Research Methodology is the systematic plan to conduct research. It is the blueprint of carrying out research in a systematic and scientific manner. It includes the numerous ways of data collection and the techniques of data analysis but the selection of techniques must be logical. The present chapter describes the research methodology used in this research work. The rationale behind undertaking this study and the significance of this research along with the research objectives has been presented. As the study is about developing the empirical model of customer loyalty by establishing the inter-relationship among various variables of customer loyalty, so the step-wise research design comprising of population and sampling, measures and procedure used for conducting the present study is discussed.

3.1 RATIONALE BEHIND THE STUDY

The Indian mobile telecommunications sector is growing at a fast pace accompanied by the rapid increase in tele-density and subscribers base. The revenue generated by the sector has been expanding over the years thereby ever enhancing its contribution to India’s GDP. With rapid technological advancements in the country and government’s committed aspirations towards digital infrastructure, the sector can play a significant role in shaping the future by not only making the country more connected but also generating millions of jobs. Given its dynamic nature and the huge potential it holds, the Indian mobile telecom sector conceives a strong case for further research to enable its various stakeholders in better decision making.

Customer satisfaction and customer loyalty are the cornerstones of any industry. The existing literature reveals a number of research studies regarding the measurement of customer satisfaction and customer loyalty. These studies deal with different variables affecting customer satisfaction and customer loyalty but no study has been conducted to incorporate all the variables and to diagnose the inter-relationships among these variables affecting customer satisfaction and loyalty. This gap is filled in the present study by using an aggregated approach where all the variables affecting customer satisfaction and loyalty are identified and their inter-relationships are analyzed by developing a wholesome empirical model by taking inputs from all the significant areas of customer experience and perception. This will go a long way in understanding their stand-alone significance and their simultaneous
impact on customer loyalty. For the development of this model the data related to customer experience and perception is collected from two telecom circles.

Since different telecom circles are characterized by different and unique set of issues, the problems faced by each need to be studied in the light of its own specific characteristics. Therefore, two telecom circles i.e. Delhi and Haryana, classified on the basis of their revenue potential have been selected for the present study. The results of the separate telecom circles and the comparison amongst themselves reflect deep insights into the finer nuances of these circles which prove to be significant to the telecom service providers functioning in these circles along with the other stakeholders.

3.2 SIGNIFICANCE OF RESEARCH

The present study has been undertaken keeping in view the strategic relevance of understanding customer satisfaction and customer loyalty in the ever-evolving world of disruptive technologies. With continuous developments in the telecommunications sector, it is imperative for the academia to deliver research based results which can serve as crucial determinants of industry decisions. It is important to achieve an academia-industry interface where the research efforts can offer solutions and insights as to how to best manage the challenges they face on a regular basis.

Research findings have always formed the backbone of industry decisions. The present study holds the significance of providing a comprehensive model of customer loyalty which is applicable to the current scenario of the Indian telecom sector. This model is developed owing to the fact that the contribution of telecom sector to India’s GDP, revenue generated by it and the role it plays in Indian economy are ever enhancing. The dynamic nature of Indian telecommunications sector and the huge potential it holds, makes the Indian telecom sector a strong case for further research. The multiple facets of this sector play an important role in accomplishing the objectives of telecom companies. The wholesome empirical framework developed in the study helps in understanding the simultaneous impact of all the relevant factors on customer loyalty. Therefore, all major antecedents of customer loyalty in telecommunications sector have been considered in this research.
The most notable significance of the study lies in the fact that its results when seen separately for the selected telecom circles and then compared amongst themselves provide deep actionable insights applicable to the unique set-ups of these circles. These insights go a long way in devising effective customer retention strategies. Thus, it is safe to say that the findings of this study will prove to be of significant value to the telecom service providers functioning in the selected circles along with the other stakeholders involved.

3.3 RESEARCH OBJECTIVES

The purpose of this study is to empirically investigate the variables of customer satisfaction and customer loyalty within the Indian telecommunications sector. In the present study, the aim is to explore and assess the customer satisfaction and customer loyalty among prepaid telecom service subscribers considering all the drivers including:

*Perception based drivers* such as overall quality of service, price charged by the telecom service provider, value for money and image of the telecom service provider

*Experience based drivers* such as various customer touch points like network, call centre, recharging, data, tariff, value added service, advertisements etc.

In line with the aim of the study and the stated problem, the following specific research objectives have been formulated:

1) To develop an empirical model of customer loyalty based on identifying inter-relationships among all its key drivers.

2) To evaluate the impact of each of the drivers on customer loyalty by utilizing the developed framework.

3) To comparatively assess customer loyalty and customer satisfaction among prepaid mobile telecommunication customers of key service providers from Delhi and Haryana telecom circles.

4) To comparatively assess customer satisfaction on key perceptual dimensions along with various customer touch points like network, call centre, complaint resolution etc. among Delhi and Haryana telecom subscribers.
5) To examine the difference in customer satisfaction and customer loyalty on the basis of customer demographics, Socio-Economic Class (SEC), stated average monthly expenditure on telecom service and stated Age on Network (AON).

6) To examine the difference in customer satisfaction and customer loyalty levels for different brands owned by respective telecom service subscribers.

7) To find out the Net Promoter Score (NPS) for each of the telecom service providers.

3.4 RESEARCH DESIGN

To achieve the above stated objectives, a quantitative research design has been used. The primary data has been simultaneously collected from the current prepaid mobile users of two telecom circles, Haryana and Delhi using a structured questionnaire. The data thus obtained is compared using quantitative tests of significance namely, Student’s t-test and Analysis of Variance (ANOVA).

3.4.1 Population and Sample Size

The population under study constitutes the Indian prepaid mobile users belonging to two telecom circles: Haryana and Delhi. The population comprises of prepaid mobile users of different telecom companies like Airtel, BSNL, Idea, Vodafone, etc. The total sample size taken for this study is 770, selecting 385 respondents from each circle. The target respondents are selected using a combination of convenience and snowball sampling techniques. This sample size has been calculated taking 95% confidence level and an error margin of 5%, as given by the below mentioned formula:

\[ n = \frac{Nx}{(N-1)E^2 + x} \]

where,

\[ x = Z[(\frac{c}{100})^2 r(100-r)] \]

\[ E = \sqrt{\frac{(N-n)x}{n (N-1)}} \]

In the above formula, \( n \) represents sample size, \( E \) represents margin of error, \( N \) is the population size, \( r \) is the responses that one is interested in, and \( Z(c/100) \) is the critical value for the confidence level \( c \).
During the data analysis stage, 36 inconsistent cases are removed according to the Mahalanobis distance statistics by AMOS software and thus, the effective sample of 734 is used for data analysis (W1). Taking into consideration the population size and sample size, the obtained results can be seen at 95% confidence level.

### 3.4.2 Variables under Study

The chief objective of the present study is to develop an empirical model of customer loyalty in the context of Indian mobile telecommunications sector. This model aims to encompass all the major dimensions of the total customer experience and perceptions which contribute to overall customer loyalty.

**Figure 3.1: Variables under study**

Therefore, the major variable of this study is customer loyalty. Other variables under study are the ‘Perception-based drivers’ which include Quality, Value for Money, Price, and Image and the ‘Experience-based drivers’ including Network, Recharging, Tariff, Short Messaging Service (SMS), Value Added Services (VAS),
Research Methodology

Data Services, Call Centre, and Advertisement. The diagrammatic representation of the variables under study is given in Figure 3.1.

### 3.4.3 Questionnaire Design

For the purpose of data collection, a structured questionnaire is constructed. The first part of the questionnaire focuses on the demographic information about the participants of the study. The demographic information collected through the questionnaire includes age, gender, occupation, education, area of residence (urban, sub-urban, or rural), and total household income. The information collected from this part of the questionnaire leads to sample categorization.

Figure 3.2 depicts the flow chart including various steps followed in the construction of structured questionnaire.

![Figure 3.2: Steps followed in construction of structured questionnaire](image)

A step-wise process is undertaken in the development, refinement and validation of the structured questionnaire measuring customer loyalty. So as to develop a reliable and valid customer loyalty measure, all the necessary scale
development procedures are adopted. The methodology used to develop this questionnaire follows the recommendations of Churchill (1979), Gerbing and Anderson (1988), and Saxe and Weitz (1982). An extensive literature review is carried out prior to item generation in order to conceptualize constructs and identify important dimensions of customer loyalty. In total, 200 items were identified from the literature. Before proceeding to construction of any questionnaire, it is important to tap into the beliefs and perceptions of the target population with respect to the variables of interest. In this study, the target population is prepaid telecom service subscribers belonging to Haryana and Delhi telecom circles. The exploratory and qualitative interviews are conducted with the participants to better understand the relevant drivers of customer loyalty perceived as important by them. These personal in-depth interviews, comprising of open-ended questions are conducted with twenty two participants as out of randomly selected thirty, eight participants were not available. Each interview is conducted for about 30 to 40 minutes. The questions on the following issues are asked in the interviews:

- the factors and aspects of telecom services considered important by the customers in order to be truly loyal to their respective service provider,
- the key factors that influence the customers’ perceptions of loyalty with mobile telecommunications services

These in-depth interviews highlight some new items apart from literature. 21 items are identified from these interviews. The responses obtained through interviews provide valuable insights about customer loyalty and various aspects of telecom services perceived as important by the respondents. Both the in-depth interviews with target respondents and the review of existing literature on customer loyalty offer valuable information about the critical dimensions of customer loyalty. These dimensions can be categorized under two groups:

i)  *Perception-based drivers* which include Quality, Value for Money, Price, and Image.

ii)  *Experience-based drivers* which include Network, Recharging, Tariff, Short Messaging Service (SMS), Value Added Services (VAS), Data Services, Call Centre, and Advertisement.
The items for perception and experience-based drivers are compiled from the inferences extracted through the review of the research studies as well as from the responses of exploratory interviews. In the screening process out of 221 total items identified from literature as well as in-depth interviews, each item is checked in terms of consistency and less significant items are discarded. In total, 60 items are generated at this stage. These items are further incorporated in structured questionnaire which measures the customers’ perceptions regarding service quality on a five-point Likert scale, with 5 denoting excellent/strongly agree and 1 denoting poor/strongly disagree (Boulding et al., 1993 and Babakus and Boller, 1992).

After the questionnaire design, a pilot study is conducted with a small sample of 50 respondents. This is done in order to identify the gaps in questionnaire. During this exercise, the respondents are asked to provide feedback on clarity and flow of questionnaire items. This feedback goes a long way in ascertaining face validity of items. On the basis of the feedback received, minor but necessary modifications are made in the items.

Opinions from subject matter experts are sought in order to assess the validity of questionnaire. A panel of three marketing judges evaluated the items for content validity and suggested modifications in some items while complete removal of other items which they considered to be redundant, double-barreled, and ambiguous. This process resulted in the elimination of 6 items, leaving the final 54 items in the questionnaire. These 54 items are arranged as contributing to ten distinct constructs. The detailed list of these constructs and items falling under each construct is presented in chapter 4. In this manner, the questionnaire is finalized and is administered to the respondents by following a combination of face-to-face and web-based methodologies.

3.4.4 Sample Categorization

The sample is derived from two telecom circles, Delhi and Haryana only. The total sample size taken for this study is 734, out of which 373 (50.81%) respondents are from Haryana telecom circle and 361 (49.18%) respondents are from Delhi circle. The telecom circle for any customer is determined by the place from which he or she has purchased the telecom connection. For instance, a person who has bought his
connection from a retailer located in Delhi will be considered to belong to ‘Delhi’ telecom circle. Also, the telecom customers who have subscribed to more than one telecom operator are requested to respond to their primary telecom connection i.e. the one that they mostly used.

The sample is further categorized on the basis of demographic characteristics. On the basis of gender, the total sample of 734 participants comprised of 417 (56.81%) female and 317 (43.18%) male respondents from both the telecom circles.

On the basis of the area of residence, the total sample can be further categorized into three categories: Urban, Rural and Sub-urban. 488 (66%) respondents are from urban area of residence, forming the biggest chunk followed by 146 (20%) respondents belonging to rural area of residence. Only 100 (14%) respondents belong to the sub-urban area of residence. This sample distribution is shown in Figure 3.3:

![Sample distribution on the basis of area of residence](image)

**Figure 3.3: Sample distribution on the basis of area of residence**

Another basis of sample categorization is age. All the respondents are assigned one out of five categories formed on the basis of age. These categories are shown in Figure 3.4. This figure shows the ‘age’ wise sample distribution. Out of the total sample of 734 prepaid mobile users, 172 (24%) respondents belong to age category A, 308 (42%) respondents belong to age category B, 127 (17%) respondents to category C, 98 (13%) respondents to category D while only 29 (4%) respondents belong to age category E.
The sample is also categorized on the basis of Socio-Economic Class (SEC). For this purpose, SEC classification by Market Research Society of India is used. This classification takes into account the educational level and occupation of the chief wage earner of the household and segment them into eight classes i.e. A1 to E2 where SEC A and SEC B refer to high socio-economic class, SEC C refers to mid and SEC D and SEC E refer to low socio-economic class. This categorization is represented in Figure 3.5.

The figure reveals that out of 734, 197 (27%) respondents belong to socio-economic class A1, 283 (38%) respondents belong to class A2, 42 (6%) respondents belong to class B1, 94 (13%) respondents belong to class B2, 52 (7%) respondents belong to class C, 56 (8%) respondents belong to class D. Only 8 (1%) respondents belong to socio-economic class E1 and only 2 (0%) respondents belong to socio-economic class E2.
Further the sample is categorized on the basis of ‘Age on Network’ criterion i.e. the duration for which the respondent has been using the network. This categorization is represented in Figure 3.6.

Figure 3.6 shows that 68 (9.26%) respondents have the minimum age on network and belong to category A, 71 (9.67%) respondents belong to category B, 134 (18.25%) respondents belong to C, 222 (30%) respondents belong to D, and 239 (33%) respondents belong to category E, who have the maximum age on network i.e. more than 5 years.
Another basis of sample categorization is the average monthly expenditure on telecom service. On this basis, five categories are identified which are shown in Figure 3.7.

![Sample distribution on the basis of average monthly expenditure on telecom service](image)

**Figure 3.7: Sample distribution on the basis of average monthly expenditure on telecom service**

Figure shows that out of the total 734 respondents, 76 (10.35%) respondents belong to the average monthly expenditure category A who spend minimum amount on telecom services; 252 (34.33%) respondents belong to the average monthly expenditure category B, 245 (33.37%) respondents belong to category C, 123 (17%) respondents belong to category D, and 38 (5%) respondents belong to category E with maximum monthly expenditure.

Further, the sample is categorized according to the telecom operator subscription criterion. Eight telecom operators are considered including A – Aircel, B – Airtel, C – BSNL, D – Idea, E – MTNL, F – Reliance, G - Tata Docomo, and H – Vodafone. Diagrammatically this categorization is presented in Figure 3.8.
As seen in Figure 3.8, out of the total 734 respondents, 22 (2.99%) respondents subscribe to telecom service provider Aircel, 213 (29.01%) respondents to Airtel, 84 (11.44%) respondents to BSNL, 100 (14%) respondents to Idea, 14 (2%) respondents to MTNL, 29 (4%) respondents to Reliance, 78 (11%) respondents to Tata Docomo, and 194 (26%) respondents subscribe to telecom service provider Vodafone. So, out of the selected sample, Airtel enjoys the maximum subscription base as compared to MTNL with minimum subscription.

Therefore, the above description portrays the demographic characteristics of the sample depicting the originality and range of the data taken as a base for model development. The variations shown in the sample categorization ensure that the sample is truly representing the population and provides the foundation to generalize the applicability of the model across sampling frames taken from other telecom circles.

At the end, the above detailed discussion on research design used in the present research work is highlighted step-wise in Figure 3.9:
3.5 DATA ANALYSIS TECHNIQUES

After the data collection, the data is analysed. Apart from using some basic statistical techniques like mean, standard deviation, standard error, t-test etc., some advanced techniques like SEM, Factor Analysis, ANOVA and Tukey’s Honest Significant Difference (HSD) test are also used in this research which are described as under:
3.5.1 Mean, Standard Deviation and Standard Error

The descriptive statistics including mean, standard deviation and standard error is described as under:

(a) Mean is used to estimate the average value when the data is collected using an interval or ratio scale. It is obtained by summing up all the elements in a set and dividing the sum by the number of elements.

\[ \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \]

where \( X_i \) represents the observed values of the variable \( X \); \( n \) is the number of observations (sample size) and \( \bar{X} \) is arithmetic mean of data.

(b) Standard Deviation is the square root of variance and is calculated as

\[ SD = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}} \]

(c) Standard Error (SE) of mean is the standard deviation of the sampling distribution of the mean or proportion.

\[ SE = \frac{SD}{\sqrt{n}} \]

3.5.2 \( t \) – Test

\( t \)-test is based on the Student’s \( t \) statistic. The \( t \) statistic assumes that the variable is normally distributed and the mean is known. Assume that the random variable \( X \) is normally distributed, with population mean \( \mu \) and unknown population variance \( (SD)^2 \), which is estimated by the sample variance \( s^2 \).

Then,

\[ t = \frac{\bar{X} - \mu}{SD_{\bar{X}}} \]

is \( t \) distributed with \( n-1 \) degrees of freedom.

where \( SD_{\bar{X}} \) is the standard deviation of the sample mean, \( \bar{X} \). \( SD_{\bar{X}} \) is estimated as

\[ SD_{\bar{X}} = \frac{s}{\sqrt{n}} \]
3.5.3 Structural Equation Modelling (SEM)

In order to achieve the stated research objectives, the quantitative research design is undertaken for the present study. The population under study consists of all prepaid telecom service subscribers from Delhi and Haryana telecom circles in India. The target respondents are selected using a combination of purposive as well as snowball sampling technique. Initially, the total sample size is 770, wherein 385 respondents each from Delhi and Haryana telecom circles are surveyed for the study by using a structured questionnaire.

Taking into consideration the population size of target respondents and the above mentioned sample size, the results obtained from the study can be seen at 95% confidence level while the margin of error is kept at 5%. The survey data is analyzed by using a two-step structural equation modelling technique. The first step assesses the measurement model validity by using confirmatory factor analysis (Hair et al., 2009), whereas the second step assesses the structural relationships among the research variables in the hypothesized model (Kline, 2011). IBM Analysis of Moment Structures (AMOS) statistical package version 23 is applied for the purpose of model development using structural equation modelling (SEM) technique. During the process of model development, 36 inconsistent cases are found which are removed from the total sample by using Mahalanobis distance statistics, as generated by AMOS. Thus, the effective sample size is confined to 734 cases. The adequacy of sample size is vital when the statistical power of an analysis is taken into account in case of covariance-based methods, especially in SEM. A sample of 200 or greater is regarded as adequate for SEM analysis (Hoe, 2008; Hooper et al., 2008; Kline, 2011). The SEM technique encompasses two processes – measurement model and structural model. The measurement model involves a Confirmatory Factor Analysis (CFA), which deals with how well the measured or observed variables represent the latent variables. Confirmatory Factor Analysis (CFA) is a statistical strategy specifically designed to identify and explore hypothetical constructs as manifested by the observed variables. It is used to verify the factor structure of a set of observed variables. It allows the researcher to test the hypothesis that a relationship between observed variables and their underlying latent constructs exists (Suhr, 2006). The method requires a strong empirical or conceptual foundation to guide the specification.
and evaluation of the factor model. CFA can be used for a variety of purposes, such as psychometric evaluation, the detection of method effects, construct validation, the evaluation of measurement invariance and examination of the latent structure of a test instrument during the process of scale development.

On the other hand, the structural model incorporates multiple regression analysis and path analysis and it models the relationships among the latent variables (Chen et al., 2011). Multiple regression is an extension of simple linear regression. It is used when we want to predict the value of a variable based on the value of two or more other variables. The variable we want to predict is called the dependent variable. The variables we are using to predict the value of the dependent variable are called the independent variables. Multiple regression analysis is one of the most widely used statistical procedures for prediction by finding the linear combination of a set of predictors that provides the best point estimates of the dependent variable across a set of observations. Predictive accuracy is calibrated by the magnitude of the $R^2$ and the statistical significance of the overall model. It is also used to draw conclusions about individual predictor variables (Mason and Perreault, 1991). The extension of multiple regression is Path Analysis. Its aim is to provide estimates of the magnitude and significance of hypothesized causal connections between sets of variables. It is a method employed to determine whether or not a multivariate set of non-experimental data fits well with a particular causal model. Structural Equation Modelling (SEM) is an integration of Confirmatory Factor Analysis and Path Analysis methods. The procedure of modelling using SEM technique can be summarized as follows:

i) Define the structural and measurement components (e.g. measured variables and latent variables) to set up a hypothetical model.

ii) Assess the verification of hypothetical model and develop it with modifications, if necessary.

iii) Assess the verification of the final model and interpret it (Chen et al., 2011)

The chief objective here is to model the various contributory factors of customer loyalty using SEM, highlighting that errors might occur during this modelling process which are attenuated while validating the model. Ultimately, the results are confirmed based on the survey findings and are explained in chapter 4.
SEM is a comprehensive statistical approach to test relationships among observed and latent variables (Hoyle, 1995). The main reason behind using SEM for the study was the capability of modelling complex dependencies and latent variables which is also regarded as its main merit (Nachtingall et al., 2003). Also, SEM helps in proper assessment of various interrelated relationships, has the capability to represent unobserved constructs in these relationships and accounts for measurement error as well (Hair et al., 2009). The developed model takes care of reducing measurement errors, which diminishes biases in the SEM model. Structural model analysis tests relationships amongst latent exogenous and endogenous constructs, as well as relationships among the latent endogenous constructs. Another important capability is that SEM can easily detect multicollinearity problems (Folmer et al., 2010; Radosevic and Yoruk, 2013; Xiong et al., 2014). It also permits modifications within a justifiable setting (Hair et al., 2014). Hence, SEM is aptly suited for analyzing the type of data which is used in the present study.

Now-a-days, the second generation method of multivariate analysis i.e. Structural Equation Modelling (SEM) is fastly being accepted in academic research. Given the efficiency of analysis using SEM, it greatly helps in making proper interpretation of the results. It, therefore, guides in making right and precise decisions. Several researchers including Anderson and Gerbing, 1988; Baron and Kenny, 1986; Bollen, 1989; Byrne et al., 1989; and Joreskog and Sorbom, 1993, emphasized that SEM should be accepted as the preferred method for academicians, and even proclaimed that SEM would be of great help for the coming generation of researchers as well.

Structural Equation Modelling (SEM) consists of two types - Variance Based Structural Equation Modelling (VB-SEM) and Covariance Based Structural Equation Modelling (CB-SEM). These two packages differ greatly from each other in terms of their statistical approaches, namely the non-parametric testing and the parametric testing, the objective of the study namely exploratory and confirmatory, and more importantly the algorithm employed, namely Generalized Least Square (GLS) Estimation and Maximum Likelihood Estimation (MLE) (Esposito, 2009). The algorithm employed in VB-SEM, popularly known as PLS-SEM (Smart-PLS and Warp-PLS), is Generalized Least Squares (GLS) while the algorithm employed in
CB-SEM (AMOS, etc.) is the Maximum Likelihood Estimator (MLE). These two types of algorithm differ greatly from each other in terms of efficiency of their statistical estimates for path coefficients (Hair et al., 2014). PLS-SEM is meant for exploratory research while CB-SEM is meant for confirmatory research (Hair et al., 2011). Now-a-days, most of the statisticians would not compare the capability of the two SEM approaches - VB-SEM and CB-SEM, but instead stress that the two complement one another. (Awang et al., 2015).

3.5.3.1 The fitness of model

In CB-SEM, there are two models involved namely measurement model (measuring individual latent construct) and structural model (measuring the inter-relationships among latent constructs). The measurement model should be assessed first prior to modelling structural model (SEM). The assessment of measurement model is done through the Confirmatory Factor Analysis (CFA) for unidimensionality, validity and reliability of items measuring the construct. However, the unidimensionality assessment should be made first prior to assessing validity and reliability. In unidimensionality step, the items having low factor loading (less than 0.6) should be deleted while redundant items could either be deleted or constrained. (Afthanorhan, 2014; Zainudin, 2012, 2014).

All measurement models need to achieve certain fitness indices as stated in the literature (Zainudin, 2014). The fitness indices reflect how well the measurement model of a construct fits to the data collected from the field. Therefore, the assessment of measurement model in the first place is crucial for modelling SEM itself. However, such assessment is only available in CB-SEM rather than in PLS-SEM. In fact, the CB-SEM produces many fitness indices which reflect the appropriateness of a measurement model but it is enough for scholars to report only a few indices such as Chi-square (Wheaton et al., 1977), Root Mean Square of Error Approximation (Browne & Cudeck, 1993), Goodness of Fit Index (Joreskog & Sorbom, 1984), Adjusted Goodness of Fit Index (Tanaka & Huba, 1985), Comparative Fit Index (Bentler, 1990), Tucker-Lewis Index (Bentler & Bonett, 1980), Normed Fit Index (Bollen, 1989), and $\chi^2/df$ (Marsh & Hocevar, 1985). Table 3.1 presents the description of each of the fit indices:
## Table 3.1: Description of Model Fit Indices

<table>
<thead>
<tr>
<th>Fit Index</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$</td>
<td>To assess the magnitude of discrepancy between the sample and fitted covariance matrices</td>
</tr>
<tr>
<td>RMSEA</td>
<td>Sensitive to the number of estimated parameters in the model. It will choose the model with the lesser number of parameters.</td>
</tr>
<tr>
<td>GFI</td>
<td>Calculate the proportion of variance that is accounted for by the estimated population covariance</td>
</tr>
<tr>
<td>AGFI</td>
<td>Adjusts the GFI based upon degree of freedom and this fitness index increases with increase in sample size</td>
</tr>
<tr>
<td>NFI</td>
<td>To assess the model by comparing the chi-square value of the model to the chi-square of the null model</td>
</tr>
<tr>
<td>CFI</td>
<td>Assumes that all latent variables are uncorrelated (null/independence model) and compares the sample covariance matrix with this null model</td>
</tr>
<tr>
<td>$\chi^2/df$</td>
<td>The estimation process is dependent on the sample data</td>
</tr>
</tbody>
</table>

(Source: Hooper et al., 2008)

### 3.5.4 Factor Analysis

Factor analysis is a very useful method in marketing research. It is a general term which denotes a class of procedures which have the objective of reducing data complexity. In marketing research, there may be a large number of variables, most of which are correlated. Factor analysis examines the relationship between interrelated variables and represents them in terms of few underlying factors. Factor analysis is interdependence technique whereas in multiple regression, discriminant analysis and analysis of variance, one variable is considered as the dependent or criterion variable and the others as independent or predictor variables. No such distinction is made in factor analysis and it examines the whole set of interdependent relationships. Factor analysis is used in following circumstances:

- To identify the nature of constructs underlying responses in a specific content area;
- To determine what sets of items ‘hang together’ in a questionnaire;
- To determine a smaller set of salient variables while classifying a group of items;
- To demonstrate the dimensionality of measurement scale as researchers often wish to develop scales responding to a single characteristic; and
Research Methodology

- To generate ‘factor means’ representing values of the underlying constructs for use in further analysis.

This technique has various applications in marketing research. It is used in market segmentation, in product research by identifying the different attributes that influence the consumers’ choice, in advertising research by analyzing the media consumption habits of the target customers, in pricing decisions by identifying the characteristics of price-sensitive consumers and so on.

3.5.4.1 Factor analysis model and statistics

In factor analysis, each variable is expressed as a linear combination of underlying factors. The covariation among the variables is described in terms of small number of common factors plus a unique factor for each variable. If the variables are standardized, the factor model may be represented as:

$$X_i = A_{i1}F_1 + A_{i2}F_2 + A_{i3}F_3 + \ldots + A_{im}F_m + V_iU_i$$

where $X_i$ is the $i$th standardized variable, $A_{ij}$ is the standardized multiple regression coefficient of variable $i$ on common factor $j$, $F$ is common factor, $V_i$ is standardized regression coefficient of variable $i$ on common factor $j$, $U_i$ is the unique factor for variable $i$ and $m$ is the number of common factors.

The unique factors are uncorrelated with each other and with the common factors. The common factors themselves can be expressed as linear combination of the observed variables.

$$F_i = W_{i1}X_1 + W_{i2}X_2 + W_{i3}X_3 + \ldots + W_{ik}X_k$$

where $F_i$ is the estimate of the $i$th factor, $W_i$ is the weight or factor score coefficient and $k$ is the number of variables.

It is possible to select the factor scores or weights. First factor explains the largest portion of total variance. Second factor accounts for the maximum of the residual variance which is uncorrelated with first variance and so on. Various statistics are associated with factor analysis which are presented as under:

**Bartlett’s test of sphericity:** It is used to examine the hypothesis that population correlation matrix is an identity matrix. Each variable correlates perfectly with itself ($r = 1$), but has no correlation with other variables ($r = 0$).
Correlation matrix: It reflects the simple correlations (r) between all possible pairs of variables included in the analysis.

Communality ($h^2$): It is the amount of variance a variable shares with all other variables. High value of communality means that not much of the variables is left over after whatever the factors represent.

$$h^2 \text{ of the } ith \text{ variable} = (ith \text{ factor loading of factor } A)^2$$

$$+ (ith \text{ factor loading of factor } B)^2 + \ldots$$

Eigen value: It is the sum of squared values of factor loadings to a factor and hence represents the total variance explained by each factor.

Factor: It is an underlying dimension that accounts for several observed variables. There can be one or more factors depending upon the nature of the study and the number of variables involved in it.

Factor loadings: These are simple correlations between the variables and the factors. These are also known as factor-variable correlations which work as a key to understand what factors mean.

Factor matrix: It contains the factor loadings of all the variables on all the factors extracted.

Factor loading plot: It is the plot of original variables which uses the factor loadings as coordinates.

Factor scores: These represent the degree to which each respondent gets high scores on the group of items that load high on each factor.

Factor scores coefficient matrix: It contains the weights or factor score coefficients which are used to combine the standardized variables to obtain factor scores.

KMO measure of sampling adequacy: The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is the index used to determine whether the factor analysis is appropriate or not. The values between 0.5 and 1 indicate that factor analysis is appropriate and values below 0.5 imply that factor analysis is not appropriate.

Percentage of variance: It is the percentage of total variance attributed to each factor.

Residuals: These are the differences between observed correlations (given in input correlation matrix) and the reproduced correlations (as estimated from the factor matrix).
Scree plot: It is the plot of eigen values against the number of factors in order of extraction.

3.5.5 Analysis of Variance

Analysis of Variance (ANOVA) is used for comparison of means among more than two categories. It is a collection of statistical models and their associated procedures (such as "variation" among and between groups) used to analyze the differences among group means. It provides a statistical test of whether or not the means of several groups are equal, and therefore generalizes the t-test to more than two groups. ANOVA is useful for comparing three or more means (groups or variables) for statistical significance. The significance of difference between the means of two samples can be judged through either z-test or t-test but ANOVA is used when two or more populations are involved. It is essentially a procedure for testing the difference among different groups of data for homogeneity. The essence of ANOVA is that the total amount of variation in a set of data is broken down into two types, that amount which can be attributed to chance and that amount which can be attributed to specific causes. There may be variations between samples and also within sample items.

In its simplest form, ANOVA must have one dependent variable (metric) and one or more independent variables (categorical) which are also known as factors. A particular combination of factor level or categories is called treatments. One-way ANOVA involves only a single factor and if two or more factors are involved the analysis is termed as n-way ANOVA. If the set of independent variables consists of both categorical and metric variables then the technique is called analysis of covariance (ANCOVA).

The data in the present study is analyzed with one-way ANOVA. The various statistics associated with one-way ANOVA are explained as under:

3.5.5.1 One – way ANOVA and statistics

The statistics associated with one-way ANOVA are described as under:

\( \eta^2 \): The strength of the effects of X (independent variable or factor) on Y (dependent variable) is measured by \( \eta^2 \), the value of which varies between 0 and 1.

\( F \) statistic: It is used to test the null hypothesis that the category means are equal.
**Mean Square:** It is the sum of squares divided by appropriate degree of freedom.

$SS_{\text{between}} (SS_x):$ It is the variation in Y related to the variation in the means of the categories of X. This presents the variation between the categories of X, or the portion of the sum of squares in Y related to X.

$SS_{\text{within}} (SS_{\text{error}}):$ This is the variation in Y due to the variation within each of the categories of X. This variation is not accounted for by X.

$(SS_y):$ It is the total variation in Y.

Wherever ANOVA depicts a significant difference between categories, Post-hoc Analysis using Tukey’s Honest Significant Difference (HSD) Test is done to better understand which categories actually differ significantly from each other on the said variables.

### 3.5.6 Tukey’s Honest Significant Difference Test

The Tukey’s Honest Significant Difference (HSD) test is a post-hoc test based on the studentized range distribution. An ANOVA test can tell if the results are overall significant, but it does not tell exactly where those differences lie. The answer lies in Tukey’s HSD test. This test can be run to find out which specific groups’ means are different. The test compares all possible pairs of means. To test all pairwise comparisons among means using the Tukey’s HSD, HSD for each pair of means is calculated using the following formula:

$$HSD = \frac{M_i - M_j}{\sqrt{\frac{MS_w}{n}}}$$

where $M_i - M_j$ is the difference between the pair of means, $MS_w$ is the mean square within and $n$ is the number of scores used in calculating the group means of interest.

In the present study, data is analyzed with the above-mentioned statistical techniques. To conduct such statistical analysis, SPSS version 21 and AMOS version 23 computer packages are used. The entire process of data analysis is presented sequentially in Figure 3.10.
3.6 ORGANIZATION OF THE STUDY

The present thesis is arranged in the form of following chapters:

Chapter 1: Customer Satisfaction and Customer Loyalty in Indian Telecom Sector: A Conceptual Framework

Chapter 2: Review of Literature

Chapter 3: Research Methodology

Chapter 4: Development of an Empirical Model of Customer Loyalty

Chapter 5: Comparative Analysis on the Basis of Sample Categorization: Results and Interpretation

Chapter 6: Major Findings, Managerial Implications and Future Research Directions