CHAPTER – 3

RESEARCH METHODOLOGY

3.1 Introduction

Managers in the service sector are under increasing pressure to demonstrate that their services are customer-focused and that continuous performance improvement is being delivered. Given the financial and resource constraints under which service organizations must manage it is essential that customer expectations are properly understood and measured and that, from the customers' perspective, any gaps in service quality are identified. This information then assists a manager in identifying cost-effective ways of closing service quality gaps and of prioritizing which gaps to focus on – a critical decision given scarce resources. Satisfaction which is vaguely defined as fulfilling the needs for which a good or service was made is viewed differently in various industries, over various demographic backgrounds, as well as for individuals and institutions. Moreover, it has a totally different approach when it comes to services and products. We have been trying to understand quality of services and satisfaction both in the area of comfort and in terms of utility that is, the product or service fulfilling the actual purpose for which it was made and bought. Bailey (1983) identified 38 factors that affected the satisfaction of consumers of computers which are customized for computer users some of which were quality of the product, flexibility, reliability, priorities determination, security and expectations. In online education structure, transparency and communication potentials influence the satisfaction of students and enhance the learning process, (Karen, 2001).

Sahim (2006) in an effort to find out whether customers were satisfied with the food services in the military hospital in Turkey realized that specific demographic characteristics were not of significance in determining the satisfaction of the patients but the appearance and taste of food. Their emphasis on demographic characteristics gives the reader the impression that they thought it was going to be an important factor. Another study in Jiangsu province, China seeking to find out the differences in food preferences between students of different socio-demographic backgrounds and characteristics stated in their literature that societal and cultural
factors as well as environmental and indigenous factors shape children’s food choices, Shi. (2005). This makes them appreciate food quality differently and often because they are not used to it, or they do not like it at all or because of some traditional beliefs associated with the different demographic characteristics.

It has however been identified that human needs, quality of services and products, the user friendly nature of product and services, and comfort assurance (Bailey, 1983) are some of the important determinants of customer satisfaction. Even though different customers will require different levels and combinations of these variables, they generally are important factors that affect customer satisfaction. Matzler (2002) went a step forward to classify factors that affect customers’ satisfaction into three categories:

1. **Basic factors**: These are the minimum requirements that are required in a product to prevent the customer from being dissatisfied. They do not necessarily cause satisfaction but lead to dissatisfaction if absent. These are those factors that lead to the fulfilment of the basic requirement for which the product is produced. These constitute the basic attributes of the product or service.

2. **Performance factors**: These are the factors that lead to satisfaction if fulfilled and can lead to dissatisfaction if not fulfilled. These include reliability and friendliness.

3. **Excitement factors**: These are factors that increase customers’ satisfaction if fulfilled but does not cause dissatisfaction if not fulfilled which include project management.

It is however a little contradiction but it is a depiction of the complexities in the concept of satisfaction that some researchers seek to explain. Indian hotels are now facing acute competition as different demands of the customers in relation to accommodation, prompt information, food quality etc. provided by hotels/workforces of the hotels. Therefore, the present study entitled “A study on Customer Profiling, Attribute Evaluation and Customer Loyalty in the Loading Industry” to identifying dominant factors which affect the performance of customers in loading industry. Hence, the study sets out to analyze the service quality dimensions of hotel guest using the SERVQUAL instrument as base in the emerging Indian market.
Applications of Factor Analysis

A. Exploratory factor analysis

A non-theoretical application. Given a set of variables, what are the underlying dimensions (factors), if any, that account for the patterns of co-linearity among the variables? Example Given the multiple items of information gathered on applicants applying for admission to a police academy, how many independent factors are actually being measured by these items?

B. Confirmatory factor analysis

Given a theory with four concepts that purport to explain some behaviour, do multiple measures of the behaviour reduce to these four factors? Example Given a theory that attributes delinquency to four independent factors, do multiple measures on delinquents reduce to measuring these four factors?

C. R and Q Factor Analysis

R factor analysis involves extracting latent factors from among the variables Q factor analysis involves factoring the subjects vis-à-vis the variables. The result is a "clustering" of the subjects into independent groups based upon factors extracted from the data. This application is not used much today since a variety of clustering techniques have been developed that are designed specifically for the purpose of grouping multiple subjects into independent groups.
Undertaking Factor Analysis

The starting point in factor analysis, as with other statistical techniques, is the research problem. In this note we will assume that our question involves exploratory research and involves a requirement to condense the data from a number of variables into a smaller set of dimensions with a minimum loss of information. Factor analysis can identify the structure of relationships among either variables or respondents by examining either the correlations between the variables or the correlations between the respondents. For example, suppose we have data on 100 respondents in terms of 10 characteristics. If the objective of the research is to summarize the characteristics, the factor analysis is applied to a correlation matrix of the variables. This is the most common type of factor analysis, and is referred to as R factor analysis. R factor analysis analyzes a set of variables to identify the underlying dimensions Factor analysis also may be applied to a correlation matrix of the individual respondents based on their characteristics. This is referred to as Q factor analysis, a method of combining or condensing large numbers of people into distinctly different groups within a larger population. The Q factor analysis approach is not utilized very frequently. This note will focus on R factor analysis

Variable Selection

Once the purpose of factor analysis is specified, the researcher must then define the set of variables to be examined. The researcher implicitly specifies the potential dimensions that can be identified through the character and nature of the variables submitted to factor analysis. For example, in assessing the dimensions of store image, if no questions on store personnel were included, factor analysis would not be able to identify this dimension. The use of factor analysis as a data summarization technique does not exclude the need for a conceptual basis for any variables analyzed.
Designing a Factor Analysis

The design of a factor analysis involves three basic decisions:

(1) Choice of the input data (a correlation matrix) to meet the specified objectives of grouping variables or respondents;

(2) the design of the study in terms of number of variables, measurement properties of variables, and the types of allowable variables; and

(3) the sample size necessary, both in absolute terms and as a function of the number of variables in the analysis.

Correlations among Variables or Respondents

The first decision in the design of a factor analysis focuses on the choice of the correlation matrix to be used. The researcher could derive the input data matrix from the computation of correlations between the variables. This would be an R-type factor analysis. Alternatively the researcher could also elect to derive the correlation matrix from the correlations between the individual respondents. In this Q-type factor analysis, the results would be a factor matrix that would identify similar individuals. In this note it is assumed that R-type will be used and that Cluster Analysis would be used if relationships between individuals is the focus of the research.

Variable Selection and Measurement Issues

Two specific questions must be answered at this point: (1) How are the variables measured? and (2) How many variables should be included? Variables for factor analysis are generally assumed to be of metric measurement. In some cases, dummy variables (coded 0-1), although considered non-metric, can be used. The researcher should also attempt to minimize the number of variables included but still maintain a reasonable number of variables per factor. If a study is being designed to assess a proposed structure, the researcher should be sure to include several variables (five or more) that may represent each proposed factor. The strength of factor analysis lies in finding patterns among groups of variables, and it is of little use in identifying factors composed of only a single variable. Finally, when designing a study to be factor analyzed, the researcher should, if possible, identify several key variables (sometimes referred to as key indicants or marker variables) that closely reflect the hypothesized underlying factors. This will aid in validating the derived factors and assessing whether the results have practical significance.
Assumptions in Factor Analysis

In addition to the statistical bases for the correlations of the data matrix, the researcher must also ensure that the data matrix has sufficient correlations to justify the application of factor analysis. If visual inspection reveals no substantial number of correlations greater than .30, then factor analysis is probably inappropriate. The correlations among variables can also be analyzed by computing the partial correlations among variables, that is, the correlations between variables when the effects of other variables are taken into account. If "true" factors exist in the data, the partial correlation should be small, because the variable can be explained by the factors. If the partial correlations are high, then there are no "true" underlying factors, and factor analysis is inappropriate.

Another mode of determining the appropriateness of factor analysis examines the entire correlation matrix. The ‘Bartlett test of sphericity’, a statistical test for the presence of correlations among the variables, is one such measure. It provides the statistical probability that the correlation matrix has significant correlations among at least some of the variables. A basic assumption of factor analysis is that some underlying structure does exist in the set of selected variables. It is the responsibility of the researcher to ensure that the observed patterns are conceptually valid and appropriate to study with factor analysis, because the technique has no means of determining appropriateness other than the correlations among variables. The researcher must also ensure that the sample is homogeneous with respect to the underlying factor structure. It is inappropriate to apply factor analysis to a sample of males and females for a set of items known to differ because of gender. When the two sub samples (males and females) are combined, the resulting correlations and factor structure will be a poor representation of the unique structure of each group. Thus, whenever differing groups are expected in the sample, separate factor analyses should be performed, and the results should be compared to identify differences not reflected in the results of the combined sample.

Common Factor Analysis versus Component Analysis

There are two basic models to obtain factor solutions. They are known as common factor analysis and component analysis. To select the appropriate model, the researcher must first understand the differences between types of variance in factor analysis of which there are three types.
(1) **Common variance** is defined as that variance: in a variable that is shared with all other variables in the analysis

(2) **Specific variance** (sometimes called unique) is that variance associated with only a specific variable.

(3) **Error variance** is that variance due to unreliability in the data-gathering process, measurement error, or a random component in the measured phenomenon

Component analysis, also known as Principal Components Analysis, considers the total variance and derives factors that contain small proportions of unique variance and, in some instances, error variance. Specifically, with component analysis, unities (1’s) are inserted in the diagonal of the correlation matrix, so that the full variance is brought into the factor matrix. Conversely, with common factor analysis, communalities are inserted in the diagonal. Communalities are estimates of the shared, or common, variance among the variables. Factors resulting from common factor analysis are based only on the common variance.

The common factor and component analysis models are both widely used. The selection of one model over the other is based on two criteria: (1) the objectives of the factor analysis and (2) the amount of prior knowledge about the variance in the variables. The component factor model is appropriate when the primary concern is about prediction or the minimum number of factors needed to account for the maximum portion of the variance represented in the original set of variables, and when prior knowledge suggests that specific and error variance represent a relatively small proportion of the total variance. In contrast, when the primary objective is to identify the latent dimensions or constructs represented in the original variables, and the researcher has little knowledge about the amount of specific and error variance and therefore wishes to eliminate this variance, the common factor model is most appropriate. Common factor analysis suffers from the problem that several different factor scores can be calculated from the factor model results in other words there is no single unique solution, as found in component analysis. When a decision has been made on the factor model, the researcher is ready to extract the initial un-rotated factors. By examining the un-rotated factor matrix, the researcher can explore the data reduction possibilities for a set of variables and obtain a preliminary estimate.
of the number of factors to extract. Final determination of the number of factors must wait, however, until the results are rotated and the factors are interpreted.

Steps in Factor Analysis

Step 1 Compute inter-correlation matrix. Compute the factorability of the matrix.

Step 2 Extract an initial solution

Step 3 From the initial solution, determine the appropriate number of factors to be extracted in the final solution

Step 4 If necessary, rotate the factors to clarify the factor pattern in order to better interpret the nature of the factors

Step 5 Depending upon subsequent applications, compute a factor score for each subject on each factor.

Criteria for the Number of Factors to Extract

The analogy for choosing the number of factors to be interpreted with focusing a microscope. Too high or too low an adjustment will obscure a structure that is obvious when the adjustment is just right. Therefore, by examining a number of different factor structures derived from several trial solutions, the researcher can compare and contrast to arrive at the best representation of the data. They suggest the following stopping criteria for the number of factors to be extracted:

Latent Root Criterion The most commonly used technique is the latent root criterion. This technique is simple to apply to either components analysis or common factor analysis. The rationale for the latent root criterion is that any individual factor should account for the variance of at least a single variable if it is to be retained for interpretation. Each variable contributes a value of 1 to the total Eigen-value. Thus, only the
factors having latent roots or Eigen values greater than 1 are considered significant; all factors with latent roots less than 1 are considered insignificant and are disregarded.

**A Priori Criterion** This can be useful when the researcher already knows how many factors to extract before undertaking the factor analysis. The researcher simply instructs the computer to stop the analysis when the desired number of factors has been extracted. This approach is useful when testing a theory or hypothesis about the number of factors to be extracted. It also can be justified in attempting to replicate another researcher's work and extract the same number of factors that was previously found.

**Percentage of Variance Criterion** The percentage of variance criterion is an approach based on achieving a specified cumulative percentage of total variance extracted by successive factors. The purpose is to ensure practical significance for the derived factors by ensuring that they explain at least a specified amount of variance. No absolute threshold has been adopted for all applications.

**Scree Test Criterion** With the component analysis factor model, the factors extracted contain both common and unique variance. Although all factors contain at least some unique variance, the proportion of unique variance is substantially higher in later than in earlier factors. The Scree test is used to identify the optimum number of factors accounts for 60 percent of the total variance (and in some instances even less) as satisfactory. That can be extracted before the amount of unique variance begins to dominate the common variance structure. The Scree test is derived by plotting the latent roots against the number of factors in their order of extraction, and the shape of the resulting curve is used to evaluate the cut-off point. The figure below shows an example. Starting with the first factor, the plot slopes steeply downward initially and then slowly becomes an approximately horizontal line. The point at which the curve first begins to straighten out is considered to indicate the maximum number of factors to extract. In the present case, the first 10 factors would qualify. Beyond 10, too large a proportion of unique variance would be included; thus these factors would not be acceptable.
Concluding points on selection criterion

Firstly In practice, most researchers seldom use a single criterion in determining how many factors to extract. Secondly some words of caution about selecting the final set of factors. If too few factors are used, then the correct structure is not revealed, and important dimensions may be omitted. If too many factors are retained, then the interpretation becomes more difficult when the results are rotated.

Objectives of the study

The present study is based on the following objectives in order to get the empirical results on customer satisfaction in Indian hotels:

1. To identify the benefits and facilities sought by customers of Indian hotel industry.
2. To segment and profile the market for hotel industry based on psychographic, demographic and behavioural variables.
3. To review the previous studies those had conducted in this connection in order get more knowledge of customer satisfaction in hotel industry.
4. To describe the service quality dimensions in hotel business those affect the customer satisfaction using SERVQUAL model as base.
5. To separate the dominant service quality dimensions that influence customer satisfaction in Indian hotels.
6. To decide statistically the performance rank of service quality dimensions as per customers satisfaction through factor analysis.
7. To conclude the empirical facts after using factor analysis through SPSS and provide key points for betterment/satisfaction of guest in Indian hotel industry.
**Hypotheses**

The following are the statistical hypotheses of the study:

- **H₀₁**: The performance of Service Quality dimensions is same on customer satisfaction in 4 star hotels.
- **H₀₂**: The performance of Service Quality dimensions is same on customer satisfaction in 5 star Indian hotels.
- **H₀₃**: The performance of Service Quality dimensions is same on customer satisfaction in 5 star deluxe hotels.
- **Hₐ**: The performances of Service Quality dimensions are not same on customer satisfaction in three types of hotels under consideration by the study.

**The Questionnaire**

The study under consideration used modified SERVQUAL as base to measure the customer satisfaction in Indian hotels. The SERVQUAL model as the basis for the structured questionnaire because it provides information on our research questions in which we are trying to know how consumers perceive service quality in Indian hotels. In order to get the response of guests on service provided by hotels, a questionnaire is designed contains two parts, 1ˢᵗ part of the questionnaire is for **General and Demographic information**. The demographic part provides general information of respondents on age of guest, gender, education, duration of stay in hotel etc and 2ⁿᵈ part of the questionnaire is for Service perceived by guests/provided by the hotels. The quality of services is measured by 7 dimensions (Tangibles, Reliability, Responsiveness, Assurance, Empathy, Reputation and Security). There are 41 statements, which were directed to measuring service quality in the hotels in our case. The following abbreviations are used to represent the service quality dimensions.
Administering of questionnaires

As mentioned above a closed-ended questionnaire was circulated for collecting the responses of customers on service quality offered by Indian hotels. The study is using a convenience sampling technique to get the responses fit for the study. To judge the quality of services of Indian hotels, the hotels are divided into 3 categories i.e. 4 Star Hotels, 5 Star Hotels and 5 Star Deluxe Hotels. 520 questionnaires were circulated among the guests of 4 star hotels out of which 15 respondents did not indicate stay period in the hotels being rejected, 58 respondents did not completely answer the questions which were excluded from sample and considered them invalid, 23 respondents returned the questionnaires with double tick on 5 items which were not part of sample size, and 77 questionnaires had not return till the analysis. This is because some people got the questionnaires and went away with them. Thus, 351 questionnaires were finally included as sample size $n = 351$ for 4 star hotels. 1500 questionnaires were circulated in case of 5 star hotels in India out of which 29 questionnaires were not approached again, 22 questionnaires had not return with complete information, 21 questionnaires were not fill-up properly which were excluded from sample size. Finally, 1428 questionnaires were included as sample size $n = 1428$ for factor analysis for 5 star hotels. 1300 questionnaires were circulated among the customers of 5 star deluxe hotels out of which 63 respondents did not tick on all survey items which were excluded from sample, 39 respondents returned the questionnaires with double tick on many items which were not part of sample size, and 93 questionnaires had not return till the analysis. Therefore, unfortunately we received 1105 questionnaires that were complete. Hence, 1105 questionnaires were became as sample size $n = 1105$ for 5 star deluxe hotels.
Measurement

The SERVQUAL model as base is used to assess consumers’ perceptions regarding service quality in three types of hotels. The perceptions of customers (guest) are measured using a 5-point scale to rate their level of agreement or disagreement (1 for strongly disagree and 5 for strongly agree), on which the higher numbers indicate higher level of perceptions which means higher level of satisfaction. The following scales are used in study for service quality provided by three types of hotels in part-2 of the said questionnaire:

<table>
<thead>
<tr>
<th>Level of Agreement</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strongly Disagree</td>
<td>1</td>
</tr>
<tr>
<td>Disagree</td>
<td>2</td>
</tr>
<tr>
<td>Neutral</td>
<td>3</td>
</tr>
<tr>
<td>Agree</td>
<td>4</td>
</tr>
<tr>
<td>Strongly Agree</td>
<td>5</td>
</tr>
</tbody>
</table>

Perceptions are based on the actual service they received in hotels. There is direct relationship between satisfaction and service quality provided by hotels. Consumer’s satisfaction judgments are the result of consumer’s perceptions. The developed questionnaire includes forty one (41) items where four items correspond to the tangibles dimension, five items correspond to the reliability dimension, seven items to the responsiveness, nine items correspond to the assurance dimensions, seven items to empathy, four items to reputation and five items for security. Respondents were asked to indicate their degree of agreement with each of the items on five-point Likert scale. The Statistical Package for Social Sciences (SPSS) 16.0 will be used to tabulate the data gathered from the research questionnaires.

Coding

The dimensions/items are main variables used in this study and we coded these dimensions/items in order to ease our analysis of data collected. Also, demographic information was collected from respondents and these variables have to be coded as well for analysis. Here is the coding of the variables for analysis.
TANGIBLES (TA)

TA1 Ideal hotel should have modern equipments.

TA2 Physical facilities provided by hotels

TA3 The physical environment of the hotel should be clean.

TA4 Restaurants and bars of the hotel

RELIABILITY (RL)

RL1 When hotel promise to do something by a certain time, they should do so.

RL2 When a customer has a problem, hotel should show a sincere interest in solving it.

RL3 Hotel should perform the service right the first time.

RL4 They should provide their services at the time they promise to do so.

RL5 They should keep their records accurately.

RESPONSIVENESS (RN)

RN1 Employees should make information easily obtainable by the customers.

RN2 Employees should give prompt service to customers.

RN3 Employees are always willing to help customers.

RN4 Employees in a hotel should never be too busy to respond to customers' requests.

RN5 Speed of the service availed by the customer

RN6 Response for the demand in the hotel.

RN7 Giving information offering for the service

ASSURANCE (AS)

AS1 The behaviour of employees in hotel should instil confidence in customers.

AS2 Customers should be able to feel safe in their transactions with employees in the hotel.
AS3 Their employees should be well dressed and appear neat.

AS4 Their employees should be polite.

AS5 The attitude of the employees of this hotel shows me that they understand my needs.

AS6 The price of the service being charged by the hotel.

AS7 The employees of this hotel are able to answer my questions quickly and are problem solving.

AS8 The employees of this hotel are competent.

AS9 Employees of hotel should have the knowledge to answer customers' questions.

EMPATHY (EM)

EM1 Hotel should give customers individual attention.

EM2 Their operating hours should be convenient to all their customers.

EM3 Employees should give customers personal service.

EM4 They should have their customers' best interest at heart.

EM5 The employees should understand the specific needs of their customers.

EM6 Customer Insurance policy of the Hotel.

EM7 Hotel provision for the customer’s necessities.

REPUTATION (RE)

RE1 The retails stores around this hotel are conveniently located.

RE2 The layout of this hotel makes it easy for me to move around.

RE3 Booking of the room can be done at International level also.

RE4 All functioning in hotel as per government policy.

SECURITY (SE)

SE1 There are accessible fire exits at this hotel.

SE2 The medical facilities provided by doctors in hotel.
Reliability Analysis of the Survey Instrument

The study tested the reliability of the said questionnaire in three types of hotel using Reliability Coefficient (Cronbach’s alphas). Reliability analysis allows you to study the properties of measurement scales and the items that make them up. The Reliability Analysis procedure calculates a number of commonly used measures of scale reliability and also provides information about the relationships between individual items in the scale. Intra-class correlation coefficients can be used to compute inter-rater reliability estimates. Reliability refers to the property of a measurement instrument that causes it to give similar results for similar inputs. For example, consider the produce scales at grocery stores. At a given store, each of these scales was probably manufactured at the same factory. You would hope that the factory is reliable - that every scale produced at that factory would register the same weight (within a small margin of error) for the same head of lettuce.

It is important to note that reliability is not just a property of an individual produce scale. When buying groceries, you would expect the particular scale you use to be reliable and to record approximately the same weight when the same item is weighed a second time. However, the reliability of that particular scale is only of immediate importance to the customers using it, while the grocery store is concerned about all of the scales in the store, and the manufacturer is concerned about every scale produced at the factory. The deeper issue here is the reliability of the underlying process of scale manufacture. If that process is reliable, then the manufacturer can be confident that the product is reliable.
Models on Reliability

The following models of reliability are available:

- **Alpha (Cronbach)**: This is a model of internal consistency, based on the average inter-item correlation.
- **Split-half**: This model splits the scale into two parts and examines the correlation between the parts.
- **Guttman**: This model computes Guttman's lower bounds for true reliability.
- **Parallel**: This model assumes that all items have equal variances and equal error variances across replications.
- **Strict parallel**: This model makes the assumptions of the parallel model and also assumes equal means across items.

The study used Alpha (Cronbach) to test the reliability of the questionnaire. Cronbach's alpha (Cronbach, 1951) is a measure of reliability. More specifically, alpha is a lower bound for the true reliability of the survey. Mathematically, reliability is defined as the proportion of the variability in the responses to the survey that is the result of differences in the respondents. That is, answers to a reliable survey will differ because respondents have different opinions, not because the survey is confusing or has multiple interpretations. The computation of Cronbach's alpha is based on the number of items on the survey (k) and the ratio of the average inter-item covariance to the average item variance. The value of reliability coefficient must be more than .5 or close to 1 is considered good indication for reliability.

Under the assumption that the item variances are all equal, this ratio simplifies to the average inter-item correlation, and the result is known as the Standardized item alpha (or Spearman-Brown stepped-up reliability coefficient). Alpha can be calculated as:

$$\left(\frac{k}{k-1}\right)\left(1-\frac{\sum s^2}{s^2}\right)$$

Here k stands for the number of conditions contributing to a total score, and s is the standard deviation, which students have learned to calculate and interpret.
early in the most elementary statistics course. There is an $s_i$ for every column of a $p \times i$ layout (below Table), and an $s_t$ for the column of total scores (usually test scores). The formula was something that students having an absolute minimum of technical knowledge could make use of.

Person x Item Score ($X_{pi}$) Sample Matrix

<table>
<thead>
<tr>
<th>Person</th>
<th>1</th>
<th>2</th>
<th>...</th>
<th>I</th>
<th>...</th>
<th>k</th>
<th>Sum or Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$X_{11}$</td>
<td>$X_{12}$</td>
<td>...</td>
<td>$X_{1I}$</td>
<td>...</td>
<td>$X_{1k}$</td>
<td>$X_1.$</td>
</tr>
<tr>
<td>2</td>
<td>$X_{21}$</td>
<td>$X_{22}$</td>
<td>...</td>
<td>$X_{2I}$</td>
<td>...</td>
<td>$X_{2k}$</td>
<td>$X_2.$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>p</td>
<td>$X_{p1}$</td>
<td>$X_{p2}$</td>
<td>...</td>
<td>$X_{pI}$</td>
<td>...</td>
<td>$X_{pk}$</td>
<td>$X_p.$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>n</td>
<td>$X_{n1}$</td>
<td>$X_{n2}$</td>
<td>...</td>
<td>$X_{nI}$</td>
<td>...</td>
<td>$X_{nk}$</td>
<td>$X_n.$</td>
</tr>
</tbody>
</table>

**Statistical Techniques used – Factor Analysis**

Factor analysis attempts to identify underlying variables, or factors, that explain the pattern of correlations within a set of observed variables. Factor analysis is often used in data reduction to identify a small number of factors that explain most of the variance observed in a much larger number of manifest variables. Factor analysis can also be used to generate hypotheses regarding causal mechanisms or to screen variables for subsequent analysis (for example, to identify co-linearity prior to performing a linear regression analysis). Factor analysis attempts to represent a set of observed variables $X_1, X_2 \ldots X_n$ in terms of a number of 'common' factors plus a factor which is unique to each variable. The common factors (sometimes called latent variables) are hypothetical variables which explain why a number of variables are correlated with each other -- it is because they have one or more factors in common. There are a number of different varieties of factor analysis: the
discussion here is limited to principal axis factor analysis and factor solutions in which the common factors are uncorrelated with each other. It is also assumed that the observed variables are standardized (mean zero, standard deviation of one) and that the factor analysis is based on the correlation matrix of the observed variables.

**Input Data for factor Analysis**

The variables should be quantitative at the interval or ratio level. Categorical data (such as religion or country of origin) are not suitable for factor analysis. Data for which Pearson correlation coefficients can sensibly be calculated should be suitable for factor analysis.

**Karl Pearson’s Correlation Coefficient**

To know the inter-correlation among items and factors, the Pearson coefficient of correlation is calculated as:

\[
 r = \frac{N\sum XY - \sum X \sum Y}{\sqrt{N\sum X^2 - (\sum X)^2} \times \sqrt{N\sum Y^2 - (\sum Y)^2}}
\]

Where

- \( N \) = Total number of observation
- \( X \) = Dependent variable
- \( Y \) = Independent variable
- \( \sum X \) = Sum of Dependent variable
- \( \sum Y \) = Sum of independent variable
- \( \sum X^2 \) = Sum of the square of x variable
- \( \sum Y^2 \) = Sum of the square of y variable
\[ \sum XY = \text{Sum of product of both variables} \]

**Assumptions for factor Analysis**

The data should have a bivariate normal distribution for each pair of variables, and observations should be independent. The factor analysis model specifies that variables are determined by common factors (the factors estimated by the model) and unique factors (which do not overlap between observed variables); the computed estimates are based on the assumption that all unique factors are uncorrelated with each other and with the common factors.

Factor analysis can be used for exploratory or confirmatory purposes.

- **Exploratory factor analysis** seeks to uncover the underlying structure of a relatively large set of variables. The researcher's à priori assumption is that any indicator may be associated with any factor. This is the most common form of factor analysis and one which will be covered in this note.

- **Confirmatory factor analysis** seeks to determine if the number of factors and the loadings of measured (indicator) variables on them conform to what is expected on the basis of pre-established theory.

The starting point in factor analysis, as with other statistical techniques, is the research problem. In this note we will assume that our questions involve exploratory research and involve a requirement to condense the data from a number of variables into a smaller set of dimensions with a minimum loss of information.

**Analytical Procedure for study - The Factor Analysis Model**

If the observed variables are \( X_1, X_2 \ldots X_n \), the common factors are \( F_1, F_2 \ldots F_m \) and the unique factors are \( U_1, U_2 \ldots U_n \), the variables may be expressed as linear functions of the factors:

\[
X_1 = a_{11}F_1 + a_{12}F_2 + a_{13}F_3 + \ldots + a_{1m}F_m + a_{1}U_1
\]

\[
X_2 = a_{21}F_1 + a_{22}F_2 + a_{23}F_3 + \ldots + a_{2m}F_m + a_{2}U_2
\]

\[
\ldots
\]

\[
X_n = a_{n1}F_1 + a_{n2}F_2 + a_{n3}F_3 + \ldots + a_{nm}F_m + a_{n}U_n
\]

Each of these equations is a regression equation; factor analysis seeks to find the coefficients \( a_{11}, a_{12} \ldots a_{nm} \) which best reproduce the observed variables from the
factors. The coefficients $a_{11}$, $a_{12}$ ... $a_{nm}$ are weights in the same way as regression coefficients (because the variables are standardized, the constant is zero, and so is not shown). For example, the coefficient $a_{11}$ shows the effect on variable $X_1$ of a one-unit increase in $F_1$. In factor analysis, the coefficients are called loadings (a variable is said to 'load' on a factor) and, when the factors are uncorrelated, they also show the correlation between each variable and a given factor. In the model above, $a_{11}$ is the loading for variable $X_1$ on $F_1$, $a_{23}$ is the loading for variable $X_2$ on $F_3$, etc.

When the coefficients are correlations, i.e., when the factors are uncorrelated, the sum of the squares of the loadings for variable $X_i$, namely $a_{1i}^2 + a_{12}^2 + ... + a_{13}^2$, shows the proportion of the variance of variable $X_i$ which is accounted for by the common factors. This is called the communality. The larger the communality for each variable, the more successful a factor analysis solution is. By the same token, the sum of the squares of the coefficients for a factor -- for $F_1$ it would be $[a_{11}^2 + a_{21}^2 + ... + a_{13}^2]$ -- shows the proportion of the variance of all the variables which is accounted for by that factor.

The Model for Individual Subjects

Equation (1) above, for variable 2, say, may be written explicitly for one subject $i$ as

$$X_{2i} = a_{21}F_{1i} + a_{22}F_{2i} + a_{23}F_{3i} + ... + a_{2m}F_{mi} + a_2U_{2i} \quad (2).$$

This form of the equation makes it clear that there is a value of each factor for each of the subjects in the sample; for example, $F_{2i}$ represents subject $i$'s score on Factor 2. Factor scores are often used in analyses in order to reduce the number of variables which must be dealt with. However, the coefficients $a_{11}$, $a_{21}$, ..., $a_{nm}$ are the same for all subjects, and it is these coefficients which are estimated in the factor analysis.

Extracting Factors and the Rotation of Factors

The mathematical process used to obtain a factor solution from a correlation matrix is such that each successive factor, each of which is uncorrelated with the other factors, accounts for as much of the variance of the observed variables as possible. (The amount of variance accounted for by each factor is shown by a quantity called the Eigen value, which is equal to the sum of the squared loadings for a given factor, as will be discussed below). This often means that all the variables have substantial loadings on the first factor; i.e., that coefficients $a_{11}$, $a_{21}$, ..., $a_{nm}$ are all greater than some arbitrary value such as .3 or .4. While this initial solution is consistent with the aim of accounting for as much as possible of the total variance of the observed
variables with as few factors as possible, the initial pattern is often adjusted so that each individual variable has substantial loadings on as few factors as possible (preferably only one). This adjustment is called rotation to simple structure, and seeks to provide a more interpretable outcome. As will be seen in the example which we'll work through later, rotation to simple structure can be seen graphically as the moving or rotation of the axes (using the term 'axis' in the same way as it is used in 'x-axis' and 'y-axis') which represent the factors.

**Estimating Factor Scores**

Given the equations (1) above, which show the variables $X_1 \ldots X_n$ in terms of the factors $F_1 \ldots F_m$, it should be possible to solve the equations for the factor scores, so as to obtain a score on each factor for each subject. In other words, equations of the form

\[
F_1 = b_{11}X_1 + b_{12}X_2 \ldots b_{1n}X_n
\]

\[
F_2 = b_{21}X_1 + b_{22}X_2 \ldots b_{2n}X_n
\]

\[
\ldots
\]

\[
F_m = b_{m1}X_1 + b_{m2}X_2 \ldots b_{mn}X_n
\]

(3)

should be available; however, problems are caused by the unique factors, because when they are included with the common factors, there are more factors than variables, and no exact solution for the factors is available. [An aside: The indeterminacy of the factor scores is one of the reasons why some researchers prefer a variety of factor analysis called principal component analysis (PCA). The PCA model doesn't contain unique factors: all the variance of the observed variables is assumed to be attributable to the common factors, so that the communality for each variable is one. As a consequence, an exact solution for the factors is available in PCA, as in equations (3) above.

**Calculating Correlations from Factors**

It was mentioned above that an aim of factor analysis is to 'explain' correlations among observed variables in terms of a relatively small number of factors. One way of gauging the success of a factor solution is to attempt to reproduce the original correlation matrix by using the loadings on the common factors and seeing how large a discrepancy there is between the original and reproduced correlations -- the greater the discrepancy, the less successful the factor solution has been in preserving the information in the original correlation matrix. How are the correlations derived from
the factor solution? When the factors are uncorrelated, the process is simple. The correlation between variables \( X_1 \) and \( X_2 \) is obtained by summing the products of the coefficients for the two variables across all common factors; for a three-factor solution, the quantity would be \((a_{11}x \ a_{21}) + (a_{12}x \ a_{22}) + (a_{13}x \ a_{23})\). This process will become clearer in the description of the hypothetical two-factor solution based on five observed variables in the next section.

**A Hypothetical Solution**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Loadings/Correlations</th>
<th>Communality</th>
<th>Reproduced correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
<td>Factor 2</td>
<td>( X_1 )</td>
</tr>
<tr>
<td>( X_1 )</td>
<td>.7</td>
<td>.2</td>
<td>.53</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>.8</td>
<td>.3</td>
<td>.73</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>.9</td>
<td>.4</td>
<td>.97</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>.2</td>
<td>.6</td>
<td>.40</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>.3</td>
<td>.7</td>
<td>.58</td>
</tr>
</tbody>
</table>

\( \sum x^2 = 2.07 \)  
\( \sum x^2 = 1.14 \)

**The Coefficients**

According to the above solution,

\[ X_1 = .7F_1 + .2F_2 \]
\[ X_2 = .8F_1 + .3F_2 \]

…..

\[ X_5 = .3F_1 + .7F_2 \]

As well as being weights, the coefficients above show that the correlation between \( X_1 \) and \( F_1 \) is .70 that between \( X_1 \) and \( F_2 \) is .20, and so on.
**Variance Accounted For**

The quantities at the bottom of each factor column are the sums of the squared loadings for that factor, and show how much of the total variance of the observed variables is accounted for by that factor. For Factor 1, the quantity is \( .7^2 + .8^2 + .9^2 + .2^2 + .3^2 = 2.07 \). Because in the factor analyses discussed here the total amount of variance is equal to the number of observed variables (the variables are standardized, so each has a variance of one), the total variation here is five, so that Factor 1 accounts for \( (2.07/5) \times 100 = 41.4\% \) of the variance.

The quantities in the *communality* column show the proportion of the variance of each variable accounted for by the common factors. For \( X_1 \) this quantity is \( .7^2 + .2^2 = .53 \), for \( X_2 \) it is \( .8^2 + .3^2 = .73 \), and so on.

**Reproducing the Correlations**

The correlation between variables \( X_1 \) and \( X_2 \) as derived from the factor solution is equal to \((.7 \times .8) + (.2 \times .3) = .62\), while the correlation between variables \( X_3 \) and \( X_5 \) is equal to \((.9 \times .3) + (.4 \times .7) = .55\). These values are shown in the right-hand side of the above table.

**How Many Factors?**

A factor solution with as many factors as variables would score highly on the amount of variance accounted for and the accurate reproduction of correlations, but would fail on economy of description, parsimony and explanatory power. The decision about the number of common factors to retain, or to use in rotation to simple structure, must steer between the extremes of losing too much information about the original variables on one hand, and being left with too many factors on the other. Various criteria have been suggested. The standard one (but not necessarily the best) is to keep all the factors which have Eigen-values greater than one in the original solution.
Need for the Study

In general, service quality promotes customer satisfaction, stimulates intention to return, and encourages recommendations. Customer satisfaction increases profitability, market share, and return on investment (Fornell, 1992). In a highly competitive hotel industry, individual hoteliers must find ways to make their products and services stand out among the others. To achieve this, hoteliers must understand their customers' needs – and then set out to meet (or exceed) these needs. Jiju, Freenie & Sid (2004) made a research identifying the dimensions of service quality in the UK hospitality industry based on the SERVQUAL instrument. Their study had its focus on a hotel group however it doesn't mention the type of hotels or the rating of stars given to the establishment. This makes it interesting whether the factor structured proposed in their study is valid in other type of hospitality establishment and also to look whether the perceived service quality dimensions differs by countries.

Hence, the study sets out to analyze the service quality dimensions of hotel guest using the SERVQUAL instrument as base in the emerging Indian market. The research questions were what are the effectiveness of service quality dimensions that are perceived and which dimensions are the best predictors of overall service quality provided by Indian hotels. Thus, there is need now to evaluate the customer satisfaction in relation to service quality offered by three types of hotels in India. we tried to measure service quality and customer satisfaction using the SERVQUAL model from the consumer's perspective in order to know their perceptions.

Limitations of the Study

We have to define our scope of research in order to make things clear. We are focusing our study in evaluating how consumers perceive service quality in these three types of Indian hotels in general. The study under consideration is evaluated the customer satisfaction level of 4 star, 5 star and 5 star deluxe hotels. There are some limitations associated with this study that need to be discussed. Firstly, the results obtained from this study cannot be generalized to a wide range of similar situations concerning three types of hotels because of the non-probability sampling technique used even though the methodology used in this study could be applied to these similar situations. The study used convenient sampling while survey of customers about service quality provided by Indian hotels. Also, the issue of consumers' perceptions could be questioned because the sample size consisted of
respondents that come from both developing and developed economies that may differ. The said survey was limited to perceive services of the hotels whereas the research could be extended to expectation survey. But, the present study that identifying the perceptions of customers, the dimensions of service quality and their relative importance to customers for each specific segment of the hotel industry would definitely help managers in the challenge of increasing customer satisfaction.