CHAPTER 3
RESEARCH METHODOLOGY

This chapter focuses on the definition of Research, Research Design, Sampling design, sample size determination, period of study, reliability and validity test, structure of questionnaire, hypotheses of the study and data analysis plan.

3.1 Research

Research is an organised set of activities which comprises defining and redefining problems, formulating hypothesis or suggest solutions by collecting, organising and analysing data, making deductions and reaching a conclusion.

3.2 Research Design

‘A traditional research design is a blueprint or detailed plan for how a research study is to be completed – operationalizing variables so they can be measured, selecting a sample of interest to study, collecting data to be used as a basis for testing hypotheses, and analysing the results’ - Thyer (1993). The design ensures the conversion of the research question and the stated assumption/hypothesis into operational variables which can be measured and specify the method/process that would be followed to complete the above task, as efficiently and economically as possible. The research design followed is structured and more formal descriptive research design.

3.3 Source of data

The validity of any research depends on the systematic method of collecting the data, and analysing the same in appropriate order. In the present study both primary and secondary
data were collected and descriptive/quantitative research was used. The primary method of data collection was by a well structured questionnaire with a five point Likert’s scale. Secondary Data was gathered from various sources such as journals, magazines, newspapers, published sources and websites. Before the preparation of the study instrument, various literature related to this study conducted in India and Chennai were analysed and on the basis of the Oral Exploratory Survey the questionnaire is framed on a 5 point Likert scale with (1) - Highly Satisfied and (5) - Highly Dissatisfied. The data was collected at the entrance of the store after they complete their purchase during the evenings and weekends from one adult member of a household who had shopped.

3.4 Population and sample size determination

The study is limited only to Chennai city and the population of the Chennai city is 46,81,087 as per census 2011. To arrive at the sample size the mean household size of Chennai is considered. The mean size of household in Chennai city is 3.8 as per the reports of NFHS-3\textsuperscript{12}. Therefore the number of households in Chennai city is 12,31,865 (= 46,81,087 / 3.8) which is the population size for current study.

The sample size is calculated using the formulae

\[
Ss = \frac{Z^2 \times (p) \times (1-p)}{c^2}
\]

Where:

\( Z = 1.96 \) for 95% confidence level
\( p = .5 \) used for sample size needed
\( c = \) confidence interval (+-.05)
Based on the population size of Chennai, the sample size calculated on 95% confidence level and ±5% confidence interval is 384. For the current study it is decided to collect data from about 650 respondents.

### 3.5 Sampling Technique

Sampling techniques are used to select the respondents from the population and these techniques help to minimise cost while maximising generalisability. The respondents were selected on simple random sampling technique. Simple random sampling is a probability sampling and gives fair and unbiased results.

### 3.6 Selection of Location

Chennai city is one of the four major metros of India and it is the gateway to traditional and cultural rich South India. Unlike the rest of the metros, Chennaeites are more complex in their buying behaviour. Chennai has a general demographic pattern of North Chennai mostly of Industrial labours; Central Chennai predominantly covers the lower middle class to medium middle class and bureaucrats and the South Chennai predominantly the high income floating IT population. Hence it is ideal to select Chennai city for the study.
For the study purpose, Chennai city is segmented into three stratums. North Chennai mainly dominated by the industrial working class people. Central Chennai, consisting of bureaucrats, middle class and upper middle class people. The last segment, South Chennai consisting of floating population of IT sector and elite group of Chennaites.

### 3.6.1 Segmentation of location

**North Chennai** – North Chennai is composed of the following: Thiruvottiyur, Dr. Radhakrishnan Nagar, Perambur, Kolathur, Ambattur, Thiru. Vi. Ka. Nagar and Royapuram.

**Central Chennai** – Central Chennai is composed of the following: Park town, Purasawalkam, Egmore, Anna Nagar, Thousand Lights, Chepauk, Villivakkam and Harbour

**South Chennai** – Chennai is composed of the following: Theagaraya Nagar, Saidapet, Virugambakkam, Mylapore, Velachery, Alandur and Tambaram. (see appendix B)

### 3.7 Selection of Departmental Stores

For the purpose of the current study, leading departmental stores in Chennai were selected. Selection of the departmental stores was based on the market share (in terms of number of retail chains, their turnover and duration since its establishment in the Chennai city). Reliance with more than 112 branches, Spencer’s with more than 60 branches, Nilgiris with more than 30 branches, FoodWorld, More and other popular departmental stores specific to the locality were considered.

### 3.8 Pre-test

A pilot study with 60 respondents was tested for its reliability and validity and the final Questionnaire was framed on the basis of the alpha (Cronbach) values. Departmental
stores for the current study were selected from the leading stores like Reliance, Spencer’s, FoodWorld and Nilgiris across the three Chennai segments namely North, Central, and South. The reliability for the pilot study was statistically tested using Cronbach’s Alpha and it was found to be 0.892.

3.9 Reliability and Validity

3.9.1 Reliability results

Reliability as defined by ‘Joppe’, in the words of Golafshani (2003) is “The extent to which results are consistent over time and an accurate representation of the total population under study is referred to as reliability and if the results of a study can be reproduced under a similar methodology, then the research instrument is considered to be reliable”. Reliability of an instrument refers to the degree of consistency between multiple measurements of variables. It is extent to which an experiment tests or any measuring procedures yield, the same result on repeated attempts. Reliability was estimated through internal consistency method which is applied to measure the consistency among the variables in a summated scale. In the current study, the Cronbach’s Alpha co-efficient of reliability was found based on primary data. The Alpha co-efficient of reliability calculated for the 6 dimensions across the three Chennai segments and the values are listed in the Table 3.1. Most of the measures used in the instrument were already tested by the academicians and store managers.

Table 3.1 Segment-wise Reliability measure (Alpha) for the six dimensions

<table>
<thead>
<tr>
<th>Segments</th>
<th>Dimensions</th>
<th>No of items</th>
<th>Mean</th>
<th>Cronbach’s Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Chennai</td>
<td>Personal Care</td>
<td>4</td>
<td>2.16</td>
<td>0.713</td>
</tr>
<tr>
<td></td>
<td>Policy</td>
<td>11</td>
<td>2.81</td>
<td>0.719</td>
</tr>
<tr>
<td></td>
<td>Safety &amp; Security</td>
<td>2</td>
<td>2.57</td>
<td>0.739</td>
</tr>
<tr>
<td></td>
<td>Complaint Handling</td>
<td>2</td>
<td>2.42</td>
<td>0.840</td>
</tr>
<tr>
<td></td>
<td>Physical Aspects</td>
<td>6</td>
<td>2.34</td>
<td>0.709</td>
</tr>
<tr>
<td>Dimension</td>
<td>Chennai City</td>
<td>Reliability</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>----------------------------</td>
<td>--------------------</td>
<td>-------------</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>Central Chennai</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Care</td>
<td></td>
<td>4</td>
<td>2.16</td>
<td>0.713</td>
</tr>
<tr>
<td>Policy</td>
<td></td>
<td>11</td>
<td>2.84</td>
<td>0.738</td>
</tr>
<tr>
<td>Safety &amp; Security</td>
<td></td>
<td>2</td>
<td>2.48</td>
<td>0.694</td>
</tr>
<tr>
<td>Complaint Handling</td>
<td></td>
<td>2</td>
<td>2.12</td>
<td>0.707</td>
</tr>
<tr>
<td>Physical Aspects</td>
<td></td>
<td>6</td>
<td>2.32</td>
<td>0.701</td>
</tr>
<tr>
<td>Reliability</td>
<td></td>
<td>5</td>
<td>2.07</td>
<td>0.708</td>
</tr>
<tr>
<td>All variables</td>
<td></td>
<td>30</td>
<td>2.52</td>
<td>0.625</td>
</tr>
<tr>
<td>South Chennai</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Care</td>
<td></td>
<td>4</td>
<td>2.16</td>
<td>0.832</td>
</tr>
<tr>
<td>Policy</td>
<td></td>
<td>11</td>
<td>2.60</td>
<td>0.749</td>
</tr>
<tr>
<td>Safety &amp; Security</td>
<td></td>
<td>2</td>
<td>2.50</td>
<td>0.754</td>
</tr>
<tr>
<td>Complaint Handling</td>
<td></td>
<td>2</td>
<td>1.84</td>
<td>0.761</td>
</tr>
<tr>
<td>Physical Aspects</td>
<td></td>
<td>6</td>
<td>2.32</td>
<td>0.656</td>
</tr>
<tr>
<td>Reliability</td>
<td></td>
<td>5</td>
<td>2.16</td>
<td>0.721</td>
</tr>
<tr>
<td>All variables</td>
<td></td>
<td>30</td>
<td>2.37</td>
<td>0.726</td>
</tr>
<tr>
<td>Pan Chennai</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Care</td>
<td></td>
<td>4</td>
<td>2.23</td>
<td>0.795</td>
</tr>
<tr>
<td>Policy</td>
<td></td>
<td>11</td>
<td>2.66</td>
<td>0.749</td>
</tr>
<tr>
<td>Safety &amp; Security</td>
<td></td>
<td>2</td>
<td>2.52</td>
<td>0.673</td>
</tr>
<tr>
<td>Complaint Handling</td>
<td></td>
<td>2</td>
<td>2.04</td>
<td>0.769</td>
</tr>
<tr>
<td>Physical Aspects</td>
<td></td>
<td>6</td>
<td>2.54</td>
<td>0.693</td>
</tr>
<tr>
<td>Reliability</td>
<td></td>
<td>5</td>
<td>2.14</td>
<td>0.690</td>
</tr>
<tr>
<td>All variables</td>
<td></td>
<td>30</td>
<td>2.44</td>
<td>0.706</td>
</tr>
</tbody>
</table>

Source: Primary data

The instrument consist of 30 items measuring the six dimensions namely Personal Care, Policy, Safety & Security, Complaint Handling, Physical Aspects and Reliability. A 5-point Likert’s scale was used to get the subjects’ response.

It was observed from the table that the reliability estimates for the various dimensions at Pan Chennai city were: ‘Personal Care’ with 4 items had a value of 0.795, ‘Policy’ with 11 items had a value of 0.749, ‘Safety & Security’ with 2 items had a value of 0.673, ‘Complaint Handling’ with 2 items had a value 0.769, ‘Physical Aspects’ with 6 items had a value of 0.693, ‘Reliability’ with 5 items had a value of 0.690 and All variables with 30 items had a value of 0.706.
The reliability results were found acceptable. The convergent and discriminant validity test was performed on the data and found to be valid.

3.9.2 Validity results

To check the validity of the data, a Convergent validity was examined using an item assessing the overall customer satisfaction. The results (Table 3.2) showed a good correlation between the scale and this item (0.402; p < 0.001), confirming its convergent validity. The discriminant validity was examined using the item Repurchase intention. The negative correlation (-0.164; p<.001) conforms its discriminant validity.

Table 3.2 Validity test results

<table>
<thead>
<tr>
<th>Construct Name</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convergent Validity</td>
<td>Correlation 0.402 p&lt;.000</td>
</tr>
<tr>
<td></td>
<td>Overall customer Satisfaction</td>
</tr>
<tr>
<td>Discriminant validity</td>
<td>Correlation -.164 p&lt;.000</td>
</tr>
<tr>
<td></td>
<td>Repurchase intension</td>
</tr>
</tbody>
</table>

3.10 Statistical tools

Data analysis for the present study was made by using the various parametric and non-parametric statistical tests namely, descriptive analysis, t-test, ANOVA, Chi-square, f-test, Cluster analysis, correlation, and multivariate tools like Discriminant analysis and Regression analysis using IBM software SPSS v-16.

ANOVA is a collection of statistical models, and their associated procedures, in which the observed variance is portioned into components due to different explanatory variables. ANOVA allows for the study of a single factor or several factors, but will only measure one
variable. ANOVA is a fundamental technique which partition the total sum of squares into components related to the effects used in the model.

A t-test is a statistical hypothesis test in which the test statistic has a students’ t distribution if null hypothesis is true. This test is used in a situation where the research wants to compare two statistics. The basic assumptions for t-test are one random sampling, independent measurements, normal distribution and equal variance. The basic utility of a t-test is that it produces a straightforward easy to interpret results of significance.

Cluster Analysis is an advanced statistical technique used to identify relatively homogenous groups of cases (or factors) based on selected characteristics. The technique uses either distance or similarity measures to identify the groups. Distance or similarity measures are generated by the proximities between individuals based on multivariate data of the sample respondents. If the distance is low between any two individuals, they are said to belong to one group or cluster and whenever the distance between any two objects is maximum, the two entities are said to be apart from one another. Similarly, two individuals are said to belong to one cluster if the similarity between the two is the highest.

A number of clustering methods have been developed ever since the concept of clustering introduced in the data analysis. For the present study, two-step clustering analysis and the Log-likelihood distance measure have been used. The measure and the method used for the current study are widely used analysis for continuous data.

Discriminant Analysis is useful for situations where one wants to build a predictive model of group membership based on observed characteristics of each case. The procedure generates a discriminant function (or, for more than two groups, a set of discriminant functions) based on linear combinations of the predictor variables that provide the best discrimination between the groups. For the present study, discriminant analysis has been used to differentiate the three segments of Chennai (North, Central and South) using the 6 dimensions. The procedure
has a sequence of steps. The first step is the assessment of efficiency of the variables for discriminating the groups. This is done using Wilk’s Lambda. The second step involves the percentage efficiency of the classification and this is assessed using Eigen values. The next step is to obtain the standardized canonical discriminate function coefficients. This is useful when the variables are measured in different units. Forming the structure matrix is the next step and this matrix contains within-group correlations of each predictor variable with the canonical function. When the variables are measured in the same unit, un-standardized discriminate function may be used. In the next step the group centroids based on un-standardized discriminate coefficients are obtained. The final step is testing the efficiency of the discriminant function.

Regression analysis is used to test the overall relationship between the dependent variable and a set of independent variables. To check the goodness of fit the value of $R^2$ is interpreted. This measure is also called coefficient of determination of a regression equation and takes value between 0 and 1. It indicates the exploratory power of the regression model.

Variable Consistency Dominance-based Rough Set for decision-making: Rough Set Theory is the new upcoming decision make tool easy to analyse and interpret. Rough Set Theory is a mathematical approach to managing vague and uncertain data or problems related to information systems, indiscernibility relations and classification, attribute dependence and approximation accuracy, reduct and core attribute sets, and decision rules. It uses the algorithms like VC-DOMLEM to generate decision rules which are easy to understand.

Rough set theory (RS theory) has five advantages as expressed below:

(1) The RS theory does not require any preliminary or additional parameter about the data;
(2) They can work with missing values, switch among different reducts, and use less expensive or time to generate rules;
(3) They offer the ability to handle large amounts of both quantitative and qualitative data;
(4) They generate understandable decision rules and own stability; and

(5) They can model highly non-linear or discontinuous functional relationships provides a powerful method for characterizing complex, multidimensional patterns

### 3.11 Hypotheses

The following Null hypotheses (H₀) have been used for the present study to find the Customer relationship towards departmental stores of the Chennai respondents.

1. There is no significant difference between the respondents’ patronising various departmental stores and frequency of visits to the departmental store.

2. There is no significant difference between the respondents’ patronising various departmental stores and age group.

3. There is no significant difference between the respondents’ patronising various departmental stores and occupation.

4. There is no significant difference between the respondents’ patronising various departmental stores and annual family income.

5. There is no significant difference between the respondents’ patronising various departmental stores and number of household members.

6. There is no significant difference between the male and female respondents’ in overall customer satisfaction.

7. There is no significant difference between respondents’ overall satisfaction level and various departmental stores.

8. There is no significant difference between the dimension Personal care and various departmental stores.

9. There is no significant difference between the dimension Safety and security level in the store and various departmental stores.
10. There is no significant difference between the dimension store Reliability and various departmental stores.

11. There is no significant difference between the dimension store Policy and various departmental stores.

12. There is no significant difference between the dimension Complaint handling in stores and various departmental stores.

13. There is no significant difference between the dimension store’s Physical Aspects and various departmental stores.

14. There is no significant difference between respondents’ overall customer satisfaction and annual family income groups.

15. There is no significant difference between respondents’ overall customer satisfaction level and different occupation.

16. There is no significant difference between the relationship in each of the 6 dimensions and age groups.

17. There is no significant difference between the 6 dimensions and occupation.

18. There is no significant difference between the 6 dimensions and annual family income.

19. There is no significant difference between the 6 dimensions and educational qualification.

20. There is no significant difference between the respondents’ patronization level of departmental store and occupation.

21. There is no significant difference between the respondents’ patronization level of departmental store and annual family income.

22. There is no significant difference between the respondents’ patronization level and store distance.
23. There is no significant difference between the respondents’ level of willingness to change store and departmental stores.

24. There is no significant difference between the respondents’ level of willingness to change store and annual family income.

25. There is no significant difference between the respondents’ level of willingness to change store and educational qualification.

26. There is no significant difference between the respondents’ level of willingness to change store and occupation.

27. There is no significant difference between the respondents’ level of willingness to change store and number of household members.

28. There is no significant difference between the respondents’ level of willingness to change store and store distance.

29. There is no significant difference between respondents’ level of willingness to change store and duration of patronage.

30. There is no significant difference between the respondents’ share of wallet and store distance.

31. There is no significant difference between the respondents’ share of wallet and duration of patronage.

32. There is no significant difference between the respondents’ share of wallet and unplanned purchase.

33. There is no significant difference between the respondents’ share of wallet and shopping duration.

34. There is no significant difference between the respondents’ opinion gap in each of the six dimensions and occupation.
35. There is no significant difference between the respondents’ opinion gap in each of the six dimensions and educational qualification.

36. There is no significant difference between the respondents’ opinion gap in each of the six dimensions and age.

37. There is no significant difference between the respondents’ opinion gap in each of the six dimensions and departmental stores.

38. There is no significant difference between the respondents’ expectation gap in each of the six dimensions and occupation.

39. There is no significant difference between the respondents’ expectation gap in each of the six dimensions and age.

40. There is no significant difference between the respondents’ expectation gap in each of the six dimensions and educational qualification.

41. There is no significant difference between the respondents’ expectation gap in each of the six dimensions and departmental stores.

42. There is no significant difference between mean ranks of the six dimensions of customer satisfaction.

3.12 Limitations

All research has its own limitations and this study is no exception. All possible efforts were made to maintain objectivity, validity and reliability of the study, yet certain limitations need to be kept in mind whenever its findings are considered for implementation. The results are based on the responses provided by the respondents chosen through appropriate sampling technique. The views and ideas of the respondents are time dependable and it differs over time. The study is area specific and thus cannot be generalised for other parts of the country having dissimilar business environmental conditions. The study does not cover small and
medium size departmental stores. Due to vague, biased and similar responses of customers which arrived because of lack of interest and time constraint, the subjectivity cannot be ruled out. It proved to be a complex process that was better in theory than in practice for a variety of reasons. So it was simply so difficult and expensive to track and keep the high volume of records needed accurately and constantly update them periodically.