CHAPTER – 14
STATISTICS

14.1 INTRODUCTION

The branch of science known as Statistics has been used in India from ancient times. Statistics deals with collection of numerical facts i.e., data, their classification & tabulation and their interpretation. In statistics we shall try to study, in detail about collection, classification and tabulation of such data.

(a) Importance of Data :

Expressing facts with the helps of data is of great importance in our day-today life. For example, instead of saying that India has a large population it is more appropriate to say that the population of India, based on the census of 2000 is more than one billion.

(b) Collection of Data :

On the basis of methods of collection, data can be divided into two categories :

(i) Primary data : Data which are collected for the first time by the statistical investigator or with help of his workers is called primary data. As example if an investigator wants to study the condition of the workers working in a factory then fro this he collects some data like their monthly income, expenditure, number or brother, sisters, etc.

(ii) Secondary data : These are the data already collected by a person or a society and these may be in published or unpublished form. These data should be carefully used. These are generally obtained from the following two sources.

(A) Published sources

(B) unpublished sources

(c) Classification of Data :

When the data is complied in the same form and order in which it is collected, it is known as Raw Data, It is also **Crude Data.** For example, the marks obtained by 20 students of class X in English out of 10 marks are as follows :

7, 4, 9, 5, 8, 9, 6. 7, 9, 2,

0 3, 7, 6, 2, 1, 9, 8, 3, 8,

(i) Geographical basis : Here, the data is classified on the basis of place or region. For example the production of food grains of different state is shown in the following table :

S.No.	State	Production (in Tons)
1	Andhdra Pradesh	9690
2	Bihar	8074
3	Haryana	10065
4	Pubjab	17065
5	Uttar Pradesh	28095

(ii) Chronological classification : If data's classification is based on hour, day, week and month or year, then it is called chronological classification, For example, the population of India in different year is shown in following table :

S.No	Year	Production (in Crores)
1	1951	46.1
2	1961	53.9
3	1971	61.8
4	1981	68.5
5	1991	88.4
6	2001	100.01

(iii) Qualitative basis : When the data are classified into different groups on the basis of their descriptive qualities and properties, such a classification is known as descriptive or qualitative classification. Since the attributes can not be measured directly, they are counted on the basis of presence or absence of qualities. For example intelligence, literacy, unemployment, honesty etc. The following table shows classification on the basis of sex and employment.

Table Population (in lacs)

Gender →	**Male**	**Female**
Position of Employment ↓		
Employed	**16.2**	**13.7**
Unemployed	**26.4**	**24.8**
Total	**42.6**	**38.5**

(iv) Quantitative basis : if facts are such that they can be measured physically e.g. marks obtained height, weight, age, income, expenditure etc. Such facts are known as variable values. If such facts are kept into classes then it is called classification according to quantitative or class intervals.

Marks obtained	10-20	20-30	30-40	40-50
No. of students	7	9	15	6