

## CHAPTER 3

---

---

# *Research Methodology*

---

---

### **3.1 Introduction**

Research in common parlance refers to a search for knowledge. Research can also be defined as a scientific and systematic search for pertinent information on a specific topic (Kothari, 2002). The Advanced Learner's Dictionary of Current English lays down the meaning of research as "... a careful investigation or inquiry especially through search for new facts in any branch of knowledge." Redman and Mory define research as a "systematized effort to gain new knowledge." According to Clifford Woody research comprises defining and redefining problems, formulating hypothesis or suggested solutions; collecting, organizing and evaluating data; making deductions and reaching conclusions; and at last carefully testing the conclusions to determine whether they fit the formulating hypothesis. It is a sort of formal training which enables one to understand the new developments in one's field in a better way. Methodology is thus a guideline system for solving a problem, with specific components such as phases, tasks, methods, techniques and tools.

This section discusses the methodological foundations of the empirical elements of this study. It focuses on obtaining a more detailed understanding of the information seeking pattern of scientific communities to help their information needs and factors that impinge on that. This chapter includes the information regarding the sources of data, sampling procedure, tools for collection of data, construction and description of the questionnaire and data analysis.

The theory underlying any empirical research can take two forms: deductive or inductive. Using deductive theory research is conducted with reference to hypotheses (Bryman, 2008) and ideas are tested against observable empirical evidence (Neuman, 2011). Before starting a research study, it is further necessary to identify both the “meaning” and “objective” of that study (Crotty, 1998; Bryman, 2008).

Quantitative research is the “... *testing (of) objective theories by examining the relationship between variables*” (Creswell, 2009). These variables are measured in order to obtain data that can be analyzed with statistical tests.

Qualitative research is the exploring and understanding of the meaning individuals or groups ascribe to a problem (Creswell, 2009). It relies on the researcher interpreting the data, and constructing theory from initial research questions. Here researchers do not measure data; rather they look for relationships between elements of the data.

### **3.2 Population/ Size or Area of the Study**

The present study has attempted to understand the information needs and seeking behaviour of the scientific communities of Assam. The area of the study selected for the research is “Assam” state of Northeast India; which four universities (two central universities and state universities) are located, These are:

- Assam University, Silchar (Central University);
- Gauhati University, Guwahati (State University);
- Dibrugarh University, Dibrugarh (State University); and
- Tezpur University, Tezpur (Central University).

The population of the study further consists of university library users of these universities. The sample of the study was described by using random sampling technique and it was decided to take 600 samples in total (150 from each university) for the present study. The details of sampling have been given in later part of this chapter.

### ***3.2.1. Identifying the Scientific Community***

On the basis of the nature of research study, the following categories of university library users were taken into consideration which comprises:

- Under-graduate students;
- Post-graduate students;
- Research scholars; and
- Faculty members.

## **3.3 Research Method Adopted**

For this study, “Survey Method” of research was used to collect data from the scientific community library users of central and state universities of Assam; which consists of Assam University, Silchar; Gauhati University, Guwahati; Dibrugarh University, Dibrugarh and Tezpur University, Tezpur. During the survey, respondents from various science departments were surveyed for the collection of data by using questionnaire from four universities under the study. These departments consist of Computer Science, Geography, Engineering, Information Technology, Pharmaceutical Science, Earth Science, Botany, Zoology, Chemistry, Physics, Mathematics, Biotechnology and Petroleum Science.

### 3.4 Sampling Procedure for Survey

While surveying the universities and collecting data from the individual scientists various techniques have been adopted, which has described below.

“Stratified Random Sampling” technique was used in the case of data collection through questionnaire from the individual scientists in different stages. To understand the user information seeking behaviour, users have been divided into three main categories: (a) Student, (b) Faculty Member and (c) Research Scholar.

In order to cover respondents from each university, equal numbers of questionnaire have been distribution to them as per sampling procedure, which has shown in Table: 3.1. Altogether, 600 (six hundred) questionnaires were distributed to the library users. Out of 600 respondents, 150 questionnaires were distributed to the respondents of each university. The respondents have been divided into three categories, viz. UG/ PG students, research scholars and faculty members and are distributed as per Table: 3.1.

Moreover, “Purposive Sampling” is used to collect data from librarian of those universities to know specifically about the library facilities and services. Thus, 4 (0.7%) questionnaires were also distributed to each librarian of those universities.

**Table: 3.1 Sample Design for the Present Study**

University	Faculty			Research Scholar			Student	Librarian	Total
	Profe ssor	Asso. Prof.	Asst. Prof	JRF	SRF	M.Phil / Ph. D			
AU	10	15	25	10	10	30	50	1	151
GU	10	15	25	10	10	30	50	1	151
TU	10	15	25	10	10	30	50	1	151
DU	10	15	25	10	10	30	50	1	151
<b>Total</b>	<b>40</b>	<b>60</b>	<b>100</b>	<b>40</b>	<b>40</b>	<b>120</b>	<b>200</b>	<b>4</b>	<b>604</b>

### **3.5 Tools for the Collection of Data**

The tools of data collection translate the research objectives into specific questions/ items, the responses to which will provide the data required to achieve the research objectives. In order to achieve this purpose, each question/item must convey to the respondent the idea or group of ideas required by the research objectives, and each item must obtain a response which can be analyzed for fulfilling the research objectives. The problem of research is not solved unless a proper tool is selected and used for data collection. Data required to clarify all the ideas in the mind of the researcher are to be collected by the researcher properly (Bhattacharjee, Bhattacharjee & Sinha, 2013, p. 18). The following tools were used for collection of data pertaining to the present study:

- Questionnaire; which is subsequently followed by
  - Interview-schedule; and
  - Observation-method.

#### ***3.5.1. Design and Administration of Questionnaire***

A questionnaire is a research instrument consisting of a series of questions and other prompts for the purpose of gathering information from respondents. Questionnaires have advantages over some other types of surveys as it do not require as much effort from the researcher as verbal or telephone surveys, and often have standardized answers that make it simple to compile data. Questionnaires are also sharply limited by the fact that respondents must be able to read the questions and respond to them properly.

To collect the information pertaining to the information needs of the university library users of Assam a “questionnaire” was designed based on previous studies with some modifications as per the requirement of the study. At last the final questionnaires were sent to library users of various universities of Assam along with a request letter and self-addressed stamped envelope. The respondents from whom the questionnaires were not received within one month a gentle reminder was given and they were pursued for sending their responses at the earliest. The library users within the reach were visited personally and distributed questionnaire to respective respondents and passed on necessary instructions related to the filling up of the questionnaire. It was assured that the information given by them would be used for the research purpose only and would be kept confidential.

#### **3.5.1.1 Design/ Description of the Questionnaire used**

Two different types of questionnaire have been designed for the library users, which comprises of the questionnaire of “students/ research scholars” (Appendix-A) and “faculty members” (Appendix-B). Both types of questionnaire were having five sections as per details given below:

*Section A:* Consists of “Personal Identification” which comprises of eight questions;

*Section B:* Consists of “Library Visit and Library Usage Pattern” which comprises of nine questions;

*Section C:* Consists of “Information Searching Pattern” which comprises of nine questions;

*Section D:* Consists of “Internet Literacy and E-resource Search Strategies” which comprises of eleven questions;

*Section E:* Consists of “Problem and Suggestion” which comprises of three questions.

Further, separate questionnaire (Appendix-C) was designed for librarian to elicit information pertaining to the library services and library management issues. The questionnaire librarian has also been divided by dividing into five sections as enumerated below:

*Section A:* Consists of “Personal Details” which comprises of six questions;

*Section B:* Consists of “Library Collection Details” which comprises of one question;

*Section C:* Consists of “Status of Technical House Keeping Operation” which comprises four questions;

*Section D:* Consists of “Status of Library Automation/ Computerization” which comprises of seven questions;

*Section E:* Consists of “Status of Networking/ Resource Sharing Facility of the Library” which comprises of five questions.

*Section F:* Consists of “Status of Human Resource Strength in Library” which comprises of one question.

*Section G:* Consists of “Library Usage Statistics on Daily Basis” which comprises of one question.

*Section H:* Consists of “Provision of User Orientation Program from the Library” which comprises of two questions.

*Section I:* Consists of “Opinion of Librarian Regarding Other Aspects” which comprises of four questions.

### **3.5.1.2. Use of Point Scales**

A three point scale of “not at all sufficient”; “partially sufficient” and “sufficient” have been adopted in some questions in order to identify how far library collections are sufficient to meet their requirement. Some questions on library visits, awareness of Internet, “Yes” or “No” type have also been incorporated. Further, in some questions multiple responses are allowed for respondents to know the usage pattern, satisfaction levels, etc.

#### **3.5.1.2.1. Value of the five point scales**

In the questionnaire of “students/ research scholars- questionnaire” and “faculty members- questionnaire” five point scales is used for “*not used*” to “*extremely use*”; where 1 stands for “not used” and 5 stands for “extremely use”. Similarly, five point scales is also used for “*not an issue/ problem*” to “*extreme problem*”; where 1 stands for “not an issue/ problem” and 5 stands for “extreme problem”.

### **3.5.1.3. Use of Open Ended Questionnaire**

In the questionnaire for “Librarian” few open-ended questionnaires were asked to know the future plan of library, list of e-resources statistic usage by library users, etc.

### **3.5.2. Personal Interview**

Interviews are particularly useful for getting the story behind a participant's experiences. The interviewer can pursue in-depth information around the topic. It may be useful as follow-up to certain respondents to questionnaires, e.g., to further investigate their responses (McNamara, 1999). Thus to have a face to face dialogue and to collect more reliable data, personal interviews were arranged with pre-intimation to the interviews.

There are three fundamental types of research interviews: structured, semi-structured and unstructured. Structured interviews are, essentially, verbally administered questionnaires, in which a list of predetermined questions is asked, with little or no variation and with no scope for follow-up questions to responses that warrant further elaboration (Gill *et al.*, 2008). Conversely, unstructured interviews do not reflect any preconceived theories or ideas and are performed with little or no organization (May, 1991). Semi-structured interviews consist of several key questions that help to define the areas to be explored, but also allows the interviewer or interviewee to diverge in order to pursue an idea or response in more detail (Britten, 1999).

In this study, semi-structure questionnaire was designed for taking interview of the scientific community of Assam by approaching to them personally in most of the cases. It has been done in a phased manner through different stages to cover each university. Further, in some cases self-addressed stamped questionnaires were sent personally or via post to the addresses of the scientists concerned and the duly filled in ones were collected later on through the post. Some of the respondents and librarian have also returned duly filled in questionnaire through e-mail.

### ***3.5.3. Personal Observation***

Observational research is a social research technique that involves the direct observation of phenomena in their natural setting. To study the information needs of library users personal observation method also being used. Observation methods are useful to researchers in a variety of ways. They provide researchers with ways to check for nonverbal expression of feelings, determine who interacts with whom, grasp how participants communicate with each other, and check for how much time is spent on various activities (Schmuck, 1997; Marshall & Rossman, 1995).

In this study, this method was also used to get actual picture about the facilities available at university and to know what strategies usually scientific communities feel while accessing information from library. Further, some personal information seeking patterns also were observed among the scientific community while they were searching information or documents.

## **3.6 Stages/ Time Period of Data Collection**

The questionnaire were distributed and collected from the scientists in four different stages through different modes during the period from **January, 2014 to September, 2014** in different geographical location; which is stated below:

*First Stage:* Assam University, Silchar (From January, 2014 to March, 2014)

*Second Stage:* Gauhati University, Guwahati (From April, 2014 to July, 2014)

*Third Stage:* Tezpur University, Tezpur (From June, 2014 to August, 2014)

*Fourth Stage:* Dibrugarh University, Dibrugarh (From July, 2014 to September, 2014)

### **3.7 Statistics Used for the Analysis of Data**

Statistics is concerned with the science of uncertainty and can help the scientist deal with these questions. Many classical methods (regression, hypothesis testing, parameter estimation, confidence intervals, etc.) of statistics developed over the last century are familiar to scientists and are widely used in many disciplines (Efron and Tibshirani, 1991).

The SAGE Dictionary of Statistics (2004) defines Statistics as “a characteristic such as a mean, standard deviation or any other measure which is applied to a sample of data. Applied to a population exactly the same characteristics are described as parameters of the population”.

Wegman (1988) defines computational statistics as a collection of techniques that have a strong “focus on the exploitation of computing in the creation of new statistical methodology.” The SAGE Dictionary of Statistics (2004) defines statistical package as “an integrated set of computer programs enabling a wide range of statistical analyses of data usually using a single spread sheet for the data. Typical examples include SPSS, Minitab, etc. Most statistical analyses are now conducted using such a package”.

#### ***3.7.1 Using SPSS software***

In this study various statistical analysis methods are used for the analysis of data by using SPSS software in some cases. SPSS is a computer program used for survey authoring and deployment, data mining, text analytics, statistical analysis, and collaboration and deployment. Statistics included in the base software:

- Descriptive statistics: Cross tabulation, Frequencies, Descriptive, Explore, Descriptive Ratio Statistics
- Bivariate statistics: Means, t-test, ANOVA, Correlation (bivariate, partial, distances), Nonparametric tests
- Prediction for numerical outcomes: Linear regression
- Prediction for identifying groups: Factor analysis, cluster analysis and discriminate.

The following statistical techniques were employed for this study to analyse and interpret the data:

- Simple percentage;
- Mean, Median, Mode;
- Inter-Quartile Range;
- Standard Deviation;
- Chats;
- Cross-Table Analysis;
- Chi-square Test Analysis;
- Rating Scale Analysis;
- Correlation between variable Analysis; and
- Various other relevant statistics.

### **3.7.2 Simple Percentage Calculation**

Percentage analysis is the method to represent raw streams of data as a part in 100 (i.e. percent) for better understanding of collected data. The earliest method used in analysis is percentage methods. We can calculate data is the following way:

$$\text{Percentage} = \frac{(\text{No. of respondents for each response}) \times 100}{(\text{Total No. of respondents})}$$

### 3.7.3 Mean

Mean is the central tendency of a collection of numbers taken as the sum of the numbers divided by the size of the collection. On the other words, mean of a set of numbers is their average.

Let us consider we have sample space  $a_i$  (where  $a_i = \{a_1, a_2, a_3 \dots a_n\}$ ) and  $n =$  total Number of the sample. Then the arithmetic mean  $A$  is defined via the equation:

$$A := \frac{1}{n} \sum_{i=1}^n a_i.$$

### 3.7.3 Median

The median (Mdn) is the middle number in a sequence of numbers. To find the median, we have to organize each number in order or sequence; then number in the middle is the median value. If the total number of the size of the collection ( $n$ ) is an odd number, then the formula is given below:

$$\text{Mdn} = \left(\frac{n+1}{2}\right)^{\text{th}} \text{ term}$$

If the total number of the size of the collection ( $n$ ) is an even number, then the formula is given below:

$$\text{Mdn} = \left\{ \frac{\left(\frac{n}{2}\right)^{\text{th}} + \left(\frac{n+1}{2}\right)^{\text{th}}}{2} \right\} \text{ term}$$

### 3.7.4 Mode

The mode is the number that occurs most often within a set of numbers.

### ***3.7.5 Inter-Quartile Range***

The interquartile range (IQR) is a measure of statistical dispersion. It is equal to the difference between the upper and lower quartiles. In other words, the IQR is the 1st quartile subtracted from the 3rd quartile.

Quartiles divide a rank-ordered data set into four equal parts. The values that divide each part are called the first, second, and third quartiles; and they are denoted by Q1, Q2, and Q3, respectively.

Q1 = the "middle" value in the first half of the rank-ordered data set.

Q2 = the median value in the set.

Q3 = the "middle" value in the second half of the rank-ordered data set.

$$***IQR = Q3 - Q1***$$

The IQR is a measure of dispersion. It shows whether the responses are clustered together or scattered across the range of possible responses.

### ***3.7.6 Dispersion***

Dispersion (variability) of data, i.e. how spread out data is, as measured by the variance and its square root, the standard deviation.

### ***3.7.7 Standard Deviation***

The standard deviation of a random variable, statistical population, data set, or probability distribution is the square root of its variance. It is algebraically simpler, though in practice less robust, than the average absolute deviation (Gauss, 1816; Walker, 1931).

$$\text{Mean Deviation} = x - \bar{x}$$

Where; x = the value of score;  $\bar{x}$  = mean of the score value

Standard Deviation;  $\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$

Where;

$\sum(x - \bar{x})^2$  = sum of the squares of mean deviations

### **3.7.8 Charts**

A chart, also called a graph, is a graphical representation of data, in which "the data is represented by symbols, such as bars in a bar chart, lines in a line chart, or slices in a pie chart". Charts are often used to ease understanding of large quantities of data and the relationships between parts of the data. Charts can usually be read more quickly than the raw data that they are produced from. Chart helps to communicate the data easily to the viewer without consuming much time. It is done with more care and well planned before representing the data in a pictorial form.

A bar chart is a chart with rectangular bars with lengths proportional to the values that they represent. The bars can be plotted vertically or horizontally. In a grouped bar chart, for each categorical group there are two or more bars. These bars are color-coded to represent a particular grouping; whereas, a pie chart shows percentage values as a slice of a pie. (<https://en.wikipedia.org/wiki/Chart>)

### **3.7.9 Cross-Table**

Cross-tabulation analysis, also known as contingency table analysis, is most often used to analyze categorical (nominal measurement scale) data. A cross-tabulation is a two (or more) dimensional table that records the number (frequency) of respondents that have the specific characteristics described in the cells of the table. Cross-tabulation tables

provide a wealth of information about the relationship between the variables.  
([www.qualtrics.com/wp-content/uploads/2013/05/Cross-Tabulation-Theory.pdf](http://www.qualtrics.com/wp-content/uploads/2013/05/Cross-Tabulation-Theory.pdf))

### **3.7.10 Chi-square Test**

The Chi-square test is an important test amongst the several tests of significance developed by statisticians. The Chi-square test was used to compare an observed group of frequencies with an expected group of frequencies. This led to deduce the expected frequencies from the null hypothesis. The level of significance was set at 0.001 or 0.05. The statistical analysis of the data of the present study was done by statistical package. Chi-square, symbolically written as  $\chi^2$  is a statistical measure used in the context of sampling analysis for comparing a variance to a theoretical variance.

It involves the simple two-step operation for calculating  $\chi^2$  for the general case where there is one dimension of categorization. As we are measuring through this procedure is the degree to which an observed pattern of frequencies differs, overall, from a mean chance expected values pattern.

$$\chi^2 = \left\{ \frac{(\text{Observed Frequency} - \text{Expected Frequency})^2}{\text{Expected Frequency}} \right\}$$
$$= (O_f - E_f)^2 / E_f$$

Where,  $O_f$  = Observed frequency,  $E_f$  = Expected frequency

### **3.7.11 Rating Scale Analysis**

A “Likert scale” is the sum of responses to several Likert items. These items are usually displayed with a visual aid, such as a series of radio buttons or a horizontal bar

representing a simple scale. A “Likert item” is a statement that the respondent is asked to evaluate in a survey. When responding to a Likert questionnaire item, respondents specify their level of agreement or disagreement on a symmetric agree-disagree scale for a series of statements. Thus, the range captures the intensity of their feelings for a given item. The five response categories are often believed to represent an Interval level of measurement. But this can only be the case if the intervals between the scale points correspond to empirical observations in a metric sense (Reips & Funke, 2008).

#### 3.7.11.1 Use of 3 or 5-point scales

In certain sections of the analysis, point scales have been used to convert the collected data into scores for making the analysis more empirical. The calculated scores have been drawn to ascertain the variations amongst the variables under considerations by finding out mean deviations, standard deviations and IQR.

To calculate the Likert scale rating, we can use median and inter-quartile range (IQR) of each item. The median is a measure of central tendency whereas IQR is a measure of dispersion; it shows whether the responses are clustered together or scattered across the range of possible responses. A relatively small IQR, is an indication of consensus. By contrast, larger IQRs might suggest that opinion is polarised, i.e., respondents tend to hold strong opinions either for or against that topic.

#### ***3.7.12 Correlation***

This is used to verify strong or weak relationship between two variables. This means that changes in one variable lead to correlated with changes in the second variable. This also used to verify relationship which may be directly or indirectly i.e. positive or negative between those two variables. Thus, it is essential for a scientific

study to process data and analyzes them in accordance with the method laid down for the purpose in order to arrive at the possible result.

### **3.8 Coding**

All researchers collect data for their study, in order to make sense of the data, it must be analyzed. Analysis begins with the labelling of data as to its source, how it was collected, the information it contains, etc. Codes allow the researcher to reduce large quantities of information into a form that can be more easily handled, especially by computer programs (<http://web.csulb.edu/~msaintg/ppa696/696codes.htm>). Specific codes have been deployed at the time of data generation works to understand classification and grouping of categories of Scientists and their characteristics. These codes are used while making data entry in SPSS software.

### **3.9 Citation Style**

A citation is a reference to a published or unpublished source (not always the original source). More precisely, a citation is an abbreviated alphanumeric expression embedded in the body of an intellectual work that denotes an entry in the bibliographic references section of the work for the purpose of acknowledging the relevance of the works of others to the topic of discussion at the spot where the citation appears. (<https://en.wikipedia.org/wiki/Citation>). Throughout the course of the study, it is decided to use APA, (6<sup>th</sup> edition) Style while making text and reference citations.

### **3.10 Conclusion**

Previously many studies had been conducted on information needs and information seeking behaviour among various groups and different methods and techniques had been used by previous studies, while collecting information from the target users. But, yet there was an untouched area, where this study has been carried out among the scientific communities of Assam. The methodology adopted in the present study was also found to be accepted in many such studies as literature review shows us. The above discussed approaches were found to be appropriate, feasible, flexible and adaptable for achieving the goal or objectives of the study.