. The temperature in Barrow, Alaska at 12:00 pm on any given day.
2. The make of automobile driven by each faculty member.
3. Whether or not a 6 volts lantern battery is defective.
4. The weight of a lead pencil.
5. The length of time billed for a long distance telephone call.
6. The brand of cereal children eat for breakfast.
7. The type of book taken out of the library by an adult.

Levels of Measurement I

- They offer a quantitative definition of the variable attributes.
- Understanding different scales of measurement allows you to see the different types of data you can gather.
- And helps determine the kind of statistical analysis required
- Data can also be classified by levels of measurement, also called the scales of measurement. There are four levels of measurement: nominal, ordinal, interval, and ratio.
- The term **nominal level** are variables that are simply labelled with no specific order. eg. religious affiliation, gender,hall of residence etc.

UNIVERSITY OF GHANA

Levels of Measurement II

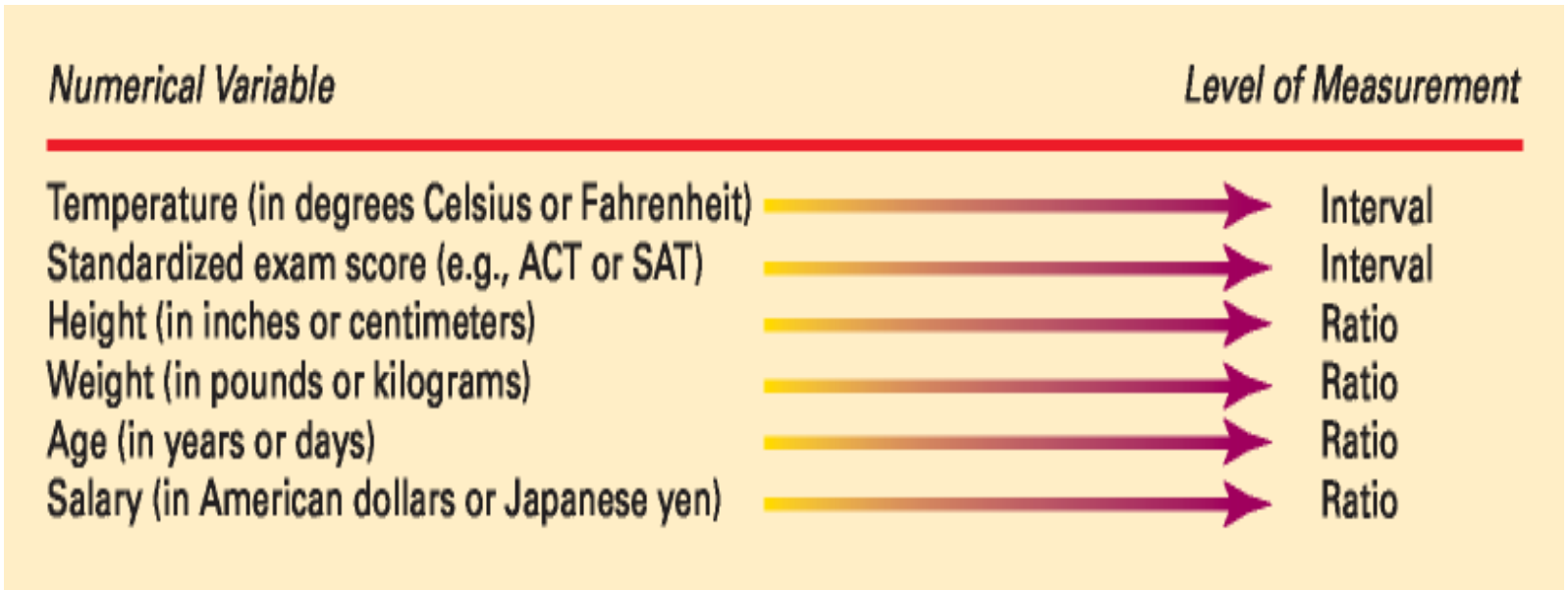
- In an **ordinal level**, kind of categorical data with a set order or scale. A student rank**s** her level of satisfaction from 1-5 of life on campus. Eg. ; ranking peoples opinion on a subject matter usually, 5,7 or 10 point Linkert scale. 1. Very satisfied, 2. Satisfied 3. Indifferent 4.Dissatisfied 5. Very dissatisfied.
- another eg. Child(0-12yrs), Teenager(13-19yrs), Youth(20-35), Middle age(36-58), Old(59yrs+).

UNIVERSITY OF GHANA

Levels of Measurement III

- **Interval data** classifies and orders a measurement. it specifies that a distance between each interval on a scale is equivalent from low interval to high interval.
- It represents a higher level of measurement than ordinal scale.eg 90 to 100 equivalent to 110 to 120.
- Does not represent a true zero
- can add, substract and multiply but cannot find ratio between them.
- it can represents values below zero. eg temperature, you can measure temperature below zero degree celsius such as -10 degrees
- The **ratio scale** it represents true zero. eg. height and weight, age,money
- can add, subtract, multiple and can find ratio.

Levels of Measurement IV



Example 3

Identify each of the following as examples of (1) nominal, (2) ordinal, (3) discrete, or (4) continuous variables:

1. The length of time until a pain reliever begins to work.
2. The number of chocolate chips in a cookie.
3. The number of colors used in a statistics textbook.
4. The brand of refrigerator in a home.
5. The overall satisfaction rating of a new car.
6. The number of files on a computer's hard disk.
7. The pH level of the water in a swimming pool.
8. The number of staples in a stapler.

Microsoft Excel Terms

- When you use Microsoft Excel, you place the data you have collected in worksheets.
- The intersections of the columns and rows of worksheets form boxes called cells.
- If you want to refer to a group of cells that forms a contiguous rectangular area, you can use a cell range.
- Worksheets exist inside a workbook, a collection of worksheets and other types of sheets, including chart sheets that help visualize data.

UNIVERSITY OF GHANA

Microsoft Excel Terms

- When you use Microsoft Excel, you place the data you have collected in worksheets.
- The intersections of the columns and rows of worksheets form boxes called cells.
- If you want to refer to a group of cells that forms a contiguous rectangular area, you can use a cell range.
- Worksheets exist inside a workbook, a collection of worksheets and other types of sheets, including chart sheets that help visualize data.

UNIVERSITY OF GHANA

Designing Effective Worksheets

- You should associate column cell ranges with variables.
- You do not skip any rows as you enter data, so column cell ranges will never contain any empty cells.
- Place all the variables on a worksheet that is separate from the worksheet containing the statistical results.
- Allow the user to be able to explicitly see the chain of calculations from the starting data.
- Create two copies of your worksheets: one optimized for the screen, the other for the printer.

UNIVERSITY OF GHANA

Lecture 1 Summary

In this lecture, we have

- Introduced key definitions:
  - Population vs. Sample
  - Primary vs. Secondary data types
  - Categorical vs. Numerical data
- Examined descriptive vs. inferential statistics
- Reviewed data types and measurement levels
- Discussed Microsoft Excel terms and tips

Lecture 2: Data Reduction Techniques

UNIVERSITY OF GHANA

Introduction

- Every statistical effort begins with the process of collecting the necessary data to be used in the study.
- The management of data is a major challenge to organisations of all kind.
- Data in its natural form is large or meaningless.
- The process of putting data in such a way that meaning can be made is known as **Data Reduction**.
- To determine the significance use of the data, it must first be organise into some form so that at a mere glance one can visualise the data and draw reasonable conclusions.

Data Reduction Techniques

- Statistical tools or techniques that are useful for organising data include:
  1. Frequency Tables
  2. Cross tabulation for bivariate data (Contingency tables)
  3. Stem and leaf display
  4. Bar charts
  5. Pie Charts
  6. Scatter plots
  7. Histograms
  8. Box-and-whisker plots
  9. Line graphs
  10. Ogives and Frequency polygons

Frequency Distribution

- The method of organising the raw data into separate classes with a count of the number of elements in each class is known as **Frequency Distribution**.
- The classes are mutually exclusive (i.e. non-overlapping).
- Example: Suppose twenty students of STAT 111 are asked to choose the preferences of the following football clubs: Arsenal, Chelsea, Liverpool and Barcelona. The preferences of these twenty students can be display on table as follows:

Football Clubs	No. of Students
Arsenal	5
Barcelona	7
Chelsea	3
Liverpool	5

Frequency Distribution

- When each class in the table consiste of a single value, the distribution is called ungrouped frequency distribution table.
- Conversely, if each class consists of a range of values , then its called grouped frequency distribution table.
- Examples of Ungrouped and Grouped frequency distribution tables.

Marks	Frequency	Marks	Frequency
10	6	0-10	6
20	8	11-20	8
30	7	21-30	7
40	11	31-40	11
50	3	41-50	3

Types of Frequency Tables

- There are three types of Frequency distribution tables, namely;
  1. Frequency distribution table (normal)
  2. Relative frequency distribution table
  3. Cumulative frequency distribution table
- **Relative frequency table:** It is the ratio of each frequency to the total frequency

  i.e. $RF = \dfrac{\text{frequency of class i}}{\text{total frequency}} = \dfrac{f_i}{\sum f_i}$
- When the relative frequency is multiplied by 100, the result is the percentage frequency.
- **Cumulative Frequency Distribution table:** is obtained by summing the frequencies (relative frequencies) of all classes up to the specific class.

Example

- The following are the ages of 100 level 100 students from the department of Statistics and Actuarial Science.

Ages	Tally	Freq.	Relative Freq.	Cumulative Freq.	C.R.F.
16		10	0.10	10	0.10
17		15	0.15	25	0.25
18		25	0.25	50	0.50
19		19	0.19	69	0.69
20		11	0.11	80	0.80
21		7	0.07	87	0.87
22		5	0.05	92	0.92
23		3	0.03	95	0.95
24		3	0.03	98	0.98
25		2	0.02	100	1.00

Grouped Frequency Distribution

- Grouped data is data given in intervals. Grouped frequency distributions are mostly used for continuous data.

- The number of classes or intervals may be determine by Sturge's rule which says that

  Number of Intervals, $I = 1 + 3.322 \log_{10} n$,
  where n = number of observations.

- To determine the class width, we use the formula

$$\text{Class width} = \frac{\text{Range}}{\text{No. of Classes}}$$

Example

Given the following set of data, we want to construct a grouped frequency distribution. Find the number of classes or intervals required and the class width.

26	18	21	34	18
38	22	27	22	30
25	25	38	29	20
24	28	32	18	33
21	18	19	24	26
18	23	24	27	32

Contingency Tables

- A contingency table, sometimes called a two-way frequency table, is a tabular mechanism with at least two rows and two columns used in statistics to present categorical data in terms of frequency counts.
- More precisely, an $r \times c$ contingency table shows the observed frequency of two variables, the observed frequencies of which are arranged into r rows and c columns.
- The intersection of a row and a column of a contingency table is called a **cell**.
- Example: A two rows and five columns contingency table, showing the results of 2200 adults classified by two variables, namely gender and favorite way to eat ice cream.

Gender	Cup	Cone	Sundae	Sandwich
Male	592	300	204	104
Female	410	335	180	75

Stem-and-leaf plots

- A good way to present both continuous and discrete data for sample sizes of less than 200 or so is to use a stem-and-leaf plot.
- This plot is similar to a bar chart or histogram, but contains more information.
- As with a histogram, we normally want 5–12 intervals of equal size which span the observations.
- However, for a stem-and-leaf plot, the widths of these intervals must be 0.2, 0.5 or 1.0 times a power of 10, and we are not free to choose the end-points of the bins.
- Each score is divided into a **stem** consisting of the first digit or digits and a **leaf** consisting of the remaining digits.
-

Example

- The raw data below show sum of scores of 30 students in two STAT 112 interim assessments. Organised the data as stem-and-leaf display.

- Use a stem unit of Tens and a leaf unit of ones.

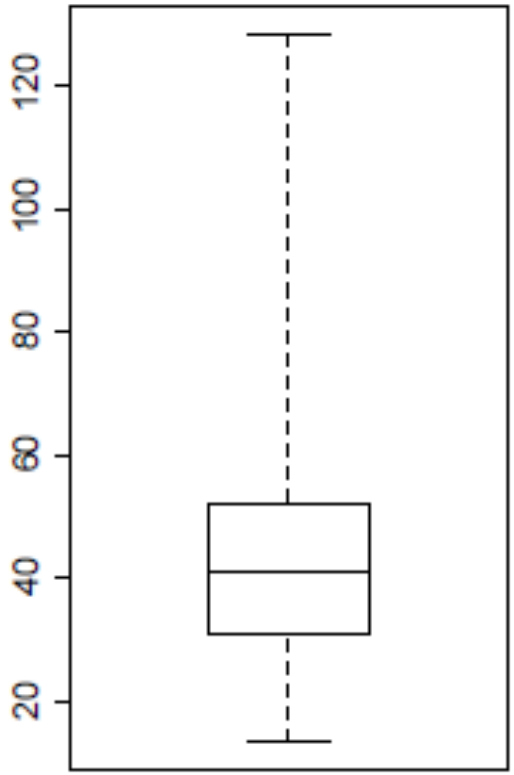
200	204	209	210	211	212
217	218	218	222	227	227
217	228	230	231	233	237
238	241	242	242	243	247
251	252	253	254	256	260

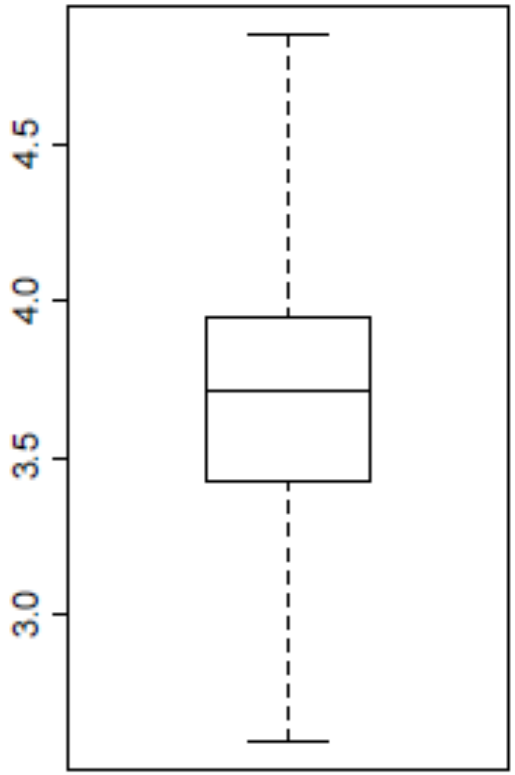
Box-and-whisker plots

- This is a useful graphical description of the main features of a set of observations.
- There are many variations on the box plot. The simplest form is constructed by drawing a rectangular box which stretches from the lower quartile to the upper quartile, and is divided in two at the median.
- From each end of the box, a line is drawn to the maximum and minimum observations.
- These lines are sometimes called whiskers, hence the name.
- The box plot provide an image of the concentration of the data and also show how far the extreme values are from most of the data.
- In brief, the box plot gives a good and quick picture of the data.

Example

- The following figures below show box-and-whisker plots of two different datasets.





Example

- For example, consider the dataset below and use it to construct a box-and-whisker plot.

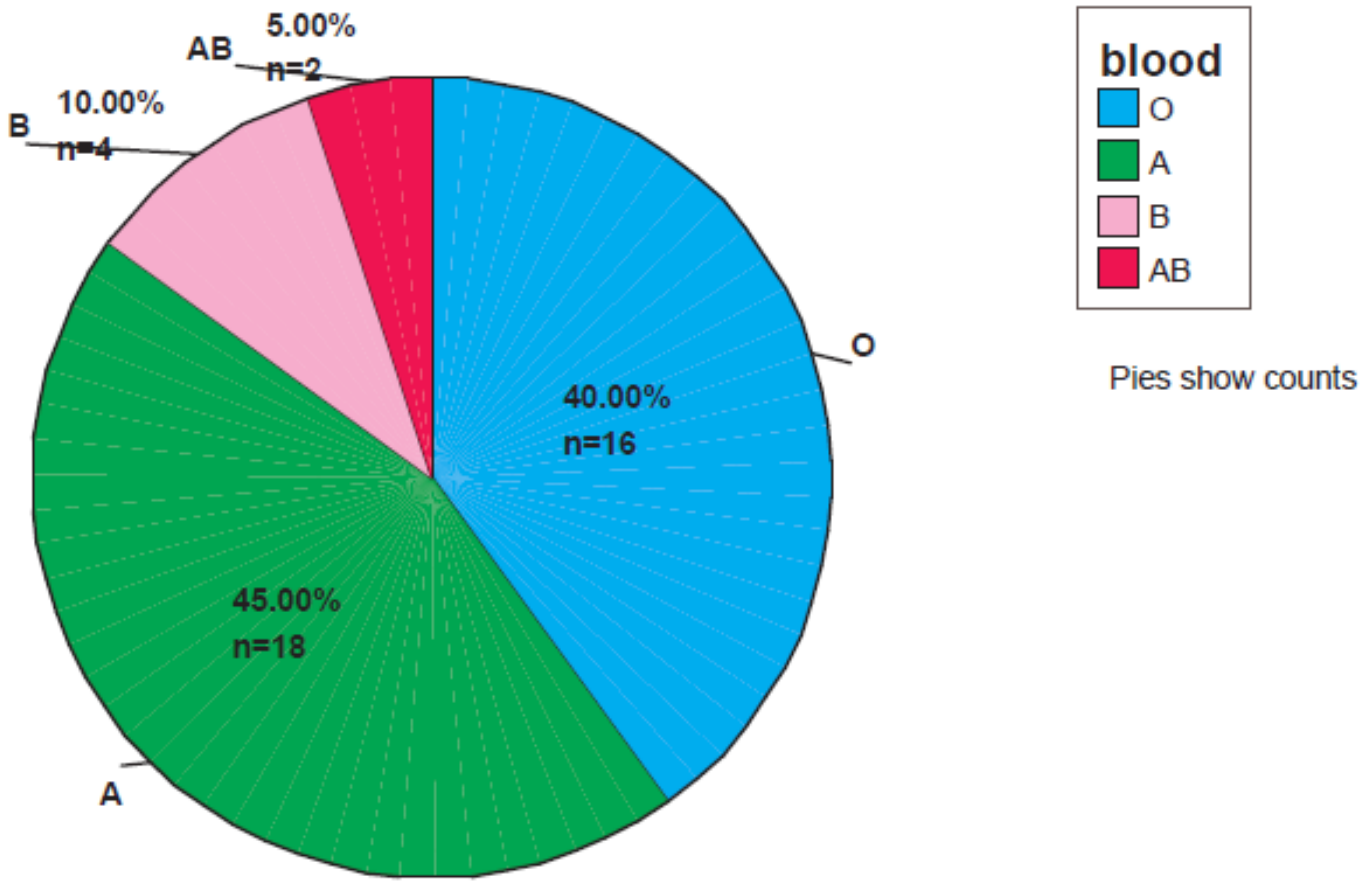
1	1	2	2	4	9	10
6	6.8	7.2	8	8.3	10	11.5

Pie Chart

- Pie charts are useful when only categorical variable is measured and shows the percentage of each categorical variable with respect to the whole or total frequency.
- A pie chart is a circular graphical representation of categorical variables in segments.
- Example: Let the blood types of 40 persons be displayed in the frequency table as follows. Use the information to draw a pie chart.

Blood Type	Frequency	Percentage
O	16	40
A	18	45
B	4	10
AB	2	5
Total	40	100

Example cont'd

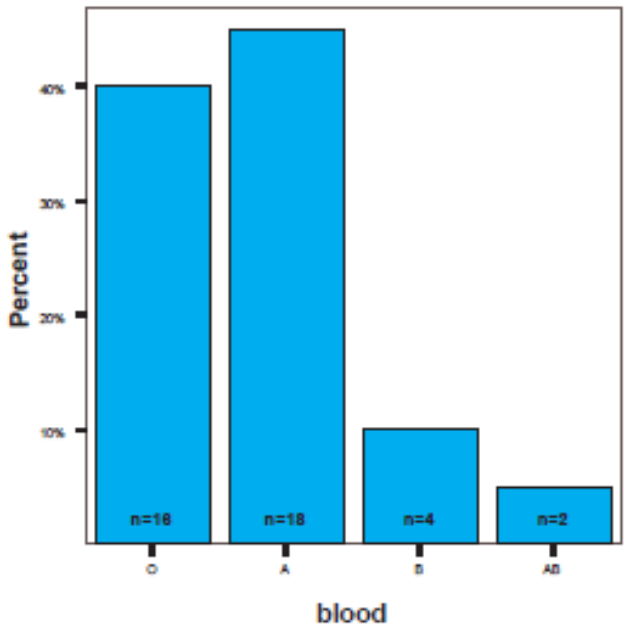


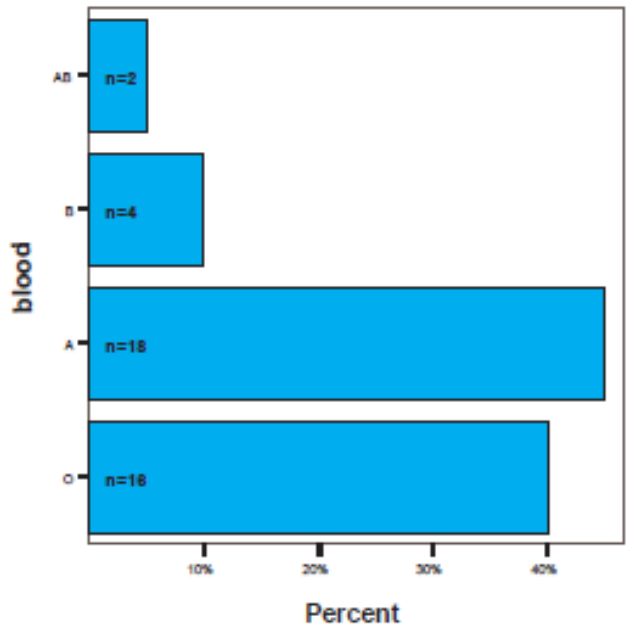
Bar Charts

- A bar charts is a graphical representation of nominal or categorical variables in bars.
- It can be also used to represent two or three categorical variables simultaneously.
- The horizontal axis is specified by the categorical variables and vertical axis represents the frequencies or percentages.
- The bars are separated to show the fact that each class or variable is a separate category.
- Types of Bar charts
  1. Simple bar chart
  2. Multiple bar chart
  3. Component or Stack bar charts

Example of Bar chart

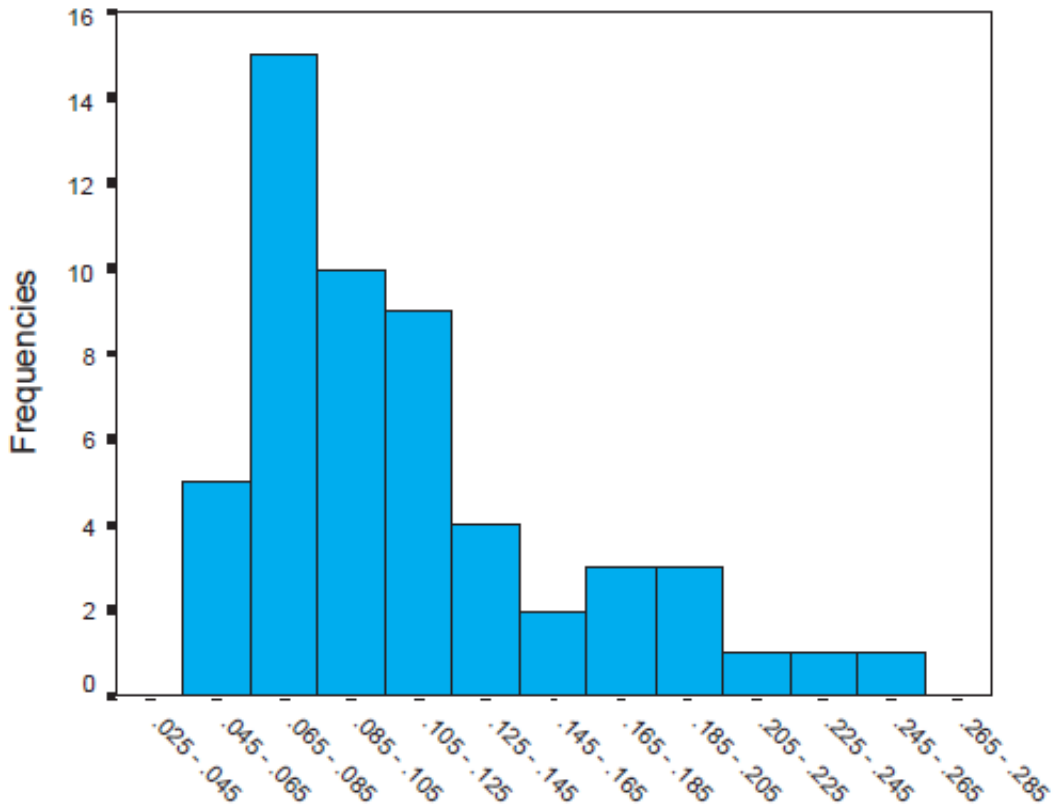
- using the blood type data, a simple bar chart of the data is displayed below:





Histogram

- A histogram consists of a set of rectangles having, bases on a horizontal axis, with centers at the class marks and lengths equal to the class interval size and areas of the bars proportional to the class frequencies.

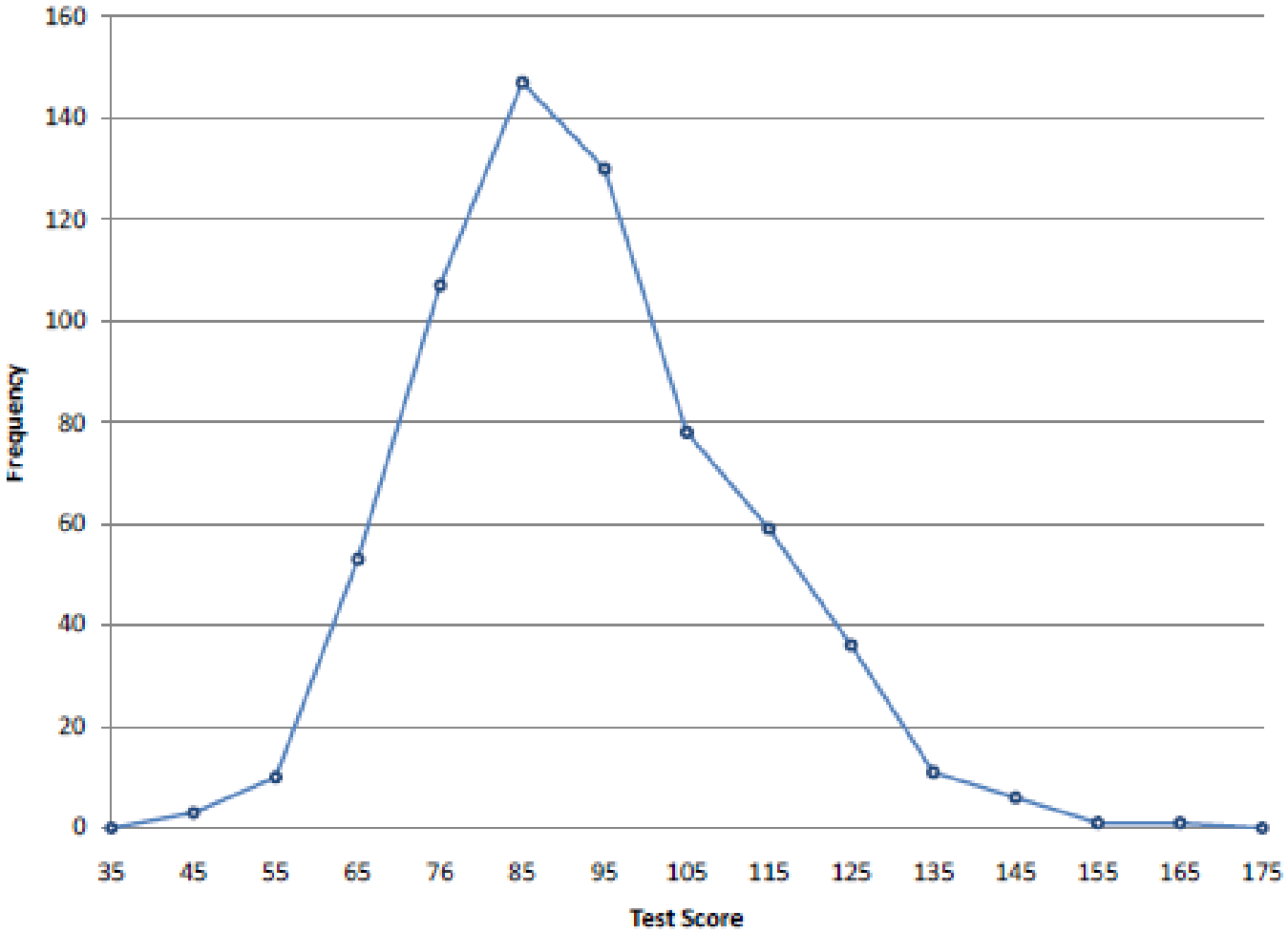


Example

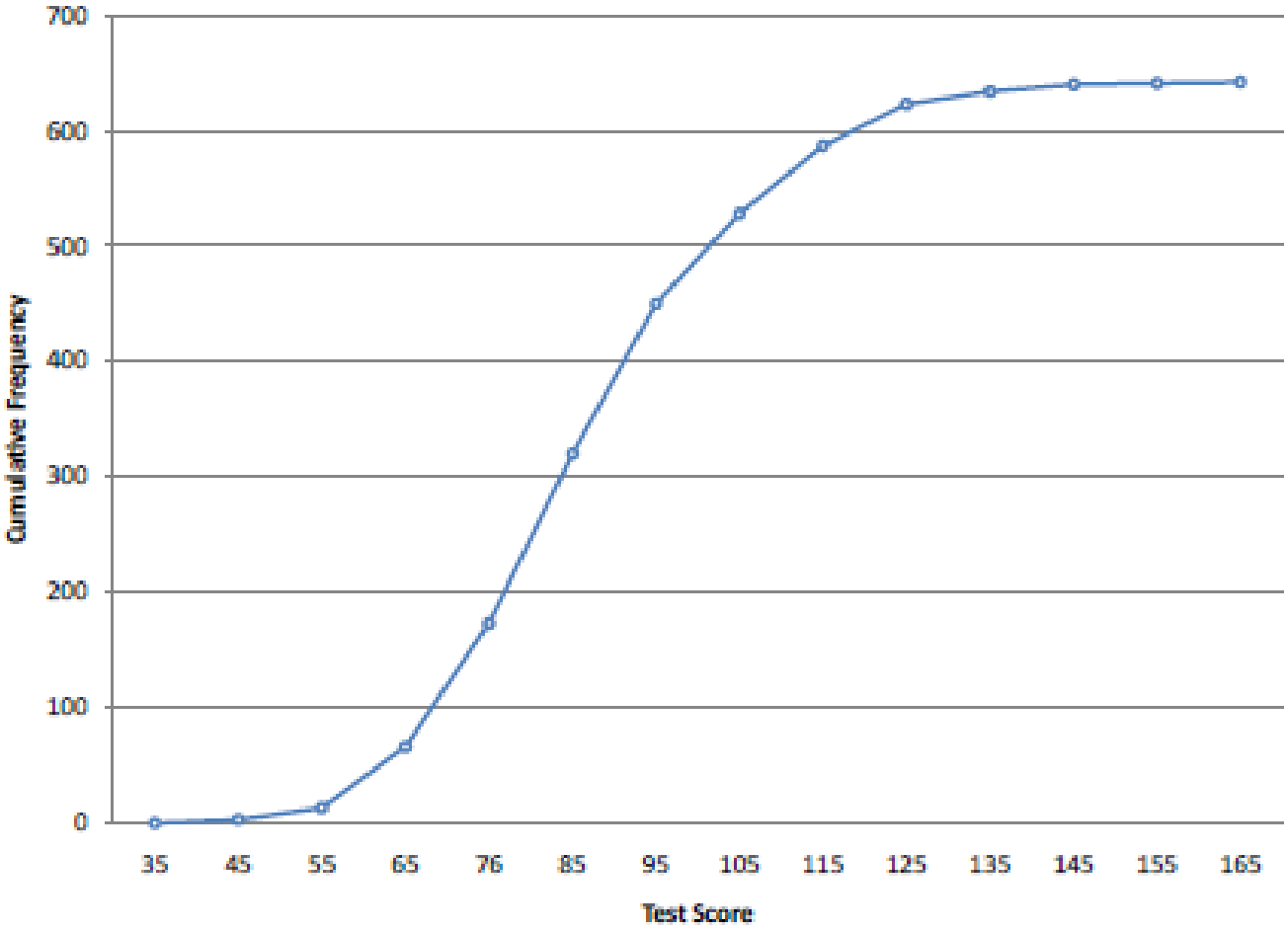
- The table below shows a frequency distribution of heights of 100 male students at the University of Ghana.
- Use the information to construct a histogram for the students.

Heights (inches)	No. of Students
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8
Total	100

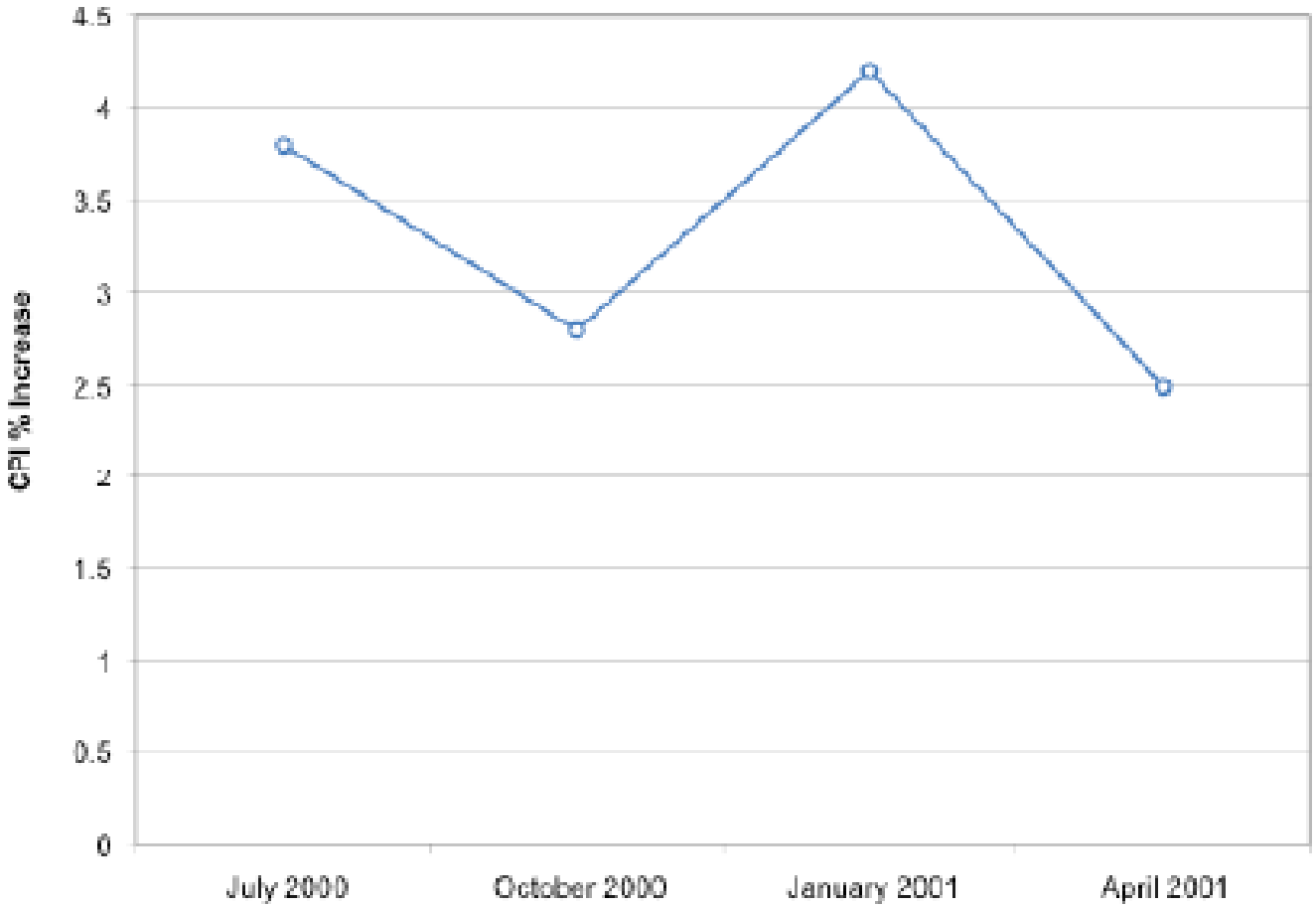
Frequency polygon for the psychology test scores



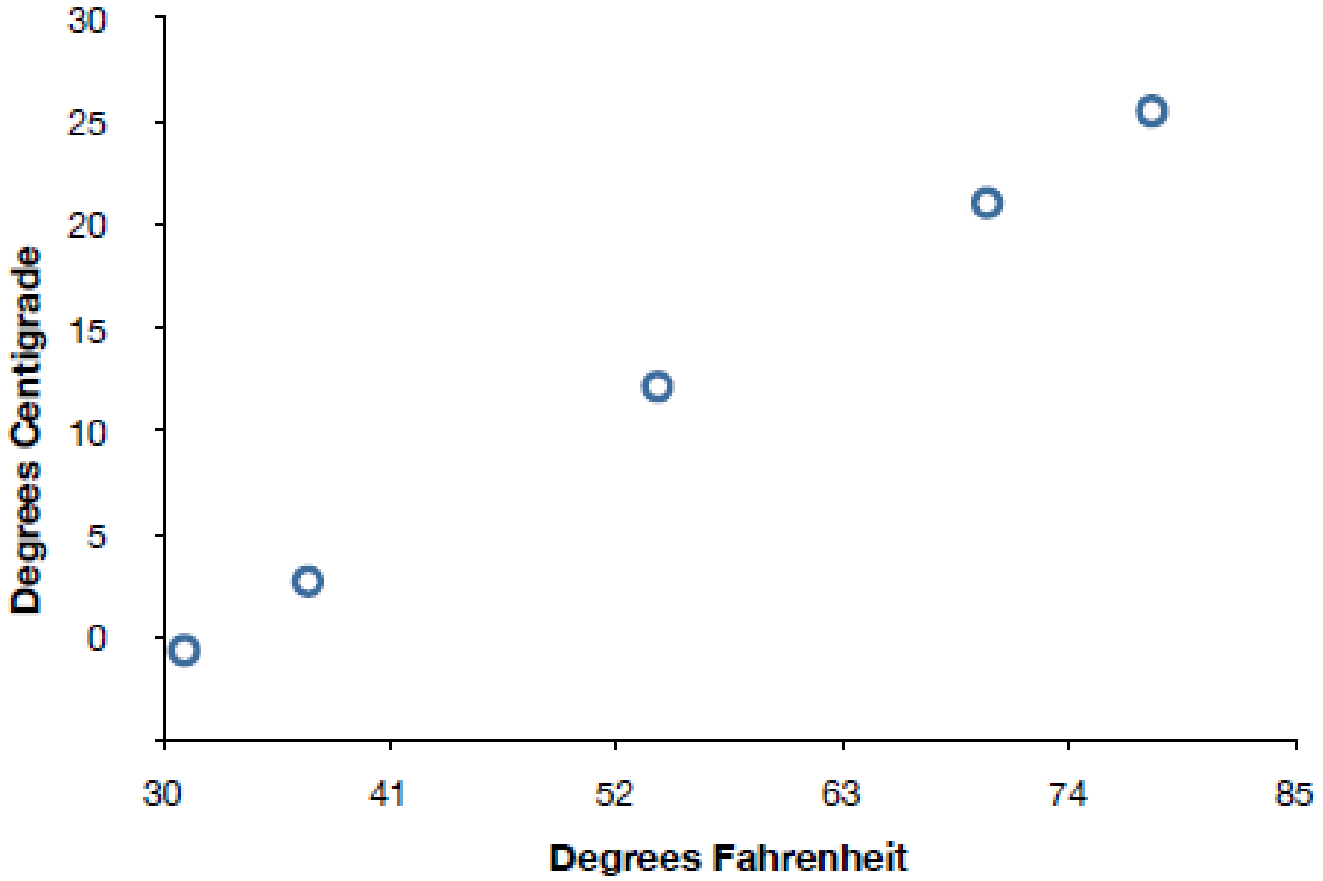
Cumulative frequency polygon for the psychology test scores



A line graph of the percent change in the CPI over time



A Scatter plot of Degrees Centigrade against degrees Fahrenheit



Recap of Data reduction

- **Discrete data**: values of discrete data are restricted to only distinct numbers (but not measurement). for eg. 0,1,2,... but not -3,2.5,etc. it is usually found by counting
- **Continuous data**: values of continuous data can be any real number. continuous data is usually found by measuring (not counting) and sometimes rounded to a suitable degree of accuracy. eg. height, weight etc.
- A frequency distribution is a tabular display of data summarized into relatively small number of intervals
- it helps in the analysis of large amount of statistical data
- They work with all types of measurement scales

UNIVERSITY OF GHANA

Construction of frequency table

1. Sort the data in ascending order.
2. Calculate the range of the data, defined as Range = Maximum value − Minimum value.
3. Decide on the number of intervals in the frequency distribution, k.
4. Determine interval width as Range/k.
5. Determine the intervals by successively adding the interval width to the minimum value, to determine the ending points of intervals, stopping after reaching an interval that includes the maximum value.
6. Count the number of observations falling in each interval.
7. Construct a table of the intervals listed from smallest to largest that shows the number of observations falling in each interval.

In Step 4, when rounding the interval width, round up rather than down, to ensure that the final interval includes the maximum value of the data.

Construction of frequency table

As the above procedure makes clear, a frequency distribution groups data into a set of intervals.[7] An **interval** is a set of values within which an observation falls. Each observation falls into only one interval, and the total number of intervals covers all the values represented in the data. The actual number of observations in a given interval is called the **absolute frequency**, or simply the frequency. The frequency distribution is the list of intervals together with the corresponding measures of frequency.

To illustrate the basic procedure, suppose we have 12 observations sorted in ascending order: −4.57, −4.04, −1.64, 0.28, 1.34, 2.35, 2.38, 4.28, 4.42, 4.68, 7.16, and 11.43. The minimum observation is −4.57 and the maximum observation is +11.43, so the range is +11.43 − (−4.57) = 16. If we set $k = 4$, the interval width is 16/4 = 4. Table 1 shows the repeated addition of the interval width of 4 to determine the end-points for the intervals (Step 5).

Construction of frequency table

Table 1 Endpoints of Intervals

−4.57	+	4.00	=	−0.57
−0.57	+	4.00	=	3.43
3.43	+	4.00	=	7.43
7.4	+	4.00	=	11.43

Thus the intervals are [−4.57 to −0.57), [−0.57 to 3.43), [3.43 to 7.43), and [7.43 to 11.43].[8] Table 2 summarizes Steps 5 through 7.

Construction of frequency table

Table 2 Frequency Distribution

Interval				Absolute Frequency
A	−4.57	≤ observation <	−0.57	3
B	−0.57	≤ observation <	3.43	4
C	3.43	≤ observation <	7.43	4
D	7.43	≤ observation ≤	11.43	1

Note that the intervals do not overlap, so each observation can be placed uniquely into one interval.

7 Intervals are also sometimes called classes, ranges, or bins.

8 The notation [−4.57 to −0.57) means −4.57 ≤ observation < −0.57. In this context, a square bracket indicates that the endpoint is included in the interval.

Relative frequency

- **Definition of Relative Frequency.** The **relative frequency** is the absolute frequency of each interval divided by the total number of observations.

The **cumulative relative frequency** cumulates (adds up) the relative frequencies as we move from the first to the last interval. It tells us the fraction of observations that are less than the upper limit of each interval. Examining the frequency distribution given in Table 3, we see that the first return interval, −44 percent to −42 percent, has one observation; its relative frequency is 1/87 or 1.15 percent. The cumulative frequency for this interval is 1 because only one observation is less than −42 percent. The cumulative relative frequency is thus 1/87 or 1.15 percent. The next return interval has zero observations; therefore, its cumulative frequency is 0 plus 1 and its cumulative relative frequency is 1.15 percent (the cumulative relative frequency from the previous interval). We can find the other cumulative frequencies by adding the (absolute) frequency to the previous cumulative frequency. The cumulative frequency, then, tells us the number of observations that are less than the upper limit of each return interval.

Stem and Leaf Plot

DATA									
15	16	21	23	23	26	26	30	32	41
Stem	**Leaf**								
1	5	6							
2	1	3	3	6	6				
3	0	2							
4	1								

- the basic idea behind a stem and leaf plot is to divide each data point into a stem and a leaf.
- used to display quantitative data

Other summary charts

- Histogram : graphical equivalent of a frequency distribution: It is a bar chart that has been grouped into a frequency distribution
- Cumulative frequency distribution and frequency polygon

Exercises

1. Make a stem and leaf plot for the numerical values below and find how many members are in the a) 10s b)20s c)30s d)40s

26	19	21	13	28	20	17	26	23	28
22	17	32	41	35	12	30	25	22	32
34	36	24	27	33	13	18	21	26	39

2. The following represents the number of marks scored by some pupils in a school in an examination

66	63	67	46	62	42	32	46
59	60	47	38	58	64	66	49
65	60	59	58	52	50	68	66
43	66	53	61	66	56	62	50
43	60	53	61	66	56	62	50
62	59	62	46	58	57	40	44

Represent the information on a stem and leaf plot and use to answer the questions below.

a) What is the total number of pupils in the class?
b) What is the highest mark in the examination?
c) What is the lowest mark in the examination?

Exercises

4. The following table gives the frequency distribution of the marks obtained in a class test by a group of 64 students.

Marks	Frequency
2	9
3	14
4	13
5	10
6	5
7	8
8	2
9	3

a) Draw a bar chart for the distribution

b) A pupil is chosen at random from the class, what is the probability that

   a. The pupil obtained 7 marks.

   b. Obtained 3 marks?

   c. Obtained more than 6 marks?

   d. Obtained less than 5 marks?

5. The salaries of workers in a company are as follows:

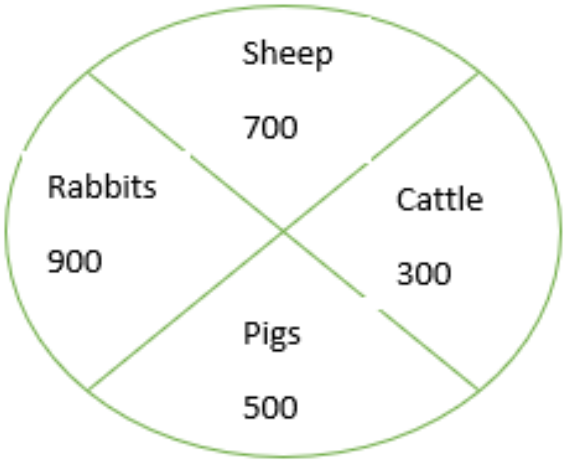
Worker	Salary (GHc)
Manager	360
Accountant	300
Supervisor	240
Secretary	90
Receptionist	90

a) Draw a pie chart to illustrate the information

b) What percentage of the salary represents that of an Accountant?

Exercises

6. the following pie chart shows the number of rabbits, sheep, cattle and pigs on a farm.



a) How many animals are on the farm?
b) What angle represents the number of sheep on the farm?
c) What percentage of the total number of animals are rabbits?
d) Calculate the angle that represents number of pigs and cattle

UNIVERSITY OF GHANA