

CHAPTER-III

RESEARCH METHODOLOGY

The sources and methods of collecting data, information, problems and analysis are discussed in this chapter. For the purpose the previous research studies were reviewed and secondary data was collected to enhance the understanding of the area and the development process. Actually, it was a process to make an assessment of the issues raised in banking habits of rural people and their status in progress of banking sector. Based on the understanding from literature review, the data needs were established and options for data gathering techniques were considered. Questionnaire survey was adopted to gather information regarding the problem of the study.

The present chapter is divided into two parts. In the first part, development of research model or plan of study, which highlights that why this study is planned or what is targeted to achieve and what need to be done. It represents the whole picture of the study. Second part is the application of research plain in the field. This is the main part of the chapter. It also represents the tools and techniques used in implementing the research model.

Selecting a topic was to streamline the whole idea for doing research to pursue the objectives of the study. The availing literature was reviewed before selecting the topic. It was to access what is needed to know more about the research issue? No doubt sometimes the researchers simply follow their curiosity and their drive to know. They become interested in a particular topic and pursue it.

Sometimes they choose a topic simply because funds are available for that purpose. At other times some social problems become a pressing issue and the social scientist want to gather data that will help people better understand and perhaps help solve it. But none of the reasons was behind this research. The topic was selected as this area is grossly under researched.

Most of the research was confined to study the growth of banking sector in rural areas, comparative performance of banking sectors in urban and rural set up or performance of cooperative and credit societies in agricultural growth or rural development. But the underlying relationship of banking habits of rural people is least understood. So by reviewing the available literature on different issues in the development of banking habits like customer relationship, door canvassing, motivation, provision of online services.

It became clearer that the services offered by the banking institution would attract the customers. Therefore, It is necessary to review the facilities or services offered by the banking institution and banking habits of the rural people. Moreover, the rural people intent to save their savings in conventional methods like keep the money in pot or in trunk box or with relatives rather than in banking because where the bank will be closed suddenly without intimation. Such tendencies do not allow them to go for banking.

Though the technology and economy is developing, still in rural areas many people found to be out of banking services even though bank branches are operating in those areas. Therefore all these issues provide help in selecting the topic for this thesis.

OBJECTIVES

The present study is taken up with the following broad objectives,

- To study the profile of customers and banking environment prevailing in the Thiruvallur District, TamilNadu.
- To find out the level of awareness and to identify the problems in availing banking services among the Customers in rural areas.
- To analyze the factors motivating to use banking services among the customers in Thiruvallur district, TamilNadu.
- To find out the relationship between the demographical variables and the level of awareness among the Respondents.
- To offer suggestions to minimize the gaps in the practice of banking habits and the services offered to the customers in Thiruvallur district, TamilNadu.

Hypothesis

During the process of literature review, many predictions were highlighted. It was not easy to make it clear that what results can be expected options from the study. Keeping in view all this, at this stage it was necessary to formulate hypothesis. The hypothesis is a statement of what is expected to find according to predictions from theory. It predicts a relationship between or among different variables. In this study, it is assumed that increasing bank branches and extension of banking services would attract rural people to have flexibility in banking operations and made them to be the customer.

Hypothesis 1:

Hypothesis 1 is used to find out whether the banking services are gender biased or not. With the help of ANOVA table it is has been calculated.

Hypothesis 2:

In hypothesis 2 the calculation is done to find out whether the aged Persons are aware of banking services well than younger persons.

Hypothesis 3:

Hypothesis 3 is used to know whether educational standard improves better access to banking services.

Research Design

Descriptive method is being used up in this research design. The research method for a study has different stages and strategies to work on the issue. The ways by which the data is collected are called research tools. These are important though different in different studies. Based on the nature of the study the most appropriate methods are used that can answer the questions related to the objectives and hypotheses of the study. In this study, the research tools are a combination of different techniques; primary data and secondary data have been used and analyzed using statistical techniques.

Sample Method

The most frequently used approaches for statistical analysis of survey data and estimation of input/output relationship using regression analysis are based on a simple random sample for data collection. Unfortunately simple random sampling is rarely practical in large scale surveys. A multi-stage procedure can greatly simplify construction of a sample frame, while ensuring that the survey population is adequately covered.

However, the conventional formulae yield biased population parameter estimates for samples that are not randomly selected. Alternative, and often far more complex formula are required to reduce or eliminate biases.

The sampling framework developed for the current study is multi-stage sampling method. In the first stage, the study area is selected under lottery ticket method, second stage, the villages in the study area are chosen on the basis of area, population and banking branches, thirdly respondents are selected based on account holders in the study area.

Sample Size

The primary data for the study is collected from the pre-tested interview schedule. The samples are chosen by the method of simple random sampling without sample replacement. 406 samples have chosen from the study area and collected samples were scrutinized to check the validity and found 47 questionnaires as defective and they are removed finally 359 sample was taken for the purpose of the study.

The sampling procedure is narrated below;

The formula for sample size is as follows:

$$n = (ZS/E)^2$$

Where,

Z= Expected confident level of 95 percent → 1.96

S= Population Standard Deviation (which is calculated from the pilot survey of 50 questions)

E= Acceptance level of sampling errors

$$n = ((1.96) (0.52)/0.05)^2$$

$$= 415.5$$

Then,

$$n^* = (N-n/N)n$$

$$n^* = ((20590-415.5)/20590) * 415.5 \left(\left(\frac{3049-308.4098}{3049} \right) 308.4096 \right)^2$$

$$= (0.979) (415.5)$$

$$= 406.7$$

Sample Size = 406

Research area

The research or Sample areas are Tiruttani, Gummidipoondi, Ponneri, Madavaram, Ambattur, Poonamalle, Pallipattu, Uthukkootai, and Thiruvallur.

All these areas are very important places of the Thiruvallur district of TamilNadu, India. These areas though comes under rural area have a wide business and educational surroundings.

Out of 32 districts in TamilNadu, Thiruvallur district is the one of the important and preferred district in TamilNadu, India.

Basically the customers are from agriculture sector and a few may include the self help groups, small and medium size business holders. The activities of the businesses are closely related to the farm sector and allied farm sector.

The customers include both male and female, belongs to different age groups, educational qualifications, occupations, levels of income, type of living place, purpose of visiting a bank, tenure of account holding with a particular bank etc. All the customers belong to rural areas and some way or the other related to farming sector.

Very few customers are into retail and self skilled businesses. These customers are also visiting a bank for farming loans and to avail themselves of the other services offered by a bank.

Pilot study and Reliability test

A pilot survey was carried out in the field situation to assess the effectiveness and consistency of the questionnaire. The pilot survey was done in the different villages of the study area. The sample was selected by using multi-stage random sampling. Many observations were raised on different questions after the pilot survey. Some questions were modified after the pilot surveys. And at the final stage many questions were found irrelevant and therefore dropped from questionnaire. This way it was tried to ensure a complete and uniform data for the study. The questionnaire used in this study was developed in English, but interviews were conducted in vernacular language (Tamil).

In order to validate the questionnaire for data collection, the data gathered from pilot survey was used and reliability test had been administered for the same. The reliability coefficient of the questionnaire was computed using Corn Bach's Alpha and the value is 0.82, which shows that the instrument is reliable.

Reliability Statistics

Case Processing Summary

		N	%
Cases	Valid	360	99.7
	Excluded ^a	1	.3
	Total	359	100.0

Reliability Statistics

Cronbach's Alpha	Number of Items
.820	172

Data collection

The study used both primary and secondary sources of data. The primary data has been collected from the structured questionnaire. To ensure the quality of data, the researcher personally conducted the survey. The questionnaire consists of the details regarding socioeconomic background, Awareness on banking services, usage of functional services and barriers acquire them, reasons for poor banking habits, importance of the banking services, satisfaction on the banking services and motivating factors in the availing of banking services.

The secondary data was collected for the study from journals, magazines, reports, working papers and research dissertations and websites.

Tools and Techniques used for Analysis

For the purpose of analyzing the data, appropriate statistical techniques are used in consultation with the research supervisor. In the analysis part of the present thesis was made by using the various parametric and non-parametric statistical tests namely, Percentage Analysis, Chi-square analysis with hypothesis, Fried man two way ANOVA with null hypothesis, Inferential statistics, Factor analysis, Rank analysis, F-test, Kruskal-Wallis one-way analysis of variance by ranks and Principal component analysis.

Chi-square analysis

Chi-squared test is used as a statistical tool in this project. Also it invokes no assumption about the form of original distribution from which the observations are made. In this method we test if two attributes considered are dependent or not.

Formula

$$\psi^2 = \sum \left(\frac{(O-E)^2}{E} \right)$$

Where O – Observed Frequency

E – Expected Frequency

Kruskal-Wallis one way analysis of variance:

Kruskal Wallis one-way analysis of variance by ranks (named after William Kruskal and W Allen Wallis) is a non-parametric method for testing equality of population medians among groups. It is identical to a one way analysis of variance with the data replaced by their ranks. It is an extension of the Mann Whitney U test to 3 or more groups.

Since it is a non-parametric method, the Kruskal Wallis test does not assume an ordinary population, unlike the analogous one way analysis of variance. But, the test does assume an identically shaped and scaled distribution for each group and except for any difference in medians.

Method

1. Rank all data from all groups together i.e., rank the data from 1 to N ignoring group membership. Allocate any tied values the average of the ranks they would have received had they not been tied.

2. The test statistic is given by
$$K = (N-1) \frac{\sum_{i=1}^g n_i (\bar{r}_i - \bar{r})^2}{\sum_{i=1}^g \sum_{j=1}^{n_i} (r_{ij} - \bar{r})^2}$$
, where

n_i is the number of observations in group i

r_{ij} is the rank (among all observations) of observation j from group i

N is the total number of observations across all groups $\bar{r}_i = \frac{\sum_{j=1}^{n_i} r_{ij}}{n_i}$,

$\bar{r} = (N+1)/2$ is the average of all the r_{ij} .

Notice that the denominator of the expression for K is exactly $(N-1)N(N+1)/12$.

So,
$$K = \frac{12}{N(N+1)} \sum_{i=1}^g n_i (\bar{r}_i - \bar{r})^2$$

3. A correction for ties can be made by dividing K by $1 - \frac{\sum_{i=1}^G (t_i^3 - t_i)}{N^3 - N}$,

where G is the number of groupings of different tied ranks and t_i is the number of tied values within group i that are tied at a particular value. This is the correction usually makes little difference in the value of K unless there are a large number of ties.

4. Finally, the p-value is approximated by $\Pr (X_{g-1}^2 \geq K)$. If some n_i 's are small (i.e., less than 5) the probability distribution of K can be quite different from this chi-square distribution. When the table of the chi-square probability distribution is available, the important value of chi-square, $X_{\alpha;g-1}^2$, can be derived by entering the table at $g - 1$ degrees of freedom and looking under the desired significance level. Null hypothesis of equal population medians would then be rejected if $K \geq X_{\alpha;g-1}^2$. Appropriate multiple comparisons would then be performed on the group medians.

Principal component analysis:

In principal component analysis there are two objectives: data reduction and interpretation of the components.

Let X is a multivariate normal random vector with mean μ and covariance matrix Σ . We want linear combinations of the p random variables that represent the rotated coordinate system in the directions that maximize the variability of the data set.

Let

$$\begin{aligned}
Y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p = a_1^T X \\
Y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p = a_2^T X \\
&\vdots \\
Y_p &= a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pp}x_p = a_p^T X
\end{aligned}$$

In practice we want to maximize the variance with respect to a_1 so that we have a unique solution. For a non null solution to exist we must satisfy the equation $|\Sigma - \lambda_1 I| = 0$. Since X is a random vector with a covariance matrix Σ it has characteristic roots $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$. Notice $(\Sigma - \lambda_1 I)a_1 = 0$ implies a_1 is a characteristic vector of Σ corresponding to the characteristic root λ_1 . So it is convenient to normalize the coefficients vectors. The contribution of the first principal component Y_1 to the total variation is given by

$$\frac{\text{Explained}}{\text{Total Variance}} = \frac{\lambda_1}{\sum_1^p \sigma_{ii} = \sum_1^p \lambda_i} = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \dots + \lambda_p}$$

Analysis of variance (ANOVA)

Analysis of variance (ANOVA) is a collection of statistical models and their associated procedures and in which the observed variance is divided into components due to different explanatory variables. We show the model for a simplified ANOVA with one type of treatment at different levels.

$$SS_{\text{Total}} = SS_{\text{Error}} + SS_{\text{Treatments}}$$

The number of degrees of freedom (abbreviated *df*) can be partitioned in a similar way and specifies the chi-square distribution which describes the associated sums of squares.

$$df_{\text{Total}} = df_{\text{Error}} + df_{\text{Treatments}}$$

The F-test

The F-test is used for comparisons of the components of the total deviation. For instance, in one-way or single-factor ANOVA statistical significance is tested for by comparing the F test statistic

$$F = \frac{\text{variance of the group means}}{\text{mean of the within group variances}}$$

$$F^* = \frac{\text{MSTR}}{\text{MSE}}$$

Where
$$\text{MSTR} = \frac{\text{SSTR}}{I-1}$$

I = number of treatments

And
$$\text{MSTR} = \frac{\text{SSE}}{n_T - I}$$

n_T = total number of cases.