Chapter: 3

Research Methodology

This chapter comprises research methodology and its application in answers to research problem. We need to have a good research methodology in place when trying to figure out a new ideas. “Methodology” implies more than simply the methods you intend to use to collect data. It is often necessary to include a consideration of the concepts and theories which underlie the methods. When we describe methods it is necessary to state how we have addressed the research questions or hypotheses. The methods should be described in enough detail for the study to be replicated, or at least repeated in a similar way in another situation. Every stage should be explained and justified with clear reasons for the choice of your particular methods and materials.

This chapter particular attempt will be made to define the research problem profiling of the population and the sample taken from the specified population for this study.

3.1 Object of Research - “An analysis of Distance Education in the Implementation of Universal Education in India”

3.1.1 Sub Objectives

3.2.1. To analyze the effect of distance education in Indian education system.
3.2.2. To analyze the challenges of distance education
3.2.3. To analyze the Participation of women in distance education.
3.2.4. To analyze the Scope of Distance education in future.
3.2.5. To analyze the quality of education in distance mode

3.2 Significance of study: The Study is significance for the following theoretical and practical concerns

3.2.1 Theoretical concern: Theoretically, This study seeks to explore the effect of distance education upon Indian education system. How much its help to increase the improve literacy in higher education. It also helps in understanding what students think about distance education and what factor attracts them toward enrolled in distance education
2. The finding of the study reveal to analyze the perception of student toward the distance education
3. The finding of the study reveals to analyze the role of distance education to educate women’s
4. It will helpful to the researcher who plan to undertake in depth research studies and practices in different dimensions of distance education.
5. The finding of the study will help in finding some hidden and unexplored facts of distance education in India.

3.2.2 Practical Concerns:
1. It may be l helpful for the Government of India to improve the quality of education on different dimensions of distance education.
2. It will helpful to the those student who are in dilemma because of rumors about of distance education.

3.3 Hypothesis Formulation-
Ordinarily, when one talks about hypothesis, one simply means a mere assumption or some supposition to be accepted or rejected. But for a researcher hypothesis is a formal question that he intends to resolve. Thus a hypothesis may be defined as a proposition or a set of proposition set fourth as an explanation for the occurrence of some specified group of phenomena either asserted merely as a provisional conjecture to guide some investigation or accepted as highly probable in the light of established facts. Quite often a research hypothesis is a predictive statement, capable of being tested by scientific methods, that relates an independent variable to some dependent variable.

\( H_{01} \). There is no significant effect of distance education in Indian education system.

\( H_{02} \). There is no significant difference in opinion of students about quality of education in Distance Education in India.

3.4 Research Design:
I have used the conclusive research method in this study. As our study deals with the study of distance learning education so we need to meet the student. Our primary aim is to find the psychology of student, what they think about the distance learning education system. Also, we collected data from Open/Distance universities of India by using Right to Information 2005 Act for that I have framed some questions.
Conclusive research provides information that he can make a rational decision. In some instances particularly if an experiment is run, the research in many come close to specifying the precise alternative to choose in other cases specially with descriptive studies the research will only partially clarify the situation, and much will be left to the executive judgment. This study has done well while attempting to arrive at a more clear description of an apparent problems. Our primary aim is to find effect of distance education in Indian education system and Quality of Education.

3.5 Sampling Process: -

3.5.1 Non-Probability Sampling Method: - This method does not provide all the items in the universe with a known change of being included in the sample. The selection process is at least partially subjective. Actually our research project is based on Distance learning education system, so we used Convenience sampling method for my research project.

Our research project is based on Distance learning education system, so we used Convenience sampling method for collecting data.

3.5.2 Convenience Sampling:
Convenience sampling is a non-probability sampling technique where subjects are selected because of their convenient accessibility and proximity to the researcher.

3.6 Area of Research
We collected Secondary data from IGONU, 11 State Open Universities and 71 Dual mode Universities from various part of India. Who provided distance education from 2009-10 to 13-14. Also collected primary data from students enrolled in distance mode of education to know their opinion regarding quality and other parameters of distance education.

3.7 Sample Unit: Sampling unit contains the particular units being studied and which generally reflects the desired response that is most of the traits that the universe has.

3.7.1 Secondary Data Source
- Total 83 universities selected to collect data from all over India.
  1 Nation Open University
  11 State Open University
71 Dual Mode Universities

3.7.2 Primary Data sources

- Total 500 respondents fill the online and offline questionnaires, who are/were a part of distance education from various parts of the country.
- Link of online questionnaires, -
  https://docs.google.com/a/shobhituniversity.ac.in/forms/d/e/1FAIpQLSeJyP8_QG6islJMikwZC0uYv3rMst8hyLnFer0stS28RW-Zng/viewform

Figure 3.1 Link of online questionnaires
3.7 Sources of Data Collection

As with any kind of research, the researcher aware about the value of secondary data. The secondary data provided the researcher information regarding the activities scope and opinions of other researcher and experts in the initial stage. The data further guide the researcher in defining the variables of the study. The secondary data provide useful and necessary information supplementing the qualitative and quantitative aspects of research finding.

For this purpose secondary data will be collected from all associated sources that include

(i) Data collected from various universities by using “Right to Information Act 2005”
(ii) Data from Distance learning council
(iii) Data from UGC,MHRD,DEB,NKC Etc.
(iv) Books
(v) Research journals.
(vi) Magazines
(vii) Newspapers, published materials & websites,etc
3.9 Primary Data Collection

A preliminary study of Distance learning education was done by this method of research. For this purpose structured and non-structured questionnaire was prepared and administered to various students in different part of India. Questionnaires are commonly used to gather first-hand information from a large audience, in the form of a survey. This questionnaire is framed in such a way that it provide us continue response from the student’s available in centers and to achieve flexibility by asking certain closed ended questions. The time constraints and various other factors were also kept in mind while framing the research methodology. There are different types of questionnaires and the type of questionnaire to be used usually depends on the purpose of the survey and the type of data that has to be collected. Questionnaires are highly practical and can be carried out by any number of people, and the results can be quickly quantified as well.

3.9.1 Types of Questionnaire Used: Depending upon the nature of the questions in a questionnaire, there can be different types of questions in questionnaire –

**Open Ended Questionnaire**

Questions that allow the target audience to voice their feelings and notions freely are called open-format questions or open-ended questions. These questions are not based on pre-determined responses, giving respondents an opportunity to express what they feel is right, and often provide real, perceptual, and at times, startling proposals. Open-ended questions placed at the end of a questionnaire tend to draw accurate feedback and suggestions from respondents as well.

**Closed Ended Questionnaire**

Questions which have multiple options as answers and allow respondents to select a single option from amongst them are called closed-format or closed-ended questions. This type of questionnaire is especially useful when conducting preliminary analysis. As a fixed answer set is provided, these are ideal for calculation of statistical information and percentages of various types. Closed-ended questions help to arrive at opinions about a product or service, and sometimes, about a company, in a more efficient manner.
3.9.2 Don’t Know Response:
The Don’t know response presents special problems for data preparation. When the don’t response group is small. It is not troublesome. But there are time when it is of major, and it may even be the most frequent response received. Does this mean the question that elicited this response is useless? The answer is “It all depends.” Most don’t know answers fall into two categories.

- First there is the legitimate don’t know response when the respondent does not know the answer. This response meets our research objective: we expect don’t know response and consider them to be useful.
- The second situation a don’t know response reply illustrates the researches failure to get the appropriate information. It is reasonable to expect that some legitimate don’t know response will be made to each these question.

3.9.3 Sample of Students: Total 500 students selected from various part of India who are/were part distance education. Study based on four basis demographic factors:

Age: Age Group divided into five classes

<table>
<thead>
<tr>
<th>Age</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-20</td>
<td>231</td>
<td>46.2</td>
<td>46.2</td>
<td>46.2</td>
</tr>
<tr>
<td>21-25</td>
<td>216</td>
<td>43.2</td>
<td>43.2</td>
<td>89.4</td>
</tr>
<tr>
<td>26-30</td>
<td>31</td>
<td>6.2</td>
<td>6.2</td>
<td>95.6</td>
</tr>
<tr>
<td>31-35</td>
<td>13</td>
<td>2.6</td>
<td>2.6</td>
<td>98.2</td>
</tr>
<tr>
<td>36 &amp; Above</td>
<td>9</td>
<td>1.8</td>
<td>1.8</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>500</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
**Sex:**  Sex is classified into two male and female.

**Table 3.2: Sex and frequency:**

<table>
<thead>
<tr>
<th>Sex</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>344</td>
<td>68.8</td>
<td>68.8</td>
<td>68.8</td>
</tr>
<tr>
<td>Female</td>
<td>156</td>
<td>31.2</td>
<td>31.2</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>500</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
Qualification: Qualification group divided into four classes.

Table 3.3: Qualification and frequency

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>UG</td>
<td>343</td>
<td>68.6</td>
<td>68.6</td>
<td>68.6</td>
</tr>
<tr>
<td>PG</td>
<td>71</td>
<td>14.2</td>
<td>14.2</td>
<td>82.8</td>
</tr>
<tr>
<td>Ph.D/M.Phil</td>
<td>27</td>
<td>5.4</td>
<td>5.4</td>
<td>88.2</td>
</tr>
<tr>
<td>Others</td>
<td>59</td>
<td>11.8</td>
<td>11.8</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>500</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
Employment: Qualification group divided into five classes.

Table 3.4: Employment and frequency

<table>
<thead>
<tr>
<th>Employment</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employed</td>
<td>180</td>
<td>36.0</td>
<td>36.0</td>
<td>36.0</td>
</tr>
<tr>
<td>Employed (on Lien/Study Leave)</td>
<td>112</td>
<td>22.4</td>
<td>22.4</td>
<td>58.4</td>
</tr>
<tr>
<td>Unemployed</td>
<td>141</td>
<td>28.2</td>
<td>28.2</td>
<td>86.6</td>
</tr>
<tr>
<td>Business</td>
<td>55</td>
<td>11.0</td>
<td>11.0</td>
<td>97.6</td>
</tr>
<tr>
<td>Retired</td>
<td>12</td>
<td>2.4</td>
<td>2.4</td>
<td>100.0</td>
</tr>
<tr>
<td>Total</td>
<td>500</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>
3.10 Statistical Analysis:

3.9.1 Chi-Square Test for Independence

This lesson explains how to conduct a chi-square test for independence. The test is applied when you have two categorical variables from a single population. It is used to determine whether there is a significant association between the two variables. For example, in an election survey, voters might be classified by gender (male or female) and voting preference (Democrat, Republican, or Independent). We could use a chi-square test for independence to determine whether gender is related to voting preference.

The sample problem at the end of the lesson considers this example.

**When to Use Chi-Square Test for Independence**

The test procedure described in this lesson is appropriate when the following conditions are met:

- The sampling method is simple random sampling.
- The variables under study are each categorical.
- If sample data are displayed in a contingency table, the expected frequency count for each cell of the table is at least 5.
This approach consists of four steps: (1) state the hypotheses, (2) formulate an analysis plan, (3) analyze sample data, and (4) interpret results.

**State the Hypotheses**

Suppose that Variable A has $r$ levels, and Variable B has $c$ levels. The null hypothesis states that knowing the level of Variable A does not help you predict the level of Variable B. That is, the variables are independent.

- $H_0$: Variable A and Variable B are independent.
- $H_a$: Variable A and Variable B are not independent.

The alternative hypothesis is that knowing the level of Variable A can help you predict the level of Variable B.

Note: Support for the alternative hypothesis suggests that the variables are related; but the relationship is not necessarily causal, in the sense that one variable "causes" the other.

**Formulate an Analysis Plan**

The analysis plan describes how to use sample data to accept or reject the null hypothesis. The plan should specify the following elements.

- **Significance level.** Often, researchers choose significance levels equal to 0.01, 0.05, or 0.10; but any value between 0 and 1 can be used.
- **Test method.** Use the chi-square test for independence to determine whether there is a significant relationship between two categorical variables.

**3.10.2 Analyze Sample Data**

Using sample data, find the degrees of freedom, expected frequencies, test statistic, and the P-value associated with the test statistic. The approach described in this section is illustrated in the sample problem at the end of this lesson.

- **Degrees of freedom.** The degrees of freedom (DF) is equal to:
  \[ DF = (r - 1) \times (c - 1) \]
  where $r$ is the number of levels for one categorical variable, and $c$ is the number of levels for the other categorical variable.

- **Expected frequencies.** The expected frequency counts are computed separately for each level of one categorical variable at each level of the other categorical variable. Compute $r \times c$ expected frequencies, according to the following formula.
  \[ E_{i,c} = \frac{(n_r \times n_c)}{n} \]
where \( E_{r,c} \) is the expected frequency count for level \( r \) of Variable A and level \( c \) of Variable B, \( n_r \) is the total number of sample observations at level \( r \) of Variable A, \( n_c \) is the total number of sample observations at level \( c \) of Variable B, and \( n \) is the total sample size.

- **Test statistic.** The test statistic is a chi-square random variable (\( X^2 \)) defined by the following equation.

\[
X^2 = \sum \left[ \frac{(O_{r,c} - E_{r,c})^2}{E_{r,c}} \right]
\]

where \( O_{r,c} \) is the observed frequency count at level \( r \) of Variable A and level \( c \) of Variable B, and \( E_{r,c} \) is the expected frequency count at level \( r \) of Variable A and level \( c \) of Variable B.

- **P-value.** The P-value is the probability of observing a sample statistic as extreme as the test statistic. Since the test statistic is a chi-square, use the Chi-Square Distribution Calculator to assess the probability associated with the test statistic. Use the degrees of freedom computed above.

### 3.10.3 Measures of Association: Nominal data--Phi and Cramer’s V

- **Measures of Association** calculate the strength, and for ordinal variables the direction, of the relationship between two variables.

  - **PHI**: Used to measure the strength of the association between two variables, each of which has only two categories. (It applies to 2 X 2 nominal tables only).

  - **CRAMER’S V**: Used to measure the strength of the association between one nominal variable with either another nominal variable, or with an ordinal variable. Both of the variables can have more than 2 categories. (It applies to either nominal X nominal crosstabs, or ordinal X nominal crosstabs, with no restriction on the number of categories.)

#### Table 3.5 Interpreting the value of the Level of Association

<table>
<thead>
<tr>
<th>LEVEL OF ASSOCIATION</th>
<th>Verbal Description</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.00</td>
<td>No Relationship</td>
<td>Knowing the independent variable does not help in predicting the dependent variable.</td>
</tr>
<tr>
<td>------</td>
<td>-----------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>.00 to .15</td>
<td>Very Weak</td>
<td>Not generally acceptable</td>
</tr>
<tr>
<td>.15 to .20</td>
<td>Weak</td>
<td>Minimally acceptable</td>
</tr>
<tr>
<td>.20 to .25</td>
<td>Moderate</td>
<td>Acceptable</td>
</tr>
<tr>
<td>.25 to .30</td>
<td>Moderately Strong</td>
<td>Desirable</td>
</tr>
<tr>
<td>.30 to .35</td>
<td>Strong</td>
<td>Very Desirable</td>
</tr>
<tr>
<td>.35 to .40</td>
<td>Very Strong</td>
<td>Extremely Desirable</td>
</tr>
<tr>
<td>.40 to .50</td>
<td>Worrisomely Strong</td>
<td>Either an extremely good relationship or the two variables are measuring the same concept</td>
</tr>
<tr>
<td>.50 to .99</td>
<td>Redundant</td>
<td>The two variables are probably measuring the same concept.</td>
</tr>
<tr>
<td>1.00</td>
<td>Perfect Relationship</td>
<td>If we the know the independent variable, we can perfectly predict the dependent variable.</td>
</tr>
</tbody>
</table>

### 3.10.4 Interpret Results

If the sample findings are unlikely, given the null hypothesis, the researcher rejects the null hypothesis. Typically, this involves comparing the P-value to the significance level, and rejecting the null hypothesis when the P-value is less than the significance level.

### 3.11 Statistical Package for the Social Sciences (IBM SPSS version 20)

IBM SPSS(Statistical Package for the Social Sciences) Statistics continues to increase users' accessibility to advanced analytics through improved tools, output, and ease-of-use features. IBM SPSS Statistics 20.0 continues to deliver this trend with new mapping capabilities, improvements to existing procedures, and enhancements that increase the productivity of analysis through improved performance.
SPSS Statistics is a software package used for logical batched and non-batched statistical analysis. Long produced by SPSS Inc., it was acquired by IBM in 2009. The current versions (2015) are officially named IBM SPSS Statistics. Companion products in the same family are used for survey authoring and deployment (IBM SPSS Data Collection), data mining (IBM SPSS Modeler), text analytics, and collaboration and deployment (batch and automated scoring services). The software name originally stood for Statistical Package for the Social Sciences (SPSS), reflecting the original market, although the software is now popular in other fields as well, including the health sciences and marketing.

Statistics included in the base software:

- **Descriptive statistics:** Cross tabulation, Frequencies, Descriptive, Explore, Descriptive Ratio Statistics
- **Prediction for numerical outcomes:** Linear regression
- **Prediction for identifying groups:** Factor analysis, cluster analysis (two-step, K-means, hierarchical), Discriminate
- **Bivariatestatistics:** Means, T Test, Anova, Correlation (Bivariate, Partial, Distance), Nonparametric tests

### 3.12 Graphical and diagrammatic representation of data

- **Pie charts** are used to show frequency as well as percentile breakdown of data whenever required. Pie charts are easy to make, easy to read, and very popular. They are used to represent categorical data or values of variables. They are basically circles that are divided into segments or categories which reflect the proportion of the variables in relation to the whole. Percentages are used to compare the segments, with the whole being equal to 100%. To make a pie chart, draw a circle with a protractor. Then, convert the measures of the variables into percentages, and divide the circle accordingly. It is best to order the segments clockwise from biggest to smallest, so that the pie chart looks neat and the variable are easy to compare.

- **Line graphs** are simple to create and easy to understand. They organize and present data in a clear manner and show relationships between the data. The line graphs display a change in direction, while bar graphs display a change in magnitude.
• **Bar diagrams** used to represent response of the respondent. Bar graphs are used to present and compare data. There are two main types of bar graphs: horizontal and vertical. They are easy to understand, because they consist of rectangular bars that differ in height or length according to their value or frequency. These types of graphs serve the same purpose as line graphs: they represent time series data. However, bar graphs display a change in magnitude, and not in direction like line graphs.