

CHAPTER 2

DATABASE AND METHODOLOGY

2.1 INTRODUCTION:

Education and health has become an important strategic factor in every major economy of the world. For quickening the rate of growth and equalizing the distribution of income, education and health has been placed as a central issue in the economic infrastructure. The developing significant of education and health has been paralleled by the copiousness of socio-economic literature, and yet no body is intelligibly able to explain a linkage between education, health and labour productivity. Thus, socio-economic analysis appears to impart the understanding among countries so as to complement empirical workers on the role of education and health in productivity analysis.

2.2 SELECTION OF THE STUDY AREA

In the present study, Tribal Sub-Plan area of Southern Rajasthan has been selected as the study area for an interrogation into spatial Attributes of educational and health disparities because of several specific reasons. TSP area is considered to be one of the under developed area. The Present study considered all the 24 tehsils of Tribal Sub-Plan Area of Southern Rajasthan. TSP area has the population of 5934792 persons (Census 2011). This distinction is not a matter for jubilation, rather it is a matter of grave interest to bring in and act upon, to rigging the problems related to growth of population at a faster rate and its wider on education and health sector to such a large multitude of persons.

2.3 DATA BASE:

Data have a very critical and important role in giving the abstract shape of the unit. These also help in either comparison or contrasting units within the model of particular geographic phenomena.

2.4 SOURCES OF DATA:

The present study is based on secondary data. This study for the analysis of Levels of Education, Health and Combined levels of Development has been worked

out through transversal section data. Further study is done with the help of time series data.

In order to approximate the model used is secondary data. The study period covers 1971 to 2015 as far as secondary data is concerned. Secondary data has been used to appraise the education and health in causal factor in the region. The secondary data have been collected from the published and unpublished reports. Census data of TSP area are collected from Directorate of Census Rajasthan. Education Data Collected from Secondary & Elementary District Education Office, Rajasthan Sarva Shiksha Abhiyan (RAJSSA) office for Dungarpur, Banswara, Pratapgarh, Udaipur and Sirohi District, and MLSU. Health data Collected from Joint Directors office Udaipur, Chief Medical & Health Officer office for Dungarpur, Banswara, Pratapgarh, Udaipur and Sirohi. Deputy Director Department of Ayurveda Udaipur. All the Sources of data have been helpful for analyzing the educational and health facility, levels of development, growth of development.

The collected data have been processed in tabular form in order to deduce specific determination. Different statistical techniques used in the analysis, including the factor analysis, composite indices, correlation, ratio and percentage method. Geographical Information System (GIS) techniques have been used to analyze and represent the data through Choropleth Map, Liner Graph, Pie Diagrams and other methods.

2.5 SELECTION OF INDICATORS:

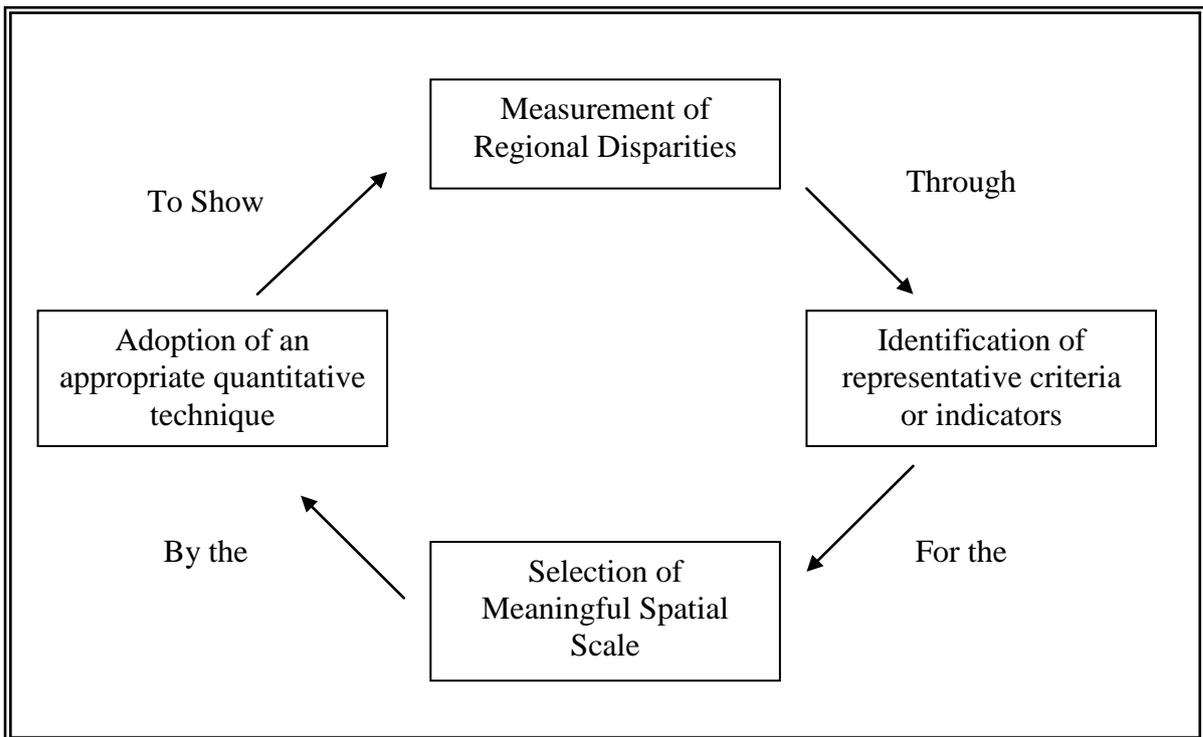
The study is basically analytical in its intervention of the data. The basic aim is to find out fluctuation in educational level and health level in an under developed tehsil. The selection of indicators under each factor of educational and health level is made with a view that it is a correspondence to the educational and health disparities in the TSP Area. Since the study area has a high level of regional disparities in educational and health level, simple and direct step of educational and health developments were attempted. The selected indicators are all correspondence to the status of education and health for the study area.

2.6 METHODOLOGY:

In the present study, the tribal sub-plan area has been selected for a query into spatial attribute of educational and health disparities. Data related to educational and health statuses have been collected at tehsil's level from secondary sources for the

analysis of the data. The factor analysis technique is used. Through this technique, factor loading and factor scores have been calculated. Factor analysis has been used to reduce the large number of indicators, many of them interrelated to a few independent implicit in attributes called factor which can be construed as being responsible for the spatial variation. Descriptive names have been given to each factor based on the indicators most closely associated with that factor and the concept expressed by the cluster of indicators. Finally, the resulting structural and regional pattern have been analyzed with the general impression and data based ideas.

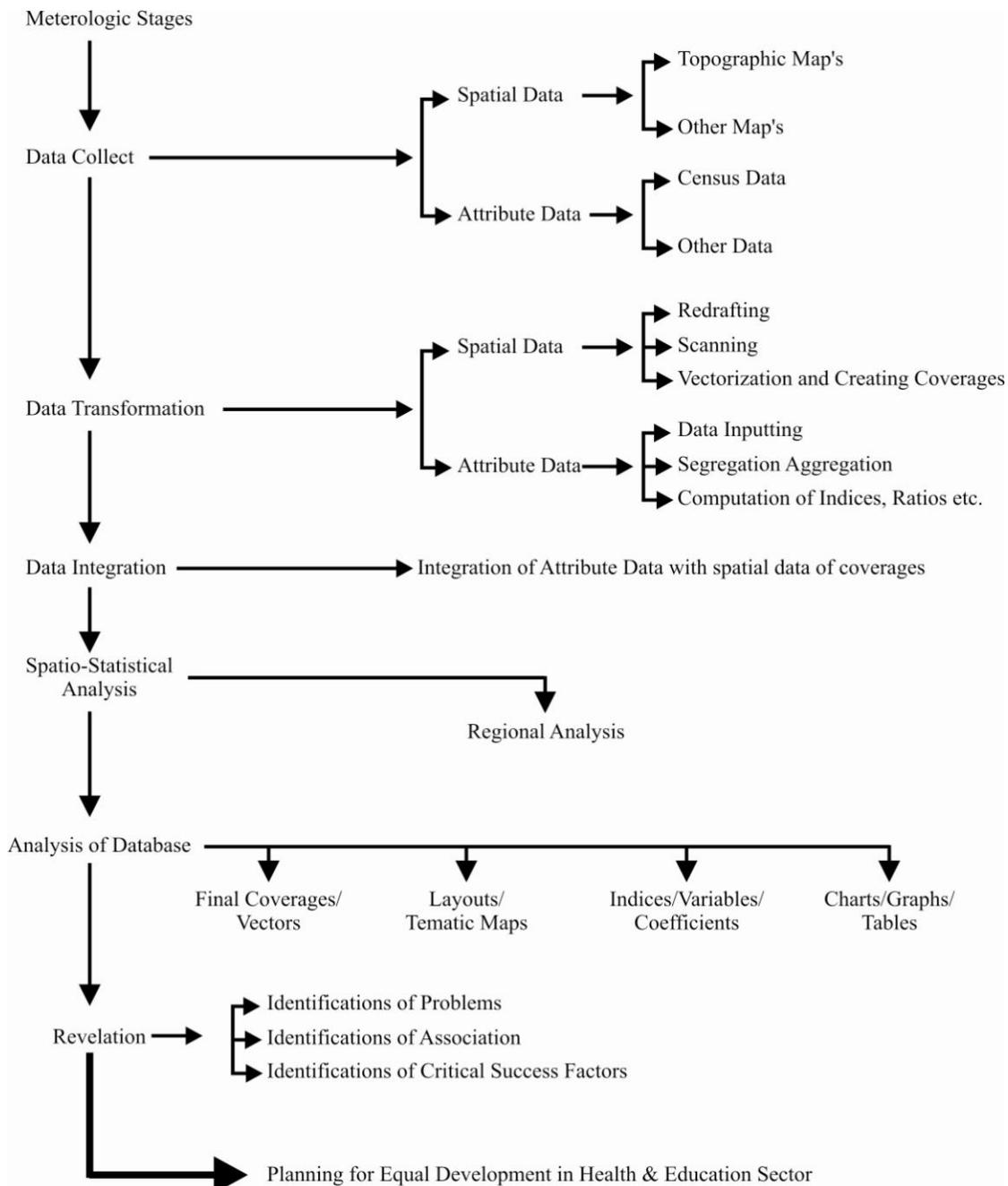
In an exercise of development, identification of regional disparities regards three major problems: Selection of capture and wise indicators to present various attributes of development; to use an objective technique for the combinations of assortment of data relating to all the indicators for every numbering unit and to develop a systematic statistical and cartographic technique to represent the process and shape of development. The importance of measurement of these regional disparities may be envisioned in the following figure.



Similar are the issues when we need to address regional disparities in educational and health development. So regarding all these prospects in methodology has been evolved to deal with these problems and so to reach at some concrete conclusion.

2.6.1 Nature of Data: -

The data related to Education, Health and Census Data of TSP area considered indicators such as EI, LR, EE, ED, HI, CBR, CDR, IMR, NIR, HR, Gender Ratio, Population, Occupational etc. The information regarding the education and health facility such as number of PS, UPS, SS, SSS, College, SCs, PHCs, CHCs, Ahs, Doctors, Beds, Nursing Staffs of TSP area taken from the tehsils levels data of various departments.



Software : SPSS and GIS has been used for analysis and layouts

2.6.2 Methods of Analysis:

The Data analyzed is based on secondary sources. The tehsils are taken as unit of analysis and interpretation. All the statistical techniques and GIS tools are used at tehsil's level only which may have limitations and scope of the study.

Scheming a suitable methodology and selection of analytical tools are important for a meaningful analysis of any research problem. This section is devoted to a description of the methodology which includes choice of the study area, period of study, collection of data and tools of analysis.

To study the objectives, the effects of education, health facilities and levels of development in the Tribal Sub-Plan area is analysed through the Composite Index, Correlation analysis and Factor Analysis (PCA) is used for the study of education and health disparities and levels of development in TSP area by using Aggregate Educational Development, Aggregate health development and Aggregate Combined development , composite, Correlation and Factor Analysis used for constructing indices of twenty four tehsils the study of education and health levels or disparities in TSP area.

2.6.3 Nature of analysis

The present study is based on qualitative and quantitative methods and techniques postscript with suitable cartographic representation. The following methods and techniques have been applied;

1. The descriptive approach has been applied for the physio-cultural analysis of the study area. For the representation of relief, interpolation method based on spot height has been used.
2. Technique of partitioned value has been applied for measuring the variations in the distribution, density, population growth of population, gender ratio, occupational structure at tehsil's level areal units, besides determining the distribution of socio-economic facilities in the district.
3. The technique of correlation and regression has been adopted to examine the relationship between average spacing and mean size of education and health at tehsil level. The relationship between education and health facilities and their correlates has measured by correlation analysis with the help of SPSS.

4. In determining the levels of educational and health development at tehsil, the functional gravity of the education and health has computed, based on composite score.

2.6.4 Composite Indices:

To determine the level of development composite index method has been used. Following steps are involved to calculate the composite index.

1. First of all mean of each indicator has been identified.

- a) Arithmetic mean of indices-

$$\bar{x} = \frac{\sum x}{N}$$

Where= $\sum x$ = sum of indices

N= number of indices

2. Standard deviation (S) of each indicator has been calculated.

- b) Standard deviation of indices

$$\sigma = \sqrt{\frac{\sum d^2}{N}}$$

Where- $d = \bar{x} - x$ – deviation from actual mean

3. Standard values has ben calculated by using the following formula

- c) Standard Score $Z_{ij} = (X - \bar{X}) / S_j$

Where \bar{X} = Mean of the jth indicators

S_j = STDEV of jth indicator

4. Gross values of each tehsils has been calculated by adding the standardized value of all indicators found in that tehsils

- d) G.V. = sum of total indicators

5. At the last composite index has been identified

- e) Composite index = G.V / N

G.V. = Gross Value

N = Number of Indicators

2.6.5 Factor Analysis

The technique of factor analysis was initially developed by psychologists early in the twentieth century as a means of analyzing the result from intelligence test (Royce, J.R. 1958) and later used by other field. It is a method of studying at the same time the complex inter-relationship between many variables as measured from different observations and summarizing prominent features of relationships in the form of a few basic patterns called factors. Although studies of geographic nature were attempted at an early date by sociologists (Hagood, M.V. et al., 1941) and the technique has been used extensively by geographers in the period since 1960. The factor analysis technique has been applied by various geographers in their studies of economic regionalization (Berry, B.J.L., 1960), Climatic regionalization of urban areas (Carey, G.W., 1966) and the analysis of good flow patterns (Berry, B.J.L., 1966).

One of the most difficult tasks in component analysis, however, is the recognition or giving some meaning to the newly produced factor in the light of the original data, because variables may burden about equally with a number of factor instead than correlate with as few factors as possible. The factors are infusion with the analysis permit mapping the basic attributes of the indicators used for the identification of spatial variations in the level of education and health sector.

The factor analysis conducted for this study is based on 33 indicators pertaining to regional disparities in the levels of education in 24 tehsils of Tribal Sub-Plan Area.

Computation for this analysis has been carried on SPSS-20. The model for the factor analysis used in the study involves the following steps:

1. Initial computation is based on transformation of the original data matrix D for 'n' observation on 'm' Variables into a standard score matrix z of $(n*m)$ order.
2. From the standard score matrix Z an order correlation matrix r has been calculated which contained productive moment correlation coefficient between each indicator and every other indicator.
3. This correlation matrix has been resolved into a factor loading matrix A of $(m*r)$. The programme employed is such that it can extract as many factors as

the number of indicators, but only three factors are calculated in the present study for case of analysis.

4. From the standard score matrix of $(n*m)$ order and factor loading matrix of $(m*r)$ order a factor score matrix of $(n*r)$ order has been obtained. These factors score (normalized to zero mean and unit variance) provide a measure of position for each observation (tehsil) on the new factors.
5. Eigen value are the sum of squared factor loading for each factor and indicates the amount and proportion of the total variance in the original data accounted for by each factor.
6. The sum of the squared factor loading across each row of the matrix A is known as communalities and tell the proportion of the total variance of each indicator which is accounted for by the 'x' factor together.

The Z-score method is also used to calculate the standardized scores. It is a linear transformation of the original data in such a way that its mean becomes zero and its standard deviation becomes unity.

Further, the result of standard score obtained for the different indicators were aggregated in order to find out the composite z-score then, the composite z-score has been categorized into high (above 0.50), medium (0.50 to -0.50) and low (above -0.50), so that status of education and health may be obtained on a common scale. For measuring the relationship between the status of education and health and level of socio-economic development z-score technique has also been applied in order to read standardization. The raw data for each variable have been computed into standard score. Lastly, Geographical Information System (GIS) and advanced computer cartographic techniques have been applied to organise the various Choropleth maps.