CHAPTER 3
RESEARCH METHODOLOGY

INTRODUCTION

In this chapter the focus of the study in terms of the geographic location and in terms of the products being studied has been discussed. This chapter also consists of the information about the sampling frame, research design and sources of data collection. The research objectives, the research questions and the research hypothesis have been discussed in detail in this chapter. The chapter further consists of descriptions of various models and techniques that have been used in the study.

3.1 FOCUS OF THE STUDY

The research is an effort to find out the attitude of the consumers towards the NFR services. This study is exploratory as well as descriptive. The research has been carried out in the Delhi NCR region. NCR is the name for the metropolitan area that includes the National Capital Territory of Delhi and the urban areas in its neighbouring states, Haryana, Uttar Pradesh, Uttarakhand and Rajasthan. It has a total area of about 33,575 km². It is the world’s largest agglomeration by area and second largest urban agglomeration by population (United nation World Urbanisation Prospect report 2011, online data). As per the 2011 census, Delhi NCR is the topmost urban agglomeration with a population of 21,753,486 followed by Mumbai Metropolitan Region (Census 2011). The vehicular population of Delhi is more than the combined vehicular population of
Mumbai, Chennai and Kolkatta (Delhi government Environment survey 2010). According to the Hindustan times report dated 3rd march 2010, almost 40 Lakh vehicles cross the Delhi road every day.

For this study we have taken into consideration only those vehicles which run on petrol or diesel, vehicles consuming CNG has not be considered. Since most of the commercial vehicles like buses, taxi and auto rickshaw has been converted into CNG powered hence they have been excluded from the study. The target population in this study was the individuals who own or drive personal vehicle like car, jeep, SUV, MUV (four wheelers), Scooter, motorcycle (two wheelers), which consume petrol or diesels and individuals who drive heavy vehicle like trucks. The different categories of consumers surveyed include:

- Two wheelers (Motorcycle, Scooter etc.)
- Four wheeler (car, jeep, SUV, MUV)
- Trucks and light commercial vehicles

3.2 SAMPLING FRAME

The sampling frame is the list of the target population. Here the sampling frame is the individual who is 18 years or above and is using / driving two wheeler/ four wheeler/ truck. The study has addressed the difference among the consumer with respect to the kind of vehicle they are driving or owning. The study has also taken into consideration the shopping status of the individual. The respondents who used to do the main shopping for their house hold were called primary shoppers and others were called the non – primary shoppers. The study had also addressed the difference of attitude among different age group of users/ drivers of different types of vehicles.
3.3 RESEARCH DESIGN

The research design used in the present study is exploratory as well as descriptive. The exploratory research has been used to find out the various variables that affects the selection of the retail outlets. Based on the inputs from the literature review seventeen factors were identified that affects the selection of the retail outlets. From these seventeen factors, hundred variables were identified. These variables were reduced to seventy five after taking the opinion of the expert and the people working in the oil retail industry. Using these seventy five variables a questionnaire was prepared to find out the variables affecting the selection of the NFR outlets in Delhi NCR. The same questionnaire was used to study the usage pattern of the NFR in Delhi NCR. The factors identified from this survey were used to measure the attitude of the consumer towards the NFR outlets.

The second study was a descriptive study. For this study a structured questionnaire was used to measure the attitude of the consumer. The questionnaire consists of questions related to belief and evaluation of belief. The mean attitude score was calculated using Fishbein’s multiattribute attitude model. Then the average attitude score was compared between different types of vehicle owner/ drivers and also between different age groups of the respondents. The average attitude score was also compared between primary and non primary shoppers and also between commercial and non-commercial vehicle.

3.4 SOURCES OF DATA COLLECTION

In order to achieve the objectives, the data was collected through the primary and the secondary sources.
3.4.1 SECONDARY DATA

This method was used to collect literature on retailing, consumer attitude, attitude measurement, fuel retailing and non fuel retailing. The secondary source of data included data from research journals, magazines, books, and company’s websites and other electronic database like Ebsco, Emrald Insight, Proquest and India stat etc. An immense review of literature and work already done at international level was also taken into account for studying the factors affecting selection of NFR outlets. From this data the various variables affecting the selection of retail store were determined and these variables were included into the questionnaire of the pilot survey.

3.4.2 PRIMARY DATA

The primary data was collected was collected using structured questionnaire. The questionnaire was developed after an extensive literature review. The primary data was collected in two stages

3.4.2.1 FIRST STAGE

In this stage data was collected using structured questionnaire which was developed using the variables identified from the literature review. Prior to conducting survey this questionnaire was reviewed extensively by two academicians and two experts from petro retail industry. It was also pilot tested on 5 car user and 10 motorcycle/ scooter user before conducting the actual survey. The suggestion obtained from the reviewer and from the pilot test was incorporated and a final survey of 550 respondents was carried out to study the usage pattern of the NFR services in Delhi NCR. This study was also conducted to find out the factors affecting the selection of the NFR outlet.
3.4.2.2 SECOND STAGE

In the second stage of the primary data collection one more questionnaire was prepared. The information obtained from the first study was used to prepare this questionnaire. A total of 402 respondents were surveyed in this stage to find out the most important factor affecting the selection of the NFR outlet. This study was also conducted to find out the attitude of the consumer towards NFR services.

3.5 RESEARCH QUESTIONS

1. What is the usage pattern of NFR services?
   a. What is the usage pattern of NFR services among two wheeler, four wheeler and truckers?
   b. What is the frequency of use of NFR services by two wheeler, four wheeler and truckers?

2. What are the factors that affect the selection of NFR outlets?

3. What is the most important factor affecting selection of the NFR outlet?

4. What is the attitude of consumer towards NFR outlets?
   a. Is attitude same among two wheeler, four wheeler and truckers?
   b. Is attitude same among commercial and non commercial vehicles?
   c. Is attitude same among different age group?
   d. Is attitude same among primary shopper and non-primary shopper?
3.6 RESEARCH OBJECTIVES

The following objectives have been set up to answer the above research questions.

1. To study usage pattern of NFR services.
2. To find out the factors affecting the selection of NFR outlets.
3. To find out the most important factor affecting the selection of the NFR outlet.
4. To measure the attitude of the consumer towards NFR.

3.7 RESEARCH HYPOTHESIS

The following hypothesis were tested to answer the research questions and achieve the objectives

H1₀: There is no significant association between types of vehicle and type of NFR services being used.

H1ₐ: There is a significant association between types of vehicle and type of NFR services being used.

H2₀: There is no significant association between frequency of use of NFR services and types of vehicle.

H2ₐ: There is a significant association between frequency of use of NFR services and types of vehicle.

H₃₀: The different variables affecting attitude towards NFR outlets do not have any significant correlation in the population.
H3: The different variables affecting attitude towards NFR outlets do have a significant correlation in the population.

H4: The data is normally distributed.

H4a: The data is not normally distributed.

H5: There is no significance difference between the average attitude of two wheeler, four wheeler and Truck users/drivers.

H5a: There is a significance difference between the average attitude of two wheeler, four wheeler and Truck users/drivers.

H6: There is no significance difference between the average attitudes of commercial and non-commercial vehicle’s respondents.

H6a: There is a significance difference between the average attitudes of commercial and non-commercial vehicle’s respondents.

H7: There is no significant difference between the average attitudes of respondents belonging to different age group.

H7a: There is a significant difference between the average attitude of respondents belonging to different age group.

H8: There is no significant difference between the average attitudes of primary and non-primary shopper.

H8a: There is a significant difference between the average attitudes of primary and non-primary shopper.

Two surveys have been done to answer the research questions and achieve the research objectives. In the first study a survey of 550 individuals was done to
find out the usage pattern of NFR services and to study the factors affecting selection of NFR outlets. In the second study, a survey of 402 individuals was done to find out the most important factor affecting the selection of the NFR outlet and to study the attitude of the consumer towards these outlets.

3.8 SAMPLING TECHNIQUES AND PROCEDURE

To achieve the above stated objectives and to answer the various research questions, seven hypotheses were proposed. To test these hypotheses, samples were collected using non-probability stratified sampling technique. In this technique, the whole population is divided into different strata or subpopulation. After that, the proportionate of each stratum is fixed as per some relevant standard. Here, the whole population was distributed in three strata, two wheeler, four wheeler, and trucks as we were interested in three categories of vehicle owner/drivers. The proportionate of each strata is based on the data of Ministry of Statistics and Programme Implementation 2011-12 survey. According to this survey, the total number of registered vehicles in India includes 72% two wheeler, 11% cars, 1% jeep, 5% goods vehicle, and 3% light motor vehicle. Keeping in mind the above data and using weighted average, we have surveyed 78.19% two wheeler, 13.03% four wheelers, and 8.69% trucks. The respondents were selected randomly from convenient locations like parking lots of the offices and commercial places. The trucker’s were selected randomly from transport area.

3.9 SAMPLE SIZE

The whole study was divided into two parts. Each part uses different questionnaire. The sample size studied and the basis of selection of particular
sample size and the distribution of samples among different cities of NCR is explained below

### 3.9.1 SAMPLE SIZE OF THE FIRST STUDY

The first study was to identify the usage pattern of the NFR services and the factor affecting the selection of the NFR outlet. Which have formed a base to study the attitude of the consumer towards the NFR services. In this study a sample size of 550 was taken which is more than five times the number of variables. According to Dennis Child, the sample size should be at least five times the number of variables. The samples were collected from Delhi, Ghaziabad, Noida and Faridabad. The sample consists of 100 respondents from Delhi, 218 from Noida, 98 from Ghaziabad and 94 from Faridabad. A total of 399 two wheeler, 67 Four wheeler and 44 truckers were studied.

### 3.9.2 SAMPLE SIZE OF SECOND STUDY

In the second study a sample size of 402 was taken. We arrive at this size using the information obtained during the primary survey. The standard deviation and the standard error obtained from the first survey was used to calculate the sample size of the final survey at 95% confidence level. The standard deviation calculated from the pilot survey was 0.493, z value at 95% is 1.96

Putting these values in the formula

\[ N = \left( \frac{Zs}{e} \right)^2 \]

Where

- \( N = \) Sample size,
- \( z = 1.96 \) (at 95% confidence level),
- \( s = \) population standard deviation (here \( s = 0.493 \))
\[ e = 0.05 \]

We get,

\[ N = \frac{(1.96 \times 0.493/0.05)^2}{2} \]
\[ = 373.47 \]
\[ = 374 \]

Thus the minimum sample size suitable for the study was 374. Keeping in mind the non response error (7% from pilot survey) a sample size of 402 was taken in the present study.

Then this sample size was divided into different strata. The size of the strata was based on the report of MoPN (Appendix Table -A18). According to this report there are 72% two wheeler, 11% cars, 5% goods vehicle and 3% light commercial vehicles in India. Based on the same data and using weighted average a total of 314 two wheeler, 53 four wheeler and 35 truckers were studied. The total respondent surveyed includes 91 from Delhi, 85 from Faridabad, 66 from Ghaziabad, 61 from Gurgaon and 99 from Noida.

### 3.10 MODELS AND TECHNIQUES

The various models and technique used in the study to are as follows:

#### 3.10.1 Fishbein’s Multi Attribute Attitude Model

Fishbein’s Multi Attribute Attitude Model measure three component of the attitude namely salient belief consumer have about an attitude, the probability that a particular objective has an important attribute and evaluation of each of the important attributes. This model can be written mathematically as-

\[ A_0 = \sum b_i e_i, \text{ where } i=1,2,3,\ldots n \]
Where

\[ A_0 = \text{the persons overall attitude towards the object.} \]
\[ b_i = \text{the strength of his belief that the object is related to attribute i} \]
\[ e_i = \text{his evaluation of feelings (liking/disliking) towards attribute i} \]
\[ n = \text{the number of relevant belief for that person} \]

This model has been used to measure the attitude of the consumer towards NFR outlets.

The other tools and techniques that have been used to carry out the research are

### 3.10.2 Factor Analysis

Factor analysis is a general name denoting a class of procedure primarily used for data reduction and summarization. In this method relationship among set of many interrelated variables are examined and represented in the terms of few underlying factors. Factor analysis is an interdependence technique and in this technique the entire set of interdependence relationship is examined. (Malhotra Naresh K., Dash Satyabhusan).

**Factor Analysis Model**

In factor analysis each variable is expressed as a linear combination of underlying factors. The amount of variance a variable share with all other variables included in the analysis is known as communality. The covariation among the variable is describe in terms of small number of common factors plus a unique factor for each variable. The general representation of factor analysis is
\[ X_i = A_{i1}F_1 + A_{i2}F_2 + A_{i3}F_3 + \ldots + A_{im}F_m + V_iU_i \]

Where

\[ X_i = \text{i}^{th} \text{ standardized variable} \]

\[ A_{ij} = \text{Standardized multiple regression coefficient of variable i on common factor j} \]

\[ F = \text{Common factor} \]

\[ V_i = \text{Standardized regression coefficient of variable i on unique factor i} \]

\[ U_i = \text{The unique factor for variable i} \]

\[ m = \text{Number of common factors} \]

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

\[ F_i = W_{i1}X_1 + W_{i2}X_2 + W_{i3}X_3 + W_{i4}X_4 + \ldots + W_{ik}X_k \]

Where

\[ F_i = \text{Estimate of i}^{th} \text{ factor} \]

\[ W_i = \text{Weight or factor score coefficient} \]

\[ k = \text{Number of variables} \]

This method has been used to find out various factors affecting selection of the NFR outlets.
3.10.3 Chi-Square test

Chi-square test is appropriate for situations in which a test for the difference between samples is required. It is used for nominal and the ordinal data. This test helps us in testing the significance of relationship between two variables.

\[ \chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{r} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \]

Where

- \( O_{ij} \) = observed number of cases categorized in the \( (i,j) \)th cell
- \( E_{ij} \) = Expected number of cases under \( H_0 \) to be categorized in the \( (i,j) \)th cell.

This test has been used to compare the usage pattern of NFR services among two wheeler, four wheeler and truckers.

The significance of difference between the frequencies of use of NFR services among different kind of vehicle was also done on the basis of chi-square test.

3.10.4 Test of Normality

Normality test was used to test whether the data follows normal distribution or not. For this first the graphical distribution (Histogram) of the sample data was obtained. If the shape of the curve obtained is bell shaped, then the data is said to follow normal distribution.

Frequency test like Shapiro-Wilk and Kolmogorov-Smirnov test was used to test the normality of the data. These test tests the null hypothesis the samples comes from a normally distributes population.
In the present study these tests has been used to test whether the samples of the second survey were normally distributed or not. The samples were found to be normally distributed.

3.10.5 Two independent sample Z- Test

The two independent sample Z-test accesses whether the mean of two samples differ significantly or not.

The general formula for this test is

\[ Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \]

Where

\( \bar{X}_1 \) = mean of first sample

\( \bar{X}_2 \) = Mean of second sample

\( \sigma_1 \) = Standard deviation of the first sample

\( \sigma_2 \) = Standard deviation of the second sample

\( n_1 \) = Sample size of first sample

\( n_2 \) = Sample size of second sample

This test has been used to compare the attitude of commercial and non-commercial vehicles. Large sample Z test was also used to compare the significance of difference between the attitude of primary shoppers and the non-primary shoppers.

Before using this test the normality of the data has been checked.
3.10.6 One Way ANOVA

The analysis of variance technique helps to draw inferences whether the samples have been drawn from populations having same means. The ANOVA technique investigates any number of factors which are supposed to influence the dependant variable of interest. The dependant variable is metric (interval or ratio scale), whereas the independent variables are categorical (nominal). If there is one independent variable (one factor) divided into various categories, we have One Way ANOVA. In ANOVA, it is assumed that each of the samples is drawn from a normal population and each of the population has an equal variance. Basically two estimates of the population variance are made. One estimate is based upon between the samples and the other one is based upon within the samples. The two estimates of variances can be compared for their equality using F statistic.

The formula for the one-way ANOVA F-test statistic is

\[
F = \frac{\text{explained variance}}{\text{unexplained variance}}.
\]

Or

\[
F = \frac{\text{between-group variability}}{\text{within-group variability}}.
\]

The "explained variance", or "between-group variability" is

\[
\sum_i n_i (\bar{Y}_i - \bar{Y})^2 / (K - 1)
\]

Where \( \bar{Y}_i \) denotes the sample mean in the \( i^{th} \) group, \( n_i \) is the number of observations in the \( i^{th} \) group, and \( \bar{Y} \) denotes the overall mean of the data.
The "unexplained variance" or "within-group variability" is

\[ \sum_{ij}(Y_{ij} - \bar{Y}_i)^2/(N - K), \]

Where \( Y_{ij} \) is the \( j^{th} \) observation in the \( i^{th} \) out of \( K \) group and \( N \) is the overall sample size. This F-statistic follows the F-distribution with \( K - 1, N - K \) degrees of freedom under the null hypothesis. The statistic will be large if the between-group variability is large relative to the within-group variability, which is unlikely to happen if the population means of the groups all have the same value.

This technique has been used to compare the attitude of owner/drivers of two wheeler, four wheeler, and truckers and to compare the attitude between different age groups of consumer.

Using the research methodology discussed in this chapter, data was obtained from the consumer. The obtained data were analyzed with the help of SPSS 16 software.