UNIT 8

SOURCES

Structure

8.1 INTRODUCTION

You have studied and learnt the basic concepts of cartography along with maps and map scale in Block 1, and map projections and its main types in Block 2 of this course. In this Unit, you will study the sources of data mainly as cartographic input being used to analyse the cause and effect relationship between two or more than two variables in geographical studies. There are mainly two kinds of data sources - primary and secondary. In the next units, you will further study and learn about the 'Census and Sample Surveys' and 'Remotely Sensed Data' in detail along with their usage and importance in the realms of cartography and geographical studies as well.

Many of you are quite familiar with the data and its meaning if you look back in hindsight of your educational journey particularly at senior secondary level. However, you may be curiously thinking about many questions related to the use of data and its sources in higher education. For example, what is data, its use and importance in geographical studies and how many kinds of data sources are there etc.? Basically, data means facts and statistics, which are converted as information for measurements and analysis purposes. Data has very high significance in geographical studies and it forms the core of

cartography. It is greatly needed in geographical studies for quantification and measurements of geographical phenomena to arrive at intended outcomes. As you know that geography is not only concerned with description of the human and earth and its environment but it is also greatly concerned with the explanations of interrelationships and causality. Therefore, the data is required for explanations and examining causality with generalization, summarization, and for drawing inferences from various situations, projections and estimations and for visualization of spatial patterns as well. The data has been classified based on various criteria, like nature and method of data collection. It is further classified in terms of data organization or structure used primarily in data management and modern cartography achieved with the help of the state-of-the-art software driven technological system which is popularly known as 'Geographical Information System'. Examples of this are internal and external data, flat data, hierarchical data, object based data and relational data, etc.

In this unit, our focus will be primarily on the data sources of all sorts of data used as cartographic inputs in geographical studies. However, it is always a good practice that we acquaint ourselves with few basic concepts before embarking into the main theme. There are two types of data i.e. qualitative data and quantitative data. Qualitative data are the ones which are non-numerical types and measured by nominal or ordinal scales whereas quantitative data are the ones which are numerical and measured by interval or ratio data measurement scales. Both the data types are used in cartography and geographical studies. Therefore, it is inevitable to study various data sources as inputs for cartography. This unit will cover various kinds of data and data sources for cartography and geographical studies.

Expected Learning Outcomes

After completing the study of this unit, you should be able to:

- Explain the importance of various kinds of data and data sources for geographical studies and thematic cartography;
- Describe and formulate a plan to collect the primary sources of data according to the need of the study; and
- Discuss and classify the secondary sources of data for geographical studies and thematic cartography.

8.2 Primary Sources

Primary data are those data which are directly collected by the researcher or user. Refer to fig. 8.1 which shows various methods and tools for collecting primary data from various sources. These are first hand data collected for the first time for special purposes. Primary survey becomes essential when there is no secondary data available for the desired subject or time or the unit of study. It is mainly collected by employing various well defined methods and data collection tools. These methods may be surveys, observations, experiments, questionnaire, and personal interviews, etc. In geographical studies, we also differentiate geographical data as physical, which are primarily spatial and socio-economic data.

There are four types of data measurement scales: nominal, ordinal, interval and ratio.

Nominal Scale is such that uses numbers to represent identities, where no number represents the size or weightage like 1 representing males and 2 representing females.

In **Ordinal Scale**, numbers represent rank order indicating the order of quality or quantity without giving the magnitude of quantity or degree of quality like 1, 2, 3 and 4 representing very good, good, moderate and poor.

Interval Scale includes the number, which have continuity and magnitude of difference. However, it indicates zero as reference point but there is no true zero in like temperature.

Ratio Scale includes numbers like interval scale but it has true zero means absence of the object/phenomena of measurement, like weight or height.

Let us learn about key data collection methods and tools in cartography. Surveys are the methods of data collection. These surveys may be divided into two categories- i) socio-economic survey and ii) physical or spatial survey. Socio-economic survey is conducted to collect the data related to the themes encompassing demography, social, economic development, political, educational and health etc. You will appreciate and learn that such survey tries to cover a whole range of socio-economic issues related with the holistic and inclusive development of the people, society and nation. Physical survey in geography on the other hand includes reconnaissance survey, virtual survey using satellite imageries, field mapping, sample collection, land use and land cover survey, landforms identification and geomorphological mapping, landform association, morphometric survey, environmental survey, ground water survey and mapping, forest, soil, agricultural, wildlife, ground water quality, cadastral or property survey etc.

There are various data collection tools. In socio-economic survey, these tools are questionnaires or schedules, observations, interviews, focussed group discussion and field reports or field diaries. Questionnaires or schedules are the primary data collection tools through direct or indirect interaction with the respondents. The mode of interaction of researcher or investigator in case of schedule is direct and face to face with respondents, whereas in case of questionnaire the interaction is indirect through online mode (like email/online app or social media) or by post. Now the Information and Communication Technology (ICT) has also enabled researchers or investigators to have direct interaction with respondents through various ways, for example, telephone/ tele-conferencing or video conferencing or other social media.

There are certain merits and demerits in using these two tools. These are good for big enquiries. Schedules give more option to the researcher for observing and recording reality, but in questionnaire the respondent may not reveal the truth since there is no observer. In schedule method, respondents are not required to be able to read or write whereas in questionnaire method, these are the prime requirements. And, schedules are costlier method than questionnaire method.

Observations are also the data collection tools. It becomes a scientific tool and the method of data collection for the researcher, when it serves a formulated research purpose, is systematically planned and recorded and is subjected to checks and controls on validity and reliability. In this method, the information is collected by way of investigator's own direct observation without asking from the respondent. Using this method has certain advantages like elimination of subjective bias, if observation is done accurately. In this method, the information related to present situation is captured. It is independent of respondents' willingness to respond and is relatively less demanding of active cooperation on the part of respondents unlike interview or the questionnaire method.

Observation models include structured (planned) and unstructured (unplanned) forms to capture the information. There are two types of observations: i) Participant Observation and ii) Non-Participant Observation. Participant observations are making observations by the researcher or investigator as a part or member of the group experiencing the same with the group. In Non-Participant Observations, the researcher or investigator makes observation as a detached emissary without any attempt on his part to experience through participation what others in the group feel.

However, it has certain limitations as well. It includes relatively higher cost and very limited information gets captured and interference of unforeseen factors is also likely to happen in the observation. In some cases, the units of investigation are rarely accessible for direct observation in the field.

Interviews are the primary data collection tools involving personal verbal interaction either on face to face mode or using electronic medias- telephone, video calls, etc. It is mostly two way communication. There are various kinds of interviews like Focused Interview, Clinical Interview and Non-Directive Interview, etc. In Focused interview, the interviewer has the freedom to decide the manner and sequence in which the questions would be asked and has also the freedom to explore reasons and motives. In Clinical Interview, the broad underlying feelings or motivations or with the course of individual's life experience, based on the interviewer's choice for level of information to be extracted are taken into consideration. In Non-Directive Interview, the interviewer's function is simply to encourage the respondent to talk about the given topic with a bare minimum of direct questioning.

There are certain merits in this method like more information with greater depth can be obtained; interviewer by his own skill can overcome the resistance; greater flexibility under this method as the opportunity to restructure questions; personal information can be obtained easily; samples can be controlled more effectively; flexibility of controlling the sample for better responses; suitable language can be used to avoid misinterpretations; and supplementary information can be collected for interpreting results. However, there are some demerits also like very expensive method depending on geographical spread of sample; possibility of the bias of interviewer as well as that of the respondent; problem of supervision and control of interviewers; certain types of respondents are not easily approachable; relatively more time-consuming; chances of imaginary information to make interview interesting;

selection, training and supervising the field-staff more complex; sometimes very difficult to make prerequisite proper rapport with respondents; etc.

Focussed Group Discussion is a primary data collection tool, which involves the gathering of people of similar backgrounds in a group to get certain information about specific topic through group discussion. It requires a moderator and ideally has 6-12 people in a group. The role of the moderator (generally researcher or investigator) allows free and open discussion by all members. The group members are required to be homogenous but close friends and relatives are to be avoided. In this method, control over dominant voice is required. Qualitative data is captured through this method.

Field Reports or Field Diaries are also important sources of primary data. The observers record their observations and measurements in the field note books and later on translated into data and information.

The primary sources of physical or spatial data are field mapping and remote sensing. Field mapping includes drawings, field measurement books and records. It may be traditional or manual like chain-tape, plane table, prismatic compass, theodolite, etc. Semi-automated or automated field mapping involves Global Positioning System (GPS) or Differential Global Positioning System (DGPS) survey, Electronic Total Station (ETS) survey, etc. Remote sensing is also a method of primary data collection, which involves aerial photography, satellite imaging, radar imaging, Shuttle Radar Topography Mission (SRTM) data, which is high resolution digital topographic data for entire globe, very-very high resolution Light Detection and Ranging (LIDAR) data or laser scanned 3D data and drone imaging/mapping.

There are some non-spatial data related to physical phenomena also which are collected through special field observations like for weather, water, environment and pollution, etc.

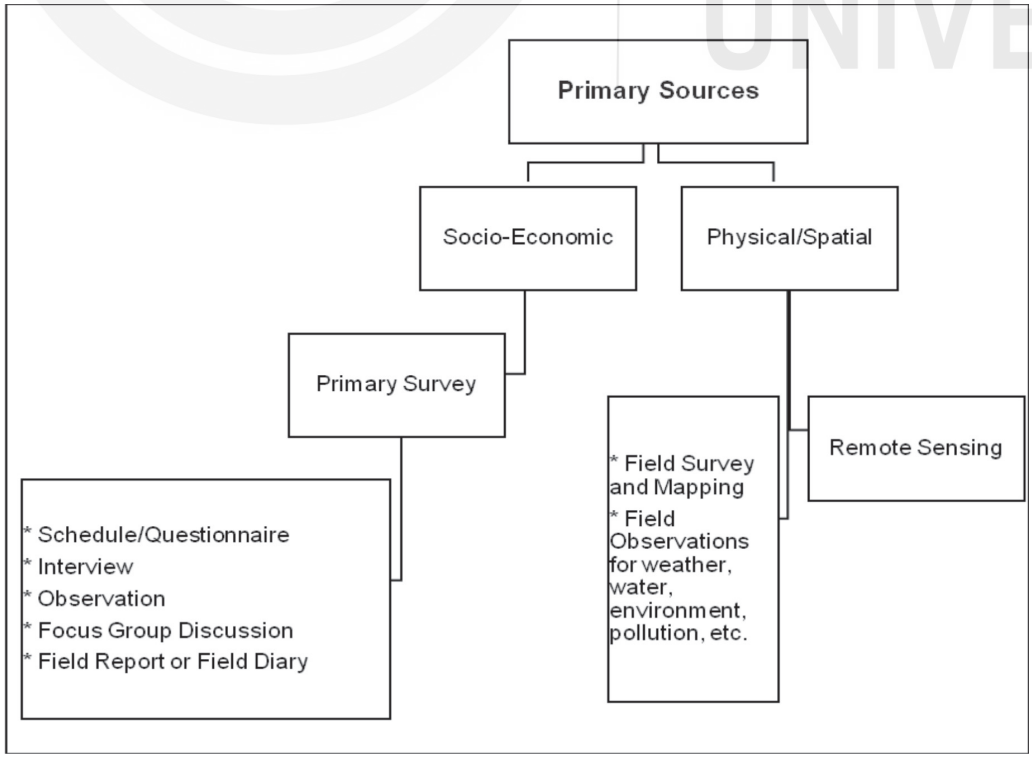

Fig. 8.1: Primary data source for geographical studies and cartography.

SAQ 1

What are various primary sources of socio-economic data? What are various primary sources of physical or spatial data?

8.3 Secondary Sources

Secondary Data are those data which have already been collected and processed or tabulated and may be published or unpublished. In other words, the data created by other than researchers or users are known as secondary sources. The secondary data can be in the form of tabular, textual, pictorial, drawings, maps and imageries. The secondary data sources include published and unpublished documents or data. These are government and non-government sources. Besides, books, journals or periodicals, reports by various government and non-government agencies are also secondary sources for other users. Nowadays, online documents are also considered as secondary sources of data.

There are national and international sources of the useful data being used in cartography and geographical studies. The major international sources are United Nations Organizations which provide various data related to demography, health, education, labour and economy, environment and climate, and many more themes or issues of contemporary relevance. Some private agencies also provide the online data related to environment and daily weather conditions for the entire globe like https://www.timeanddate.com/. However, few provide the disaggregated level data for a country.

In India, there are large number of secondary data sources. For socio-economic data, the largest and oldest has been the 'Census of India' followed by the 'National Sample Survey Organisation'. Sample Registration System has also been a source for vital data. These are provided by the government agencies. Apart from these, 'National Family Health Survey' also provides the data on various aspects of health up to district level since 1992. Though, you will study in details about all these in the next unit of this Block, we will discuss here these briefly. 'All India Educational Survey' also provides the data on school education at hamlet level. 'Unified District Information System for Education' (UDISE) also gives the data on school education at school level. Annual Report of 'University Grants Commission' (UGC) gives the statistics on the status of higher education in the country. Crime Record Bureau provides the crime data at state, district and town level. Some states also provide online crime data. Livestock census provides the data on livestock and agricultural implements. Indian Agricultural Statistics provides the district level crop-wise data on land use. Fertilizer Association of India brings out the data on fertilizer consumption. Agricultural Situation in India provides some data on agricultural production. Department of Statistics and Evaluation provides compiled data for various spheres at state level and also at district level for various states. Some private data providers like 'Centre for Monitoring Indian Economy' also compile and provides the data on various aspects. Different government departments like transport (surface, rail and roads, water and air) give data on transport and

freight. Similarly, various government departments provide data related to their domain including the data on natural resources. There are many other sources which provide various kinds of data for geographical studies and cartographic work.

There are many online data providers. Many government sources provide online data including Census of India. There is one government portal which compiles and provides the data of various sectors - data.gov.in. There are some private data providers also like www.indiastats.com, https://www.skymetweather.com/etc.

The secondary data sources related to maps and similar data for cartography and geographical research are also many and varied. National Remote Sensing Centre (NRSC), Hyderabad is the provider of all kinds of satellite imagery in the country. It also gives various atlases and real time maps for applications like land use/land cover, agriculture, drought, flood, wasteland, watershed desertification, etc. Such kinds of maps are also prepared by the Regional Remote Sensing Centres of Indian Space Research Organisation, ISRO. Similarly, Survey of India (SOI) located at Dehradun in the state of Uttarakhand also provides the toposheets at various scales. National Atlas and Thematic Mapping Organization (NATMO) publish and provide thematic atlases and planning series maps at district level for the entire country. National Bureau of Soil Survey and Land Use planning (NBSS&LUP) and All India Soil & Land Use Survey (AIS&LUS) which were later renamed as Soil & Land Use Survey of India also provides maps and data on land use

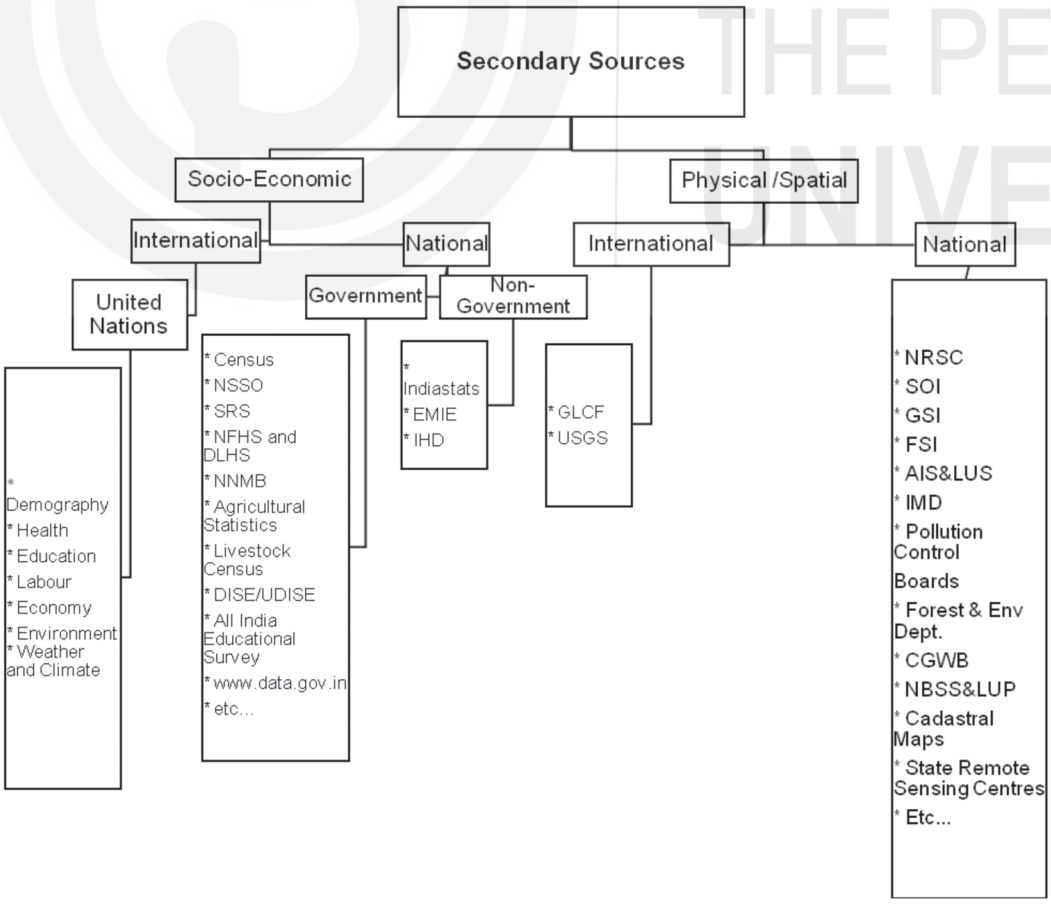

Fig. 8.2: Secondary data sources for geographical studies and cartography.

and soil. Similarly, Geological Survey of India (GSI) provides the geological maps, Indian Metereological Department (IMD) provides the weather maps and charts, Forest Survey of India (FSI) provides the forest maps in India. Census of India also provides some maps related to regions and socio-economic maps. The cadastral maps are also provided by the state governments in print and also in digital formats. Besides these, various state level agencies also provide maps and data especially state remote sensing or space applications centres provide various maps. There are also sources of tabular data on physical aspects of geography like Indian Meteorological Department (IMD), Pollution Control Boards of Central and State Governments, Weather Stations, Department of Forest and Environment, Central Ground Water Board, etc. (Fig. 8.2).

There are online sources of maps and satellite imageries. In India, Bhuvan is the platform from where satellite data and maps can be downloaded. The data from international providers are also available like The Global Land Cover Facility (GLCF) and United States Geological Survey (USGS).

SAQ 2

What are various secondary sources of socio-economic data? What are various secondary sources of physical and spatial data?

8.4 SUMMARY

In this unit, you have studied so far:

- Primary sources of socio-economic data.
- Primary sources of physical or spatial data.
- Secondary sources of socio-economic data.
- Secondary sources of physical or spatial data.
- Besides these, you have also learnt that both the sources of data i.e. primary and secondary are widely used in cartography to show the overall picture that prevails at ground level in reality at a glance in a map on the one hand. Both the data sources are also widely used in geographical studies in order to draw the rational and meaningful inferences for research, teaching and planning purposes etc. on the other hand.

8.5 TERMINAL QUESTIONS

1. Write the importance of data in cartography and geographical studies.
2. Explain the primary sources for socio-economic data.
3. Discuss the primary sources for physical or spatial data.
4. Write about the secondary sources for socio-economic data.
5. Explain the secondary sources for physical or spatial data.

8.6 ANSWERS

Self Assessment Questions

1. Various primary sources of socio-economic data are primary surveys using data collection tools such as schedules and questionnaires, observations, interviews, focus group discussion, field reports or field diaries. The primary sources of physical or spatial data are field survey and mapping and remote sensing.

2. Various secondary sources of socio-economic data are various published and unpublished documents and reports by the Government and non-government agencies including the United Nations and national and international organisations or data providers. These sources are in print or digital/online formats. The secondary sources of physical or spatial data are various international, national, and state level government and non-government agencies providing data in the form of maps and imageries (satellite, radar, laser and magnetic, etc.).

Terminal Questions

1. Your answer should cover the importance of data in cartography and geographical studies in the light of making measurements, generalization, summarization, and drawing inferences from various situations, projections and estimations and also for visualization of spatial patterns. Refer to the Sec. 9.1.

2. List all the primary sources for socio-economic data and explain those sources like primary survey with all data collection tools (schedule and questionnaire, interview, observation, focus group discussion and field report or field diaries). Refer to the Sec. 9.2.

3. Explain the primary sources for physical or spatial data from survey and field mapping and various kinds of remote sensing. Refer to the Sec. 9.2

4. List the secondary sources for socio-economic data including international, national and state level government and non-government sources. Refer to the Sec. 9.3.

5. Explain the secondary sources for physical or spatial data including all the agencies from international to national and state level with various kinds of spatial data provided by them. Refer to the Sec. 9.3.

8.7 REFERENCES/SUGGESTED FURTHER READING

- Robinson, A., Morrison, J.L., Muehrcke, P.C., Kimerling, A.J. and Guptill, S.C. (2011). *Elements of Cartography*, New Delhi: Wiley India.
- Lo, C.P. and Yeung, A.K.W. (2007): *Concepts and Techniques* of *Geographic Information Systems*. New Delhi: Prentice Hall.
- Agriculture Census Division (2015): *All India Report on Agriculture*

Census 2010-11, New Delhi: Department of Agriculture, Cooperation & Farmers Welfare Ministry of Agriculture & Farmers Welfare Government of India.

- http://dahd.nic.in/Division/statistics/animal-husbandry-statistics-division accessed on 26 April 2019.
- *19th Livestock Census-2012, All India Report*, Ministry of Agriculture, Department of Animal Husbandry, Dairying and Fisheries, Govt. of India, New Delhi.
- *History of Census in India*, http://censusindia.gov.in/Ad_Campaign/drop_in_articles/05-History_of_Census_in_India.pdf accessed on 26 April 2019.
- http://censusindia.gov.in/2011-Common/Sample_Registration_System.html
- http://censusindia.gov.in/vital_statistics/SRS/Sample_Registration_System.aspx#2
- National Sample Survey Office (NSSO), http://www.mospi.gov.in/national-sample-survey-office-nsso#
- http://mospi.nic.in/141-historical-perspective
- http://rchiips.org/pdf/dlhs4/report/APR.pdf
- IIPS (2004): *Reproductive and Child Health, District Level Household Survey, 2002-04, India*, New Delhi: MOHFW, Govt. of India.
- International Institute for Population Sciences (IIPS) and ICF. 2017. *National Family Health Survey (NFHS-4), 2015-16*: India. Mumbai: IIPS.
- www.data.gov.in
- http://schoolreportcards.in/SRC-New/AboutDISE/AboutDISE.aspx.
- http://udiseplus.gov.in/
- http://bhuvan.nrsc.gov.in/bhuvan_links.php