CHAPTER 3

RESEARCH METHODOLOGY

3.1 RESEARCH PROBLEM

The proposed research will identify the behavior of entrepreneurs or managers of the pharmaceutical firms towards pricing of their products within given structure of Indian pharmaceutical market and industry.

3.2 OBJECTIVES

In the above context, the proposed research aims at analyzing:

- To present an overview of Indian Pharmaceutical Industry and markets and to strategize means to identify anti-competitive activities prevalent in the pharmaceutical market.
- To know the behaviour of the managers for pricing decision
- To know the behavior of manager towards the competitive structure.
- Review situation for the firm whether it is price leader / price follower in the industry.
- To identify pricing strategies adopted by the firms.

3.3 HYPOTHESIS

Hypothesis 1 – Profit of the firm is depends upon sale, price and the market share of the firm.

Hypothesis 2 – Prices of the firm are influenced by sales, firm size, market concentration and experience of the firm in the market.

Hypothesis 3 – Prices are influenced by the cost and revenue of the firm.
Hypothesis 4 – There is no significant difference between sales turnover and cost of production of firms over past three years (2009 – 2011).

3.4 PERIOD OF STUDY

As far as the period of study is concern the data has been used for last three financial years that is March’ 2009, March’2010, and March’2011. For collecting the data related to the sales turnover, market share, profit and loss statement of the pharma companies.

3.5 THE SAMPLE

The Sampling frame used was the ORGIMS Company list of Pharmaceutical companies for which ORGIMS Company conducts a retail audit in India. This ORG Retail audit data is also used by the Government of India as it is the only authentic data regarding Indian Pharmaceutical Industry available in India. The ORGIMS list contains 450 Pharmaceutical companies.

The sampling method chosen is simple random sampling which is a type of probability sampling. To calculate the sample size following formula was used. If the researcher plans the results in a variety of ways or if he/she has difficulty in estimating the proportion or standard deviation of the attribute of interest, the following formula may be more useful.

In this study we determine 176 samples at 95% confident and accept 05% probability error, this calculation from Taro Yamanee formula.

\[ n = \frac{N}{1 + Ne^2} \]  

\( n = \text{sample} \)  
\( N = \text{population} \)
\[ e^2 = \text{probability of error} \]

\[
n = \frac{315}{1 + 315 (0.05)^2} \tag{3.4.2}
\]

\[ n = 176 \tag{3.4.3} \]

### 3.6 Questionnaire Design

The questionnaire was designed to match with the objectives of study and conceptual framework. A short questionnaire with conceptually clear and concise statements is judged to be desirable for both the respondents and the researcher.

The questionnaire consists of a series of questions that show in Appendix B. To ensure the accuracy, the questionnaire was developed as follows. Based on literature review and some major related researches mention in chapter 1, identify factors related to the study. Draft questionnaires based on identified factors will be developed.

### 3.7 Measurement of Conducting Questionnaire

Questionnaires with 5 point rating scale were used to measure respondents’ evaluation by asking them the degree of importance with statements in the questionnaire that ranked from (1) not at all important (or extremely unfavorable) to (5) very important (or extremely favorable). Each question consisted of many factors that mentioned above based on literature review, especially two model, internal model and five force model, to capture the construct of interest. The higher the score the more important the variables are as evaluative criteria. 5 point scales will be used to measure factors in a way such that mean scores could be calculated to show which factors have the most impact on Pharmaceutical Industry. With 5 point scales, the interval for breaking the range in measuring each variable is calculated by:
\[
\frac{5 - 1}{5} = 0.8
\]

It means items with scores fall between the ranges of:

- 4.20 - 5.00 are considered as the most important (favorable) level
- 3.40 - 4.19 are considered as the high important (favorable) level
- 2.60 – 3.39 are considered as the medium important (favorable) level
- 1.80 – 2.59 are considered as the low important (favorable) level

### 3.8 SOURCES OF DATA

#### 3.8.1 THE SECONDARY DATA

The secondary data used in the research was collected from varied sources and compiled as per the requirement of the study. Secondary data was explored through channels as follow: NPPA, Capitaline Database, Food and Drug Control Department, Gujarat.

#### 3.8.2 THE PRIMARY DATA

Interviews were conducted subject to policy makers in prices and structure of the Pharmaceutical company and The food and Drug Control Department, Gujarat. The time for interview was May’2012 to July’2012. The frame of main topics for interview is shown in Appendix 1.

Questionnaire was sent to 315 companies. Out of that 176 companies has responded for the same. These companies were sent the Questionnaire, which was structured in nature. These companies were located all over India although concentrated mostly in Mumbai, Ahmedabad, NCR (Delhi and Adjoining areas),
Hyderabad, Goa, Chennai etc. Questionnaires were sent through E-mail and follow up was done on telephone, E-mail etc. The Questionnaire was filled mostly by Top Management persons like Presidents, Directors and also by General Managers (Marketing). Questionnaire Testing & Reliability Analysis: The questionnaire was pretested on 315 companies and data was fed into SPSS software. Using SPSS reliability analysis was conducted. Two renowned Pharmaceutical Industry Market research experts were contacted.

The Experts approved the questionnaire with small modifications. The Questionnaire was suitably modified and a couple of questions were deleted and one question was added. The revised questionnaire was administered and data was collected from 176 pharmaceutical companies in India.

3.9 METHODOLOGY

3.9.1 SURVEY

A three page, two sided Questionnaire was designed keeping in mind the objectives of the study which were to find out the impact of Pricing Behavior of the Indian pharmaceutical firms and to find out the change in marketing strategies of pharmaceutical industry after implementation of the different market structure. The Literature survey and pre study consultation with industry experts were taken into account. The questionnaire consisted of few open ended questions, some questions were either using ranking scale or Likert scale, however the questions related to this paper on role of Government of India were open ended questions.
3.9.2 Data Analysis

Primary data conducted by interviews and questionnaires were used for supporting to analyze key determine factors of Pharmaceutical Industry in India. Data from questionnaires were processed in terms of frequency, mean, standard deviation (descriptive statistic).

3.9.3 Factor Analysis

Factor analysis is a statistical method used to describe variability among observed variables in terms of a potentially lower number of unobserved variables called factors. In other words, it is possible, for example, that variations in three or four observed variables mainly reflect the variations in a single unobserved variable, or in a reduced number of unobserved variables. Factor analysis searches for such joint variations in response to unobserved latent variables. The observed variables are modelled as linear combinations of the potential factors, plus "error" terms. The information gained about the interdependencies between observed variables can be used later to reduce the set of variables in a dataset. Factor analysis originated in psychometrics, and is used in behavioral sciences, social sciences, marketing, product management, operations research, and other applied sciences that deal with large quantities of data. Factor analysis is related to principal component analysis (PCA), but the two are not identical. Because PCA performs a variance-maximizing rotation of the variable space, it takes into account all variability in the variables. In contrast, factor analysis estimates how much of the variability is due to common factors ("communality"). The two methods become essentially equivalent if the error terms in the factor analysis model (the variability not explained by common factors,) can be assumed to all having the same variance.

The analysis will isolate the underlying factors that explain the data. Factor analysis is an interdependence technique. The complete set of interdependent relationships is examined. There is no specification of dependent variables, independent variables, or causality. Factor analysis assumes that all the rating data on different attributes can be reduced down to a few important dimensions. This
reduction is possible because the attributes are related. The rating given to any one attribute is partially the result of the influence of other attributes. There are two approaches to factor analysis: "principal component analysis" (the total variance in the data is considered); and "common factor analysis" (the common variance is considered) Note that principal component analysis and common factor analysis differ in terms of their conceptual underpinnings. The factors produced by principal component analysis are conceptualized as being linear combinations of the variables whereas the factors produced by common factor analysis are conceptualized as being latent variables. Computationally, the only difference is that the diagonal of the relationships matrix is replaced with communalities (the variance accounted for by more than one variable) in common factor analysis. This has the result of making the factor scores indeterminate and differs depending on the method of computation.

Meanwhile, factor scores produced by principal component analysis are not dependent on the method of computation. Although there have been heated debates over the merits of the two methods, a number of leading statisticians have concluded that in practice there is little difference (Velicer and Jackson, 1990) which makes sense since the computations are quite similar despite the differing conceptual bases, especially for datasets where communalities are high and/or there are many variables, reducing the influence of the diagonal of the relationship matrix on the final result (Gorsuch, 1983).

Factor Analysis is a very useful method of reducing data complexity by reducing the number of variables being studied. It reduces the number of variables in to fewer factors by analyzing the correlation between variables, which explains much of the original data, more economically. There are two stages in factor analysis, first is factor extraction process and second is rotation method. In the first stage factors are extracted with the help of principal components analysis. There is a rule of thumb to determine the number of factors to be extracted based on the computation to the Eigen value. The higher Eigen values of the factor, higher the amount of variance explained by the factor. The purpose of the factor analysis is to
extract the least number of factors possible, which will maximize the explained variance.

Second stage is to interpret and name the factors. This is done by the process of identifying where factors are associated with the original variables; the rotated factor matrix is used for the purpose. The factor matrix gives the loading of each variable on each of the extracted factors. This is similar to correlation matrix with loading having values between 0 and 1. The purpose is to find variables which have a high loading on one factor, but a low loading on other factors. Thereafter the combination of variables that is highly loading on other factors. Thereafter the combination of variables that is highly loaded on one factor is given a suitable name with the assumption that the factor so selected is the linear combination of the high load variables (Marketing Research, Rajendra Nargundkar, chapter 12)

3.9.4 Analysis of Variance

Analysis of variance (ANOVA) is a collection of statistical models in statistics, and their associated procedures, in which the observed variance in a particular variable is partitioned into components attributable to different sources of variation. In its simplest form ANOVA provides a statistical test of whether or not the means of several groups are all equal, and therefore generalizes t-test to more than two groups. Doing multiple two-sample t tests would result in an increased chance of committing a type I error. For this reason, ANOVAs are useful in comparing two, three or more means.

In statistics, one purpose for the analysis of variance (ANOVA) is to analyze differences in means between groups. The test statistic, \( F \), assumes independence of observations, homogeneous variances, and population normality. ANOVA on ranks is a statistic designed for situations when the normality assumption has been violated.
One-way analysis of variance (ANOVA) tests allow you to determine if one given factor, such as drug treatment, has a significant effect on gene expression behavior across any of the groups under study. A significant p-value resulting from a 1-way ANOVA test would indicate that a gene is differentially expressed in at least one of the groups analyzed. If there are more than two groups being analyzed, however, the 1-way ANOVA does not specifically indicate which pair of groups exhibits statistical differences. Post Hoc tests can be applied in this specific situation to determine which specific pair/pairs are differentially expressed. This document will provide the necessary information for you to perform these analyses within GeneSpring.

3.9.5 CORRELATION

Used the Correlation transformer to determine the extent to which changes in the value of an attribute (such as length of employment) are associated with changes in another attribute (such as salary). The data for a correlation analysis consists of two input columns. Each column contains values for one of the attributes of interest. The Correlation transformer can calculate various measures of association between the two input columns. You can select more than one statistic to calculate for a given pair of input columns.

The data in the input columns also can be treated as a sample obtained from a larger population, and the Correlation transformer can be used to test whether the attributes are correlated in the population. In this context, the null hypothesis asserts that the two attributes are not correlated, and the alternative hypothesis asserts that the attributes are correlated.

3.9.6 CORRELATION COEFFICIENT r

The correlation coefficient \( r \) is a measure of the linear relationship between two attributes or columns of data. The correlation coefficient is also known as the Pearson product-moment correlation coefficient. The value of \( r \) can range from -1 to
+1 and is independent of the units of measurement. A value of \( r \) near 0 indicates little correlation between attributes; a value near +1 or -1 indicates a high level of correlation.

When two attributes have a positive correlation coefficient, an increase in the value of one attribute indicates a likely increase in the value of the second attribute. A correlation coefficient of less than 0 indicates a negative correlation. That is, when one attribute shows an increase in value, the other attribute tends to show a decrease.

### 3.9.7 Rating Scales

Rating scales are a controversial middle case. The numbers in rating scales have meaning, but that meaning isn't very precise. They are not like quantities. With a quantity (such as dollars), the difference between 1 and 2 is exactly the same as between 2 and 3. With a rating scale, that isn't really the case. You can be sure that your respondents think a rating of 2 is between a rating of 1 and a rating of 3, but you cannot be sure they think it is exactly halfway between. This is especially true if you labeled the mid-points of your scale (you cannot assume "good" is exactly halfway between "excellent" and "fair"). Most statisticians say that you cannot use correlations with rating scales, because the mathematics of the technique assumes the differences between numbers which are exactly equal. Nevertheless, many survey researchers do use correlations with rating scales, because the results usually reflect the real world. Our own position is that you can use correlations with rating scales, but you should do so with care. When working with quantities, correlations provide precise measurements. When working with rating scales, correlations provide general indications.