



DATA COLLECTION, PROCESSING AND ANALYSIS

In the previous lesson we have learnt about the concept and approaches of local area planning for which data is a necessary condition. In this chapter we will discuss about the procedures followed in data collection processing and analysis. In our routine life we come across several information through print, audio and visual media, social gatherings and discussions. But have you ever thought how data for these information is collected, processed and analysed? The collection of data refers to a plan for gathering data, information from field situations. A set of procedure is followed to get the desired data/information from the field work in geography, to process and analyse the facts in a logical and scientific manner.



OBJECTIVES

After studying this lesson, you will be able to :

- identify the steps and issues involved in data collection;
- describe various tools and techniques of data collection;
- formulate questionnaire, schedule, rating scales etc.;
- draw sketch maps of the area to be surveyed;
- select the samples and collect primary data/information;
- collect secondary data;
- make simple tables and diagrams from the collected data;
- analyse tables, maps, diagrams, photographs and charts, and generalize the results and make suggestions.



31.1 STEPS IN DATA COLLECTION

Broadly speaking there are three major steps in data collection viz.

1. One can ask people questions related to the problem being investigated.
2. One can make observations related to places, people and organizations their products or outcomes.
3. One can utilize existing records or data already gathered by others for the purpose.

The first two steps relate to the collection of primary data while the third step relates to the collection of secondary data. The information/data collected by a person directly is known as primary data while records or data collected from offices/institutions is known as secondary data.

A. Steps in Primary Data Collection:

Collection of primary data involves the following steps :

1. Making oneself ready both mentally as well as physically for collecting primary data from field situations.
2. Keeping a field book/record book or diary for writing relevant information, doing field sketching or writing records of the occurrence of phenomenon at specific time intervals.
3. Administering questionnaire schedule to the target groups of area people across sampled sites.
4. Verifying the facts through cross checks in the answers and ground realities.
5. Integrating the observations, responses and recorded facts in a systematic and logical framework.

B. Steps in Secondary Data Collection:

The collection of secondary data involves the following steps:

1. Knowledge about the offices/institutes etc. keeping the record of relevant data is of prime importance to obtain the secondary data/information.
2. Get an official letter containing your requirements of data and purpose of data collection from your Principal/Head of the Institute? Your identity card is also an essential requirement to get an entry in the offices.
3. Keep a note book/record file to transfer data for the purpose. It could also be done with the help of photo copying systems.
4. The secondary data, thus, collected forms the basis for tabulation and processing as per need.



Notes

C. Identification of Issues:

It is very important to identify clearly the issues that are going to be assessed.

Depending upon the availability of time, cost, manpower and tools, a framework of issues to be covered need to be developed. In case of local area planning the following issues need to be considered.

1. Issues related to environmental conditions like environmental degradation, quality of human life etc.
2. Social issues like people's perception, literacy status, health hazards, incidence of crime etc.
3. Economic issues like employment, expenditure pattern, flow of goods and commodities etc.
4. Population study for agriculture, industry etc.
5. Landuse study for agriculture, industry etc.
6. Facilities and amenities available for social and economic development.
7. Problems related to growth of economy such as irrigation, means of transportation, availability of power etc.
8. Focal theme of planning like provision of basic amenities in slum areas, pollution control, clean environment in an industrial area.

31.2 TOOLS AND TECHNIQUES OF DATA COLLECTION

For data collection we make use of certain tools and follow specific techniques. The tools that help in data collection are as under:

- Observing the phenomenon and recording the details,
- Inquiring about the facts through questionnaires/schedules
- Making measurements.
- Conducting tests.
- Recording the events.

Now let us study some of these tools and techniques of data collection.

A. Questionnaires:

The questionnaires or interview schedules are the set of questions framed for the specific purpose of data collection through field work. The questionnaire serves two purposes. First, it translates the objectives of the field work into specific questions which help in the collection of necessary data. The data collected through the responses of the questions forms the basis of understanding the problem or explore the idea set by the objective. In order



Notes

to achieve these objectives, each question must communicate to the respondent the idea or group of ideas required by the objective and obtain a response which can be analysed to fulfill the objectives. The question must perform these functions with minimum distortion of the response it deals. In asking a question to the respondents, we assume that he possesses adequate knowledge, opinion or attitude. Each question should, therefore, be constructed so as to elicit a response which accurately and completely reflects each respondent's position.

The second purpose of questionnaire is to assist interviewer in motivating the respondent to communicate the required information. There are many factors which determine the respondent's willingness to engage in an interview. The questionnaire itself does much to determine the nature of interviewer-respondent relationship. Thus, the quantity and quality of data collected depends largely on the nature of questionnaire.

(a) Contents of Questionnaire:

The following two types of information should from the contents of questionnaire:

- (i) Identity or location specific contents
- (ii) Respondent centred contents

(b) Form of Questionnaire:

The form of questionnaire depends upon some of the factors such as willingness of the respondents, usefulness of the information and its level, language, sequence of questions, single idea etc.

(c) The Interview

The process of conducting interviews starts soon after the formulation of questionnaire is complete. The investigator should have a letter of introduction to explain about himself in the field. The letter of introduction must have a note that the information so collected is going to be used for the purposes of presentations and educational use only. The information will remain anonymous completely. While conducting interviews, we should help in removing the difficulties of the respondents without giving any clue as to the answer required. As far as possible we are not supposed to make any responses or show any expressions to the answers. Finally we should pay regards and express thanks to the respondents for their co-operation.

B. The Schedules

The schedules are the timed plan for a survey. It reflects time specific recording of the phenomena like traffic survey, consumer behaviour survey, precipitation



Notes

pattern etc. The investigator must record the occurrence of a phenomenon over a specific time interval. The time is an important reference of analysis. It could be in convenient units of hours, minutes or seconds depending upon the frequency of occurrences. Similarly, a phenomenon is more often associated with several elements. Hence, the record book need to have further sub divisions both on X as well as on Y axis.

1. What phenomenon to be selected and recorded in order to obtain the required information?
2. Under what conditions are observations to be made? How is the observational situation structured?
3. Can a score be assigned to the observation and what are the characteristics of that score?
4. How stable are the observations? Can the same results be obtained under the same conditions?
5. Whether the phenomenon observed has functional unity with same process?

C. Rating Scales

By the term rating scale, we mean a scale with a set of points, which describe varying degrees of dimension being observed. Rating scales are most often used in either of two ways, 1) to record the pattern at frequent intervals, or 2) to rate the entire event after it has ended. Thus, rating scales, which contain a variety of items at each point on the scale, are more efficient since they can provide more data per observer, more dimensions per unit of area and time. Investigator observes a number of acts throughout the situation, integrates them in his mind, and makes a judgment as to which point on a number of scales best described his interpretation of the varied behavior. The following examples offer an idea of rating scales.

Temperature Conditions:

Very Cold	Cold	Cool	Moderately Warm	Hot	Very Hot
0	1	2	3	4	5

Development Level:

Under Developed	VeryLow Level	Low Level	Medium Level	High Level	Very High Level
0	1	2	3	4	5

D. Field Sketches

Making of field sketches on the spot is an essential component of field survey



in geography. These are simple, rough drawings or design done rapidly to depict the ground truth on a piece of paper. Geographical facts like structure or form of physical landscape, location and site, mobility, intensity of interactions, patterns of level use, distance and directions and interdependence of certain natural or cultural objects can be depicted symbolically in the form of field sketches.

E. Photographs

Camera is one of the important equipments that is needed during the course of a field work and data collection. It is needed for taking photographs of typical features. Photographs present the view of a landscape in its totality, activity in operation and events in their occurrences. Photographs provide comprehensive data bases for analysis and interpretation. Certain aspects that need more time to record such as conditions in a slum locality, variety of landscapes, plant species, office and factory systems can be photographed and the output can be used for the explanations and analysis. Photographs are used to supplement the results.

- The formulations of questionnaires serve two purposes: (i) first translate the objectives of the field-work into specific questions which help in collection of data and (ii) the second purpose is to assist the interviewer in motivating the respondents to communicate the required information.
- Various factors which affect the form of questionnaires are (i) willingness of the respondent, (ii) the frame of reference, (iii) usefulness of the information, (iv) possibility of misunderstanding, (v) type of questions, (vi) the information level (viii) social acceptance (viii) single idea and (ix) sequence of question.
- Various precautions need to be observed while administering the questionnaire. These precautions are (i) The collection of information need to be done in an atmosphere of permissiveness, (ii) the respondent should not be kept in dark about the purpose, (iii) explain the anonymous or confidential nature of interview, (iv) socially unacceptable questions need to be avoided, and (v) the intention of the interview need to be given convincing explanations.

F. Methods of Administering the Questionnaires and Survey Schedules

The questionnaires are the set of questions framed for specific purpose of field work. Before designing the questions the purpose of specific problem is divided into various steps and phases. After this logical sequence of questions is to be developed so that desired response can be obtained. The coding of questions (each question to be given a numerical code) is another important dimension required for the transfer of data/information to



Notes

computers. The whole questionnaire is divided into schedules sets like household schedule, amenities and facilities schedule, function or activity schedule. Thus the questionnaire is a set of schedules having purpose specific questions. Schedule of time is another dimension worked out to complete the field work in given period.

Normally, the administration of questionnaire will follow a sequence of procedures in the manner given below:

1. **Building Rapport:** It refers to the atmosphere of entire relationship between respondent and interviewer. It would be necessary for him to establish a deeper kind of personal relationship with the respondent.
2. **Asking the Questions :** The interviewer's job of asking questions from the questionnaire is through the use of carefully worded questions transmitted to the respondent in verbatim which will help in achieving most of the standardization in the interview. The major aims of putting questions to a variety of respondents is to have complete and clear response about the point of investigation.'
3. **Use of the field sketches and sketch maps:** The field sketches are additional supports to the questionnaires in the collection of primary data. Field sketches supplement the set of information by producing a rough image of physical as well as cultural landscapes. These are the free hand pencil or pen drawn images on the field diary. These sketches help remembering and recollection of field relations. They also substantiate the facts as a visual presentations.

G. Collection of Information

Both the tools of registration and recording help us in the collection of primary data. With the help of these tools, we try to transfer the facts from field into data and tables. In this process of collection, there is obviously the loss of some information. Nevertheless, a good deal of satisfactory information is collected and utilized for the purpose of analysis and interpretation. Based on the set of questionnaires, schedule administered to the respondent, the desired information/data is collected. The collection of information could be a routine as well as specific purpose exercise. The routine data collection could relate to daily sales, commuting population, movements of goods etc. Similarly, recording of weather elements like temperature, air pressure, precipitation, direction of winds, cloud cover, sea conditions etc. is a routine data collection. There are many other examples of daily data collection. Based on the daily information or facts, seasonal trends and annual averages are worked out. The purpose specific data is collected at one point of time only.

H. Precautions in Collecting the Information

The task of collecting the needed genuine information is difficult one. The



Notes

collection of data from field situations is a complicated affair compared to the office or organizational situation. To get an unambiguous, unbiased and correct information from field, specific precautions need to be observed. These are related to the non-cooperation, incorrect information and tensions. The following precautions need to be observed to overcome these difficulties:

- (i) The collection of information need to be done in a friendly way. The interviewer is supposed to remain humble, polite and establish good rapport with the respondent.
- (ii) The use of words and sentences should not sound unfamiliar and causing hurt to the sentiments of the respondents. Such words and sentences need to be replaced by more appropriate words.
- (iii) Socially unacceptable questions need to be avoided. If so required, indirect information be used for the purpose.
- (iv) The respondents should not be kept in dark about the purpose of the field work. The respondent may not like to answer the questions if he is not clearly explained about the objective of the fieldwork and more specifically about his selection as sample for the data collection.
- (v) The respondent need to be assured of his/her identity and response to remain undisclosed (anonymous) and his/her cooperation to be duly acknowledged in the work.
- (vi) The intentions of the interview need to be given convincing explanations. The information collected is in no way going to affect the respondent adversely i.e., to impose a check upon his activities.

I. Selection of Samples and Sample Size

A sample is a part of a larger group or area selected for obtaining information about the whole group or area known as the universe of the study. The part of the whole is called sample and is used to ascertain the characteristics of the universe of the study. While choosing a sample, the population is assumed to be composed of individual area units or members of the group. Some of these units or members of the population selected for detailed study are called the samples. When the entire universe is taken into consideration for the study, it is known as census survey. Examples are population census, agricultural census and so on.

1. **Identification of Samples:** The identification of samples is the first task while conducting the field survey. The selection of sample should be such that it reflects the characteristics of the whole. The sample should not be identical as it leads to error.
2. **Sampling Techniques :** Samples are selected to avoid unnecessary large expenditures likely to be incurred on the total survey of all the units of universe of study. Moreover, a sample study can be completed in a lesser



Notes

time period compared to the study of universe or population. The level of accuracy also increases when we study a smaller area units and vice versa in case of the universe. The measures of assessments, estimates and projections can be better used for the purpose of planning, execution and diffusion studies. Some of the popular sampling techniques are discussed here.

- (a) **Systematic Sampling :** The items selected from the population are chosen in a regular way. Such a procedure of sampling is called a systematic sampling. For example selection of samples in a multiple of 8(8th 16th, 24th etc.), 10 (10th, 20th, 30th etc.) or any other number so decided.
 - (b) **Random Sampling:** The selection of samples, in random sampling, depends upon the chance as universe presents homogenous conditions throughout. There are two types of random sampling.
 - (i) **Simple Random Sampling:** The procedure of sampling in which each unit of universe has equal chance of being included as the sample is known as simple random sampling. For example in a survey on consumer behavior each consumer has an equal chance for being selected as a sample.
 - (ii) **Stratified Random Sampling :** This type of sampling procedure is used when considerable heterogeneity is present in the distribution. The selection of samples in such a situation is based on the division of the universe of study into homogeneous subgroups or strata. Certain aspects of study present stratified character like social structure (having groups like general population. SC population and ST population); economic structure (primary, secondary, tertiary sector etc.) Random samples are selected from each sub group based on their relative significance in the universe.
3. **Sample Size:** There are two basic requirements for the sample to fulfill. A sample must be representative and adequate. The sample is said to be representative when it reflects the various patterns and sub classes of the universe of the study. Similarly, a sample is adequate if it provides very precise result to the investigator. It is important to note that larger is the sample size, greater is the accuracy.

Usually a small sample is sufficient if the phenomenon studied is fairly homogeneous which very rarely occurs. Normally, for a field survey sample size chosen is about 5 to 10 percent of the total units of the universe.

- The sum total or aggregate from which the sample is taken and the result is derived is known as the universe or population.



Notes

- A sample is a part of a group or aggregate selected for the purpose of obtaining information about the universe.
- The procedure dealing with the selection of a part of a group from the universe to obtain information about the whole or the universe is known as sampling.
- A scheme for obtaining a suitable sample from a given universe is known as sampling design. It also indicates the size of the sample to be used keeping in view the cost involved and the precision of the result required.
- A procedure of sample selection in which units are selected at equal interval is known as simple random sampling.
- Stratified random sampling is a method of sample selection in which the universe of the study is divided in to homogeneous subgroups and simple random sample is selected from each subgroup.



INTEXT QUESTIONS 31.1

1. Give a single term to the following statements:
 - a. The data collected by asking questions from people or making observations related to the problem of investigation is known as.
 - b. The data available in records or already gathered by others for the purpose is called.
 - c. The material medium that help in data collection are called as.
 - d. Methods or ways through which data is collected are known as.
 - e. A set of questions framed for the purpose of data collection through field work is called.
2. State three important steps each for primary and secondary data collection.

A. Primary data collection	1.
	2.
	3.
B. Secondary data collection	1.
	2.
	3.
3. Match the two columns

Column A

(a) Interview

Column B

(1) A type of scale with a set of points, which describe varying degrees of dimensions.



Notes

- | | |
|------------------|--|
| (b) Schedules | (2) A rough drawing or design to present the item. |
| (c) Rating Scale | (3) A timed plan for finalizing a project or field work. |
| (d) Sketches | (4) The process of interacting with the target group in the form of questions. |
4. State the two types of information that form the contents of the questionnaire.
- | | |
|-----|-----|
| (a) | (b) |
|-----|-----|
5. Name two criteria which are necessary for the identification of a sample.
- | | |
|-----|------|
| (i) | (ii) |
|-----|------|

31.3 PROCESSING OF DATA

The processing of data/information is an essential dimension of stream lining the facts and writing of a field report. A separate account of processing is given here.

(A) Processing of primary data: The primary data collected from the field remains in the raw form of statements, digits and qualitative terms. The raw data contains error, omissions and inconsistencies. It requires corrections after careful scrutinizing the completed questionnaires. The following steps are involved in the processing of primary data.

(i) Editing of data: The editing of data can be done at two stages: field and post-field editing. The field editing is a review of reporting by the investigator for completing what has been written in an abbreviated form during interviewing the respondent.

The post-field editing is carried out when field survey is completed and all the forms of schedule have been collected together. This type of editing requires review of all forms thoroughly.

(ii) The coding of data: To keep the response with in limited alternatives, we need to assign some alphabetical or numerical symbols or both to the answers. The alternatives must be mutually exclusive i.e. defined in one concept or term only. This form of processing is known as coding. For example in a question of educational qualifications alternative choices given are: Uneducated; Below Matriculation; Matriculation & above but below Graduate; Graduate & above; Technical Diploma; Technical Degree



Notes

The alphabetical codes assigned to these alternatives could be A,B,C,D,E and F. Similarly, numerical codes to these alternatives could be 1,2,3,4.,and 5 respectively. It is necessary for the efficient analysis. Though coding exercise is a part of the formulation of questionnaire yet responses to questions need to be coded and made final at the processing stage. This simplifies the transfer of data from questionnaires to the master chart. It is a two dimensional chart in which observations are entered on one axis (X) and details of the responses on the other axis (Y). The calculations becomes easier and quicker if the details are coded and entered in the master chart or fed in the computers.

(iii) Organization of Data: The data information collected through different sources should be organized. The first task in this regard is to develop a master chart. For example in a local area survey, we record individual households in rows and the details of population, function, facilities and amenities etc. in columns. Thus a large chart is prepared that contains, practically, all relevant information/data. Finally the total of rows and columns are cross-checked. The information arranged in an ascending order is known as the array of data. The set of information related to specific entity is called the field. The following illustration demonstrates the way data is organized.

Households	Details			Population				Functions				Facilities		
	P	M	F	Agri	Ind	Trade	Service	T.V.	Phone	Vehicle				
01	20	12	08	5	-	1	12	1	1	1 Scooter				
02	17	09	08	6	-	1	1	1	1	1 Scooter				
03	9	04	05	-	-	2	1	1	2	1 Car and 1 Scooter				
04	12	06	06		1		2	1	1	1 Scooter				
05	13	07	06	2	-	-	2	1	-	1 Scooter				

(iv) Classification of data: A huge volume of raw data collected through field survey needs to be grouped for similar details of individual responses. The process of organizing data into groups and classes on the basis of certain characteristics is known as the classification of data. Classification helps in making comparisons among the categories of observations. It can be either according to numerical characteristics or according to attributes. The numerical characteristics are classified on the basis of class intervals. For example monthly income up to Rs.2000 may form its group and the respondents reporting income in the range may form its frequency. Similarly, further group can also be made like income group Rs.2000 to Rs.3000 and so on. The number of items entered against each class is known as the frequency of the class. Every class has a lower and an upper limit. The difference between the upper



Notes

and lower limits is known as the range of the class. The class intervals are mostly kept equal. Sometimes when the range of the data is too large class intervals are not kept equal, instead they are based on the perceptible gaps in the array of the data. For example settlements having less than 2000 population can be grouped as below 200 population 200-500 population, 500-1000 population and so on. In this group as class intervals are unequal.

The data is also classified on the following bases.

1. Descriptive characteristics-example land holding, sex, caste and so on.
2. Time, situation and area specific characteristics.
3. Nature of data as continuous or discrete.

(B) Presentation of data: The presentation of data could be tabular, statistical and cartographic. In case of tabular form of presentation, data related to different variables should be classified and compared. Various statistical techniques are available to derive accurate and precise results. Since techniques have a large range coupled with the limitations of their own, selection of appropriate technique needs to be made for the purpose. The construction of graphs, charts, diagrams and maps are the various forms of cartographic presentations. The data is transformed into cartographic system which is used for visual presentation. A brief account of tabular, statistical as well as cartographic presentation of data is discussed below.

(i) Tabular Presentation: It is used for summarization of data in its micro form. It helps in the analysis of trends, relationship and other characteristics of a given data. Simple tabulation is used to answer question related to one characteristic of the data whereas complex tabulation is used to present several interrelated characteristics. Complex tabulation results in two way, three way tables which give information about two or three inter-related characteristic of data. The following points may be kept in mind while constructing a table.

1. To make a table easily understandable without a text, a clear and concise title be given just above the frame of the table.
2. Each table should be numbered to facilitate easy reference.
3. Both columns and rows of the table should have a short and clear caption. They may also be numbered to facilitate the reference.
4. The units of measurement (production units)- kgs, quintals, tones, or areal units-hectare, kilometre) be indicated. If table relates to some specific time, it must be mentioned. The tables should be logical, clear and as simple as possible.



Notes

5. The source of data must be indicated just below the body of the table.
6. The abbreviated words and explanatory foot notes if any should be placed beneath the table. However, it should be used to the minimum possible extent.
7. The sequence of data categories in a table may follow alphabetical, chronological, geographical order according to magnitude of the item presented.

(ii) Statistical Presentation of data: The data collected through various sources needs to be processed statistically for precise explanations. Very often it becomes necessary to obtain a single representative value for the whole data set. The statistical measures that enable us to work out a single representative figure for the entire data distribution, is known as central tendency. Measures of central tendency help us to compare different distributions besides being representative for each distribution. These measures normally denote the central points of values, distance and occurrence in a distribution. The commonly used measures of central tendency are:

- (a) Arithmetic mean or average
- (b) Median
- (c) Mode

(a) Arithmetic Mean

It is most frequently used and is calculated by adding the sum of all individual values in a distribution and dividing the sum by the total number of individuals. For example, the production of rice per acre in five districts is 10, 8, 12, 9 and 6 quintals. The average production of rice for these districts is :

$$\frac{10 + 8 + 12 + 9 + 6}{5} = \frac{45}{5} = 9 \text{ quintals per acre}$$

The arithmetic mean is expressed in the form of equation noted below:

Where \bar{X} is the mean value,
 ΣX is the total of X values,
 N= Number of individuals/observations.

The arithmetic mean can be easily worked for small ungrouped data. However, when the number of observations are large and data is in the form of frequency



Notes

distribution of groups, arithmetic mean will be worked out with the help of following equation.

Where \bar{x} is the arithmetic mean,
 f is the frequency,
 m is the mid value of the classes

Example

Calculate the arithmetic mean from the temperature data given in the following table.

Classes (Temperatures in degree Celsius)	No. of days f	Mid values m	fm
x	f	m	fm
1-05	20	3	60
06-10	24	8	192
11-15	44	13	572
16-20	72	18	1296
21-25	76	23	1748
26-30	60	28	1680
31-35	62	33	1716
36-40	4	38	152
41-45	8	43	344
	$f = 360$ days	$fx = 7760$	

From the above

temperature

Merits of the Arithmetic Mean

1. It is easy to understand the complete idea of the distribution and simple to workout.

$$\bar{x} = \frac{\sum fm}{\sum f} = \frac{7760}{360} = 21.50^{\circ}\text{C}$$



Notes

2. It is the average of the values in a distribution. Hence, it has a balancing property in case of sample surveys.
3. It is widely used in case of normal distributions.

The arithmetic mean has certain limitations. It is affected by the extreme values especially when they are large. For example, income variations are very wide in case of Indian population.

(b) Median

Median is the middle most positional average. It is worked out by arranging data in an ascending or descending order. For example, the value of the median is worked out by adding 1 to the number of observation and the sum divided by two. It is expressed as:

For example if we are interested in working out the median latitude and longitude for the country, we must arrange these distributions in a tabular form.

Latitudinal Extent of the Mainland of India (8'4' N to 37' 6' N)

9	10	11	12	13	14	15	16	17	18
19	20	21	22	23	24	25	26	27	28
29	30	31	32	33	34	35	36	37	

The median or middle most latitude of India is 23°N which is close to the Tropic of cancer (23° 30' N,). Since mainland of India starts from 8' 4' N which is a part of 9th latitude and extends up to 37° 6' N which covers the 37° latitude completely, hence the latitudinal coverage of India is approximately 29° latitudes. The median latitude is therefore, 23°N i.e.

$$\text{Med} = \frac{N+1}{2} = \frac{29+1}{2} = \frac{30}{2} = 15$$

8° + 15° = 23°N Southern tip of India)+ 15° (median value)=23° (middle east latitude of India) Similarly, we can also workout the median value for the longitudinal extent of India. The Longitudinal Extent of India ranges between = 68°7' E to 97°25'E.

The median or middle most longitude for the country is 83°E.

69	70	71	72	73	74	75	76	77	78
79	80	81	82	83	84	85	86	87	88
89	90	91	92	93	94	95	96	97	



Notes

Longitudes are used to calculate local time, standard time of a nation and international time which is linked to Greenwich Mean Time (GMT). Indian standard time is calculated keeping $82^{\circ}30'E$ longitude as the base. The median longitude for the country is $83^{\circ}E$ which is close to the standard meridian used for Indian Standard Time calculation.

Merits of Median:

1. Being the middle most value, median remains unaffected by the extreme values in the distribution as in the case of arithmetic mean.
2. It is a partition value which divides the series into two nearly equal parts and remains the centre of gravity.
3. However, it cannot be worked out without putting data in an ascending or descending order. If data are large, it might be a time consuming and tedious job. The values of median will be erratic if one or two items are added or subtracted from the series.

(c) Mode:

It is one of the important measures of central tendency. The maximum concentration of items occurring in a distribution is considered to ascertain the mode. The value which occurs most frequently is identified as mode in case of ungrouped data. Similarly, for grouped data the mode can be calculated by identifying the class with the highest frequency. The mode denotes the centrality of the occurrence of an item in the distribution. The distribution of rural settlements in Uttar Pradesh is given below. Workout the mode for the data.

Distribution of Rural settlements in Uttar Pradesh 2001

Size of Rural Settlements	Very small (Below 500 Population)	Small (500-999)	Medium (1500-1999)	Large (2000-4999)	Very Large (5000 and above)
Proportion of distribution	16.69	23.46	47.97	10.59	1.29

Solution: Arrange the data in a sequence (either from small to large or from large to small). Put up the frequency values against each. Now compare the frequencies. The distribution registering maximum frequency is identified as 'mode'.

Merits of the Mode:

1. It is the most typical value of a series. Mode can be located easily by the inspection and can be used by common people also.
2. The occurrence of a few extreme values does not affect the mode, since it is the most typical value of series.



Notes

It is, however, not a significant measure of central tendency unless the number of observations is large. Both in case of uniform as well as skewed distributions, mode ceases to be a measure of central tendency.

Percentiles:

Percentile is a measure which divides a series into 100 equal parts. It helps to understand various classes or categories that constitute a distribution. It is expressed as:

$$\text{for ungrouped series and } P_j = L_1 + \left(\frac{P_j N / 100 - C}{f} \right) \text{ for grouped series}$$

Where P is the percentile and N is the number of observations.

There are 99 percentiles, P_1, P_2, \dots, P_{99}

L_1 = The lower limit of the j^{th} percentile class, this is frequency of this class,

C = is the cumulative frequency of the class preceding the percentile class, and

h = the magnitude of the j^{th} percentile class.

f = the frequency of the percentile class.

Distribution of Monthly Income Among Households of a locality

	Actual Number	Percentage Distribution
Economically weaker sections (Below Rs.500)	112	56.0
Lower Income Group (500-999)	41	20.5
Middle Income Group (1000-4999)	29	14.5
High Income Group (5000 and above)	18	9.0
Total	200	100=00



Distribution of Per Capita Monthly Income of the Households of a locality

Income group in Rs.	No. of Households Frequency	Cumulative Frequency
Below 500	112	112
500-999	41	153
1000-4999	29	182
5000 and above	18	200
Total	200	

Notes

Let us calculate 60th percentile as P_{60} .

Now $P_{60} =$

The 120 the income lies in the group 500–999 so that,

$l = 500, f = 41, c = 112$ and $h = 500$

$$P_{60} = 500 + \left[\frac{8}{41} \right] \times 500$$

$$= 500 + 97.56$$

Ans. = Rs. 597.56

It means that 60 percent of the monthly incomes are below Rs. 597.56 and remaining 40 percent above it.

(iii) Cartographic Presentation of Data: The primary data collected through the field survey may be presented cartographically. The representation of data in visual form refers either to time, space or to both. The cartographic presentation refers to the display of data by constructing graphs, diagrams and maps. The set of data is transformed into some form of figure which is used for illustrations. These figures could be graphic, geometric or theme specific maps. A brief discussion on different form of cartographic presentation is given here.

(a) Graphical Presentation of Data: The graph refers to the arrangement of horizontal as well as vertical lines in inch or centimeter’s divisions. These divisions are in an arithmetic sequence. A graph is used to locate the position of a given characteristic with respect to two variables represented by two axes of the graph. ‘While ordinate or X axis represents independent variable,



Notes

abscissa or Y axis represents the dependent variable. Due care is needed in the construction and interpretation of graphs. Theoretically, a phenomenon could be either increasing or decreasing or keeping constant trend of change across time. However, the observed facts may represent the change in a mixed fashion. For example, we can make use of simple line graph to represent the profile of Indian population during the past ten decades i.e. 1901 to 2001. Although we can see changes in the data, the presentation of the same on a line graph provides better comprehension.

Table 31.1 : Growth of Population in India 1901-2001

(Population in million persons)

Year	1901	1911	1921	1931	1941	1951	1961	1971	1981	1991	2001
Pop.	1238.3	1252.0	251.3	278.9	318.6	361.0	439.2	548.1	685.1	846.3	1028.73

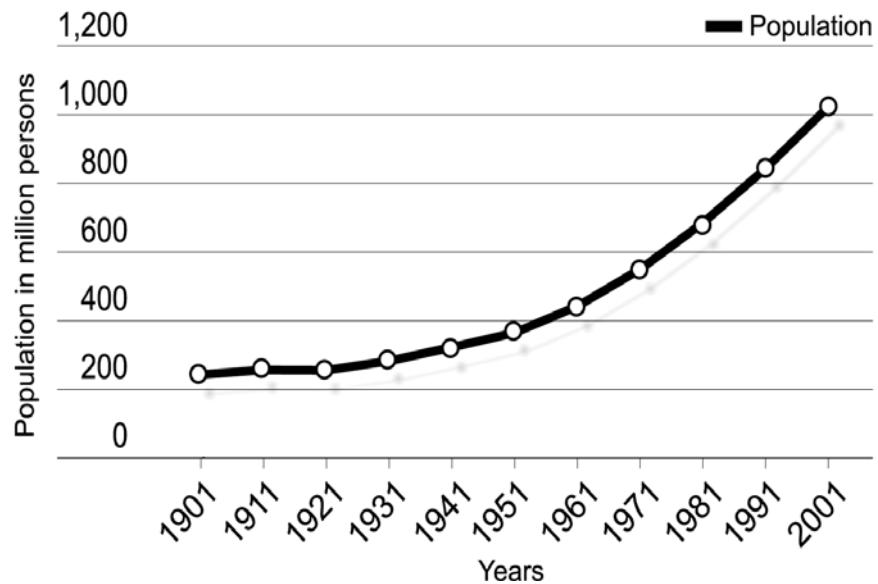


Fig. 31.1 Growth of Population in India (1901-2001)

(b) **Compound Graphs:** These graphs are being used to represent two or more dependent quantities at the same time. Different quantities represented by curves are either superimposed on the top of each other or placed on the each other in a cumulative way. For example, compound graphs of male and female population or rural and urban population can be used to represent the two segments of population. Similarly, variables having three or four segments can also be represented through compound graph. For example energy production (thermal, hydel and nuclear), migration streams (rural-rural, rural-urban, urban-rural and urban) and religious composition of population (Hindus, Muslims, Sikhs, Christians, Jains, Buddhists, etc. represent various segments of the variable.)



Table 31.2 : Sex Ratio of Population of India

(Population in million)

Years	1901	1911	1921	1931	1941	1951	1961	1971	1981	1991	2001
Male	120.9	128.3	128.5	142.9	163.7	185.5	226.2	284.2	354.3	439.2	532.1
Female	117.4	123.7	122.7	135.9	154.9	175.5	212.9	264.1	307	407.1	496.4

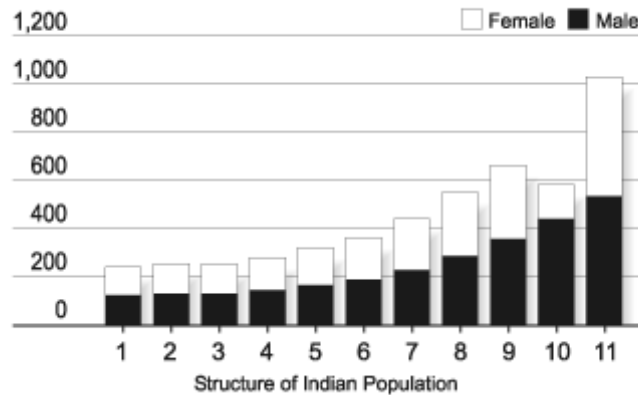


Fig. 31.2 Sex Composition of population of India (1901-2001)

Notes

(C) **Diagrammatic Presentations:** Diagrams are both graphical as well as geometric in nature. The processed data is portrayed through different diagrams for visual presentations. It is important to make use of diagrams based on their relative merit of visual presentation. The diagrams mostly refer to time or space or both the characteristics related to one location. Some of the diagrams used for the presentation of primary data are discussed below:

- (i) **Bar Diagram :** The use of column or bar has become common in representing a comparative performance of various units and growth of an individual unit. The length of bar is kept proportional to the size of production or the volume of change. Thus bar diagram is used to represent many elements at one point of time and one element across the time. The compound bar diagrams are used to represent the subclasses of an element. The block of a bar is proportionately subdivided to represent the sub classes in a compound bar diagram.
- (ii) **Pie Diagram:** The pie diagram is also known as divided circle. It is used to represent the proportion of the sub-unit of whole. The different segments of a circle represent percentage contribution of various components of data. For drawing a pie diagram, we construct a circle of any diameter. The circle is then divided into desired number of segments. i.e. angle 360 represents 100 percent. Pie diagram is generally used to represent the general land use of village, composition of shops in a



functional profile of urban areas, social composition of surveyed village, composition of total population.

- (d) **Presentation of Data through Maps:** Various types of maps can be constructed with the help of primary data. Maps related to various themes such as environment, trade, land use, production of community population, etc. can be prepared for presentation. A map is a proportional representation of some or whole part of the earth on a flat surface or piece of paper. Thus, the outline map represents the direction, distance and shape of the area, while the technique of the representation of data on the maps explains distributional characteristics. The method of preparing dot map is given below here as an example.

Dot Maps: These maps are used to show the dispersal as well as concentration tendencies (characteristics of a distribution) of the phenomena. Dot maps are related to point specific pattern of distribution unlike isopleth maps which are concerned with joining places having the same or equal values of distribution or choropleth maps which are concerned with area specific distributions instead of location specific distributions.

These maps use data to represent location specific distribution. The size of the dot is worked out considering the capacity of space on the map and the value of distribution at one point of location. The dot is assigned specific value in quantitative terms. Once the value of a dot is determined the number of dots at each location can be worked out. Dots are plotted on the map based on location specific distribution of variable. Due care needs to be taken while putting dots on the map.

Transport lines, rivers and canals, mountain tops and such other negative areas should be separated from placing dots. The final map clearly represents the concentration and dispersal of a distribution. The field data related to house-hold population, agricultural production, shop-wise daily sales or consumer pattern, unit-wise industrial production or field-wise crop can be better represented through dot maps.

For more details you are suggested to read the Practical Manual in Geography.

- Arrangement of information data either in ascending (from bottom to top) or in descending order (from top to bottom) is known as Array of data.
- Putting data on columns and rows to find the sum of the two sets for verification is called cross matching of data.
- A group of records showing similar data is called data flow.
- A set of data related to particular entity or a group is called the field.
- A complete set of information showing all basic data is known as master chart.



INTEXT QUESTIONS 31.2



Notes

1. Give single term to the following statements:
 - (a) The process of organizing data into groups or classes on the basis of certain characteristics.
 - (b) A graph used to represent two or more variables which are either superimposed or placed in a cumulative way.
 - (c) Grouping the data on certain basis.
 - (d) A measure which divides a series into 100 equal parts.
 - (e) The maps which are concerned with point specific pattern of distribution.

2. Match the following terms with the statements:

Terms	Statements
a) Array of data	(1) A person on whom questionnaire is administered.
b) Cross matching of data	(2) A complete set having all basic data.
c) Charts	(3) Arrangements of information either in ascending (from bottom to top) or in descending order (from top to bottom)
d) Respondent	(4) To put information on columns and rows to find the sum of the two sets.

3. State the three forms of data presentation.
 - (a) _____ (b) _____ (c) _____
4. Write the three forms of cartographic presentation.
 - (a) _____ (b) _____ (c) _____
5. Define the following terms.
 - (a) Pie diagram
 - (b) Median
 - (c) Coding of data
 - (d) Master chart.



31.4 INTERPRETING THE INFORMATION

Interpretation of information/data is crucial for written communication. It is an art of expressing a given data/information in a written or oral form to provide a logical explanation for the given facts. The following points should be kept in mind while interpreting the information:

- i) Clarity and explicitness of the interpretation.
- ii) Segregation of common and special features.
- iii) Focus should be clarified right in the beginning.
- iv) Organisation of the facts must be step by step.
- v) Accuracy of facts need to be checked.

I. Interpretation of a table: A table is a compact orderly arrangement of facts. It is summarized or grouped from a processed data. Interpretation of a table needs to start with the identification of minimum and maximum value i.e. ranges in the data. The difference between these two values explains the range to be comparatively smaller or larger. The smaller the range, lower the deviation and in the concentrated form is the distribution. Contrary to this, if range is larger, the interpretation will change as the distribution will be dispersed. The second step in the interpretation of a table relates to the analysis of various classes and their frequencies. The third step in the analysis of a table relates to the inferences derived. It should be brought out very clearly as what generalizations emerge from the table.

II. Interpretation of a graph: Graphs are different types and their interpretation varies significantly one another. The interpretation should be done with great care. There could be broadly two types of graphical interpretations. The first type of interpretation may deal with the amount of change with reference to time or areal units or both. The second dimensions of graphical interpretation is the trend. It is further divided into total trend and point specific trend.

III. Interpretation of a diagram: Each diagram has its own merit of presentation. It should be interpreted with regard to variables shown. A diagram highlights different levels of variables viz high, medium, low, very low etc. Interpretation of each component should be made clearly to give an idea about the performance of a variable across time and places.

IV. Interpretation of Maps: Interpretation of maps refer to area specific characteristics of a phenomenon. It could be with regard to time, intensity and community. The distributional, characteristics of a variable should be interpreted. It will bring out the distributions both in terms of volume

and area covered. Logical explanation should be given to the factor responsible for such a distribution.

- While interpreting the information certain points should be kept in mind. The points are clarity and explicitness, segregation of common and special features, focus, organization and accuracy of facts.
- The interpretation of processed data differs from one medium to another. For example, the interpretation of a table is different from diagrams, graphs and maps.



Notes

31.4 PREPARATION OF FIELD REPORT AND ITS FORMAT

Field reports are the written account of the facts and data collected from the field, its generalizations and basic conclusions. These reports are being used for comprehensive and application oriented learning. Implementation of various development schemes and plans are made depending on the conclusions derived, suggestions and recommendations made in the report. Since report forms the basis of decisions making, it needs to be comprehensive and capable of reflecting the ground truth. The field report should be prepared based on the following components:

- (a) **Introduction:** The first step in writing a field report is its introduction. The introduction includes the statement of the problem of field survey and its objectives. Methodology of the field work and the general background of the area of field survey has to be planned. The selection of samples and variables, hypothesis, processing and presenting the primary data from the part of mythology. The last part of the introduction is expected to discuss the scope and plan of the report.
- (b) **Analysis:** The value of the report is adjusted on the basis of insight and labour put in its making of a scientific and logical project. Analysis of the report is sub-divided into chapter of convenient number. Sequence of these chapters however, follows the system like 1) structure on nature of the theme of investigation. 2) Trends and patterns (both temporal as well as spatial) related to the theme of investigation. 3) correlation of associated factor influencing the problem under study 4) constraints and associated problems and 5) conclusions and suggestions. Each chapter contains logical and scientific analysis of the facts derived through the processing of data in the form of tabular and cartographic presentations besides investigators personal impressions gathered during the field work.
- (c) **The Results and recommendations:** The third and the important part of the field reports is related to deriving results and the recommendations. The generalisations made in each chapter are put together to form specific conclusions. To make suggestions more meaningful, constraints and likely problems should be worked out. Having analysed the entire theme



Notes

of survey both individually (at the level of variables) as well as collectively (at the level of groups), one is able to make final observation or to derive both broad as well as specific conclusions. The recommendation should be based on these results. Both basic as well as functional aspects of the problems should be covered by these recommendations. Before making recommendations one is expected to assess the viability and feasibility of the same. The smaller and specific is the dimension of the problem, more workable and viable is the recommendation. Similarly, the feasibility aspect of recommendation deserves to be assessed in the light of available technological, financial and social implications. The report must avoid vague and unclear recommendations. Thus, result and recommendations should touch upon finding solutions to problems faced and accelerating the pace of development.

Format of the field report: It is important to note that all field reports are special and unique in so many respects. However, there are certain formats which are common to all reports. On the basis of common characteristics, it may be summarized that a field report mainly consists of three parts. viz. (a) Parling (b) Body of the text and (c) Documentation.

(a) **The Prelims:** It consists of Title page, Preface, Table of contents, List of tables, list of maps and diagrams and list of Appendices.

Example:

Title of the Field report

Context of Field Report and Period of Survey

Name of the Investigator/Address

Name of the Project Supervisor

Name of the Institution or Organization

Year of submission

(b) **Body of the Text:** It includes from introduction to the conclusion and recommendations



Notes

Chapter Scheme:

- (1) Introduction
 - (a) Statements of the problem
 - (b) Objectives of the field work
 - (c) Methodology used
 - (i) Universe of the study
 - (ii) Selection of samples
 - (iii) Hypotheses proposed
 - (iv) Methods of data processing
 - (d) Scope and plan of the study
- 2. Nature or structure of the theme of Investigation.
- 3. Spatial and temporal trends of the problem of study. This chapter relates to understanding the area specific patterns and temporal trends.
- 4. Correlates the problem or investigations - It deals with the analysis of factors responsible for trends and patterns.
- 5. Constraints of theme of investigation - There are some basic and functional problems linked to each area. This chapter is devoted to study these problems.
- 6. Conclusions, suggestions and recommendation - This chapter summarises the findings, makes suggestions and recommendations for the development.
- (c) **Documentation:** It includes references, selected bibliography appendices, glossary of terms etc.



INTEXT QUESTIONS 31.3

- 1. State the three main parts that a field report consists.
 - (i) _____ (ii) _____ (iii) _____
- 2. Write seven points of Chapter Scheme of the field report.



Notes

**WHAT YOU HAVE LEARNT**

The data collected from the field are very extensive and unprocessed. While surveying in the field some objects remain unsurveyed and data, therefore, becomes dissimilar. Hence, there is need for processing the data properly. The different steps involved in processing data are editing, coding, organisation and classification. Only then the data becomes in the presentable form. The presentation of data could be tabular, statistical and cartographic forms. The tabular presentation could be simple or complex depending upon the variables used. Statistical presentation makes use of mean, median and mode for getting central values. Percentiles are also used to explain the coverage of a phenomenon studied. Cartographic presentation of data is made in different ways. Such as graphs, charts, diagrams, maps etc. Two variables can easily be represented by a line graph. Bar diagram is used for comparing different units. Compound bar diagram issued for representation the sub units of an element proportionately. Different types of maps are prepared with the help of primary data. The dot map is the most popular map. The dot map shows the distribution of an element. It also depicts the concentration and dispersion of the element. Isopleth map also depicts distribution of phenomenon. In this map, points of the same values are joined by curve lines. Distribution maps are also shown by shading methods.

The following points are kept in mind while interpreting the information. Clarity and explicitness, segregation of common and special features, highlight the focus, organise the matter in small paragraphs and facts should be complete and accurate.

Report is the most important component of the field work. It is a written document highlighting the conclusions drawn from the field work and data collected. The report should be extensive and related to ground realities. It should be written under the heads in a sequential orders of introduction, analysis, results and recommendations. .

**TERMINAL QUESTIONS**

1. What is data collection ? Describe any three issues that need to be covered in case of local area planning.
2. What are the tools and techniques of data collection?
3. Why is cross matching and array of data necessary in the organization of field data. Give any three reasons in support of your answer.
4. Explain any three steps in the processing of primary data.
5. What points should be kept in mind while interpreting the information.



6. Write a brief account of the Components related to the preparation of a field report.



ANSWER TO INTEXT QUESTIONS

Notes

31.1

1.
 - a) Primary data.
 - b) Secondary data
 - c) Tools of data collection.
 - d) Techniques of data collection.
 - e) Questionnaire
2.
 - A.
 1. Making oneself ready for collecting data from field situations.
 2. Keeping a field book /record book/diary.
 3. Administering questionnaire/schedule to the target group.
 - B.
 1. Acquiring knowledge about offices/ institutions etc. keeping the records of data.
 2. Getting an official letter for introduction and keeping identity card to get on entry in the office.
 3. Keeping a note book /record's file for transfer of data.
3. (a) ___ (4), (b) ___ (3), (c) ___(1) and (d)___(2)
4. (a) Identity specific contents (b) Respondent Centred Contents.
5.
 - i) The sample should be such that it reflects the characteristics of the whole.
 - ii) The sample should not be identical as it leads to error.

31.2

1.
 - (a) Classification of data.
 - (b) Compound graph
 - (c) Classification of data
 - (d) Percentile
 - (e) Dot maps.
2. (a) ___ (4), (b) ___ (4), (c) ___ (2), (d) ___ (1)
3. (a) Tabular (b) Statistical and (c) Cartographic
4. Graphical (b) Diagrammatic and (c) Maps



5. (a) A diagram which represents the share of sub-groups of an element within a circle.
- (b) The middle most position in a distribution.
- (c) To assign some alphabetical or numeral or both as the symbols.
- (d) A complete set of information showing all basic data.

31.3

1. (a) Prelims
- (b) Body of the text
- (c) Documentation
2. (i) Introduction
- (ii) Nature or structure of theme of investigation.
- (iii) Spatial and temporal trends of the problem of study.
- (iv) Data source and methodology
- (v) Correlates of the problem of investigation
- (vi) Constraints of theme of investigation
- (vii) Conclusions, suggestions & recommendations.

HINTS TO TERMINAL QUESTIONS

1. See Section 31.1
2. See section 31.2
3. See section 31.3
4. See section 31.3
5. See section 31.4
6. See section 31.4