Chapter – III

RESEARCH METHODOLOGY

The present chapter discusses in detail the methodology adopted for the present research work. It describes the design of the study, sample size, sampling design and procedure of data collection. It also discusses the various statistical techniques used for analyzing the data collected for achieving the end objectives of the present research work.

The present study is four dimensional in nature. Firstly, it attempts to study the policy framework of Indian tourism industry with special reference to Punjab. Secondly, it examines the management of tourism in Punjab and the role played by various government agencies in promotion of tourism in Punjab. Thirdly, it attempts to identify the untapped tourism destinations in the state. Lastly, it makes an effort to study the profile and perceptions of domestic and international tourists visiting Punjab.

3.1 DESIGN OF THE STUDY

In order to study the policy framework of Indian and Punjab tourism industry, mainly secondary data has been used which is gathered from various published and unpublished sources of Ministry of Tourism, Government of India and Department of Tourism, Government of Punjab. Various websites were also visited to meet the requirements of data. To examine the outcome of efforts made by central and state governments for tourism development, the data regarding tourist arrivals have been taken and analyzed to draw the results. To analyze the data various statistical
techniques have been used like mean, standard deviation, co-efficient of variation, compound growth rate and trend analysis. To check the significance of compound growth rate, the t-test has been applied.

To examine the tourism management and role played by the government agencies in promotion of tourism in Punjab, the data has been collected from various annual reports of Ministry of Tourism, Government of India and Department of Tourism, Government of Punjab. Various projects sanctioned by the central government and initiatives taken by the state government have also been taken into consideration. Information on different projects has been collected from Department of Tourism, Punjab and from e-sources of Planning Commission of the state. The role played by prominent government agencies in promotion of tourism in the state has been discussed and duly supported by the data.

To identify untapped tourism destinations in Punjab, the data has been collected from both primary and secondary sources. To collect the primary data, the domestic and foreign tourists were requested to mention in the questionnaire any unexplored area, which they have come across during their visit to Punjab. The destination, which was described as untapped tourist destination by more than three tourists has been considered in the study along with its literature and picture. To get the information about the literature of untapped destination, published and unpublished material on historical, geographical, cultural and ecological personality of Punjab, travelogues, travel brochures, tourist guide maps and industry reports etc.
have been studied. Wherever possible, physical surveys have been carried out to authenticate the secondary data.

3.2 SAMPLE SIZE

To study the profile and perceptions of domestic and international tourists’, mainly primary data has been used. To collect the data, it was required to select a sample of tourists. There has been substantial debate over the sample size needed to conduct the study. Therefore in order to gain more insight over the sample size the data regarding tourist arrival was taken and examined.

Table 3.1

Domestic and International Tourist Arrival in Punjab during 2002-2011

<table>
<thead>
<tr>
<th>Year</th>
<th>Domestic Tourists</th>
<th>Foreign Tourists</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>3,17,904</td>
<td>7,558</td>
</tr>
<tr>
<td>2003</td>
<td>11,50,015</td>
<td>4,589</td>
</tr>
<tr>
<td>2004</td>
<td>5,36,120</td>
<td>5,224</td>
</tr>
<tr>
<td>2005</td>
<td>4,31,036</td>
<td>4,353</td>
</tr>
<tr>
<td>2006</td>
<td>3,53,907</td>
<td>1,93,933</td>
</tr>
<tr>
<td>2007</td>
<td>3,68,593</td>
<td>5,470</td>
</tr>
<tr>
<td>2008</td>
<td>5,09,428</td>
<td>6,869</td>
</tr>
<tr>
<td>2009</td>
<td>53,69,995</td>
<td>1,10,404</td>
</tr>
<tr>
<td>2010</td>
<td>105,83,509</td>
<td>1,37,122</td>
</tr>
<tr>
<td>2011</td>
<td>164,16,638</td>
<td>1,50,958</td>
</tr>
</tbody>
</table>

Source: Compiled from India Tourism Statistics 2002-2011

Table 3.1 depicts the data regarding the domestic and international tourist arrivals in Punjab from 2002 to 2011.
It is evident from the data that there is a glaring gap between the number of domestic and international tourists visiting Punjab during the previous decade. This shows that the proportional representation was not possible, therefore it was decided to take a sample of 300 tourists comprising of 200 domestic tourists and 100 international tourists.

3.3 SAMPLING DESIGN

The universe of the study was the whole state of Punjab, but due to non-feasibility and time constraint, the scope of the study was restricted to five districts of the state, Amritsar, Ludhiana, Roop Nagar, Jalandhar and Patiala.

Table 3.2

District wise Tourist Arrival in Punjab, June 2009–July 2010

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>District</th>
<th>Tourist Arrival</th>
<th>Sr. No.</th>
<th>District</th>
<th>Tourist Arrival</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Amritsar</td>
<td>8759533</td>
<td>11</td>
<td>Fatehgarh Sahib</td>
<td>355316</td>
</tr>
<tr>
<td>2</td>
<td>Ludhiana</td>
<td>1939214</td>
<td>12</td>
<td>Muktsar</td>
<td>344891</td>
</tr>
<tr>
<td>3</td>
<td>Rupanagar (Ropar)</td>
<td>1467569</td>
<td>13</td>
<td>Moga</td>
<td>270299</td>
</tr>
<tr>
<td>4</td>
<td>Jalandhar</td>
<td>1155568</td>
<td>14</td>
<td>Mansa</td>
<td>168494</td>
</tr>
<tr>
<td>5</td>
<td>Patiala</td>
<td>889494</td>
<td>15</td>
<td>Barnala</td>
<td>142280</td>
</tr>
<tr>
<td>6</td>
<td>Tarn Taran</td>
<td>765312</td>
<td>16</td>
<td>Hoshiarpur</td>
<td>141624</td>
</tr>
<tr>
<td>7</td>
<td>Bhatinda</td>
<td>623259</td>
<td>17</td>
<td>Kapurthala</td>
<td>120066</td>
</tr>
<tr>
<td>8</td>
<td>S.A.S Nagar</td>
<td>548946</td>
<td>18</td>
<td>Gurdaspur</td>
<td>111600</td>
</tr>
<tr>
<td>9</td>
<td>Ferozpur</td>
<td>531855</td>
<td>19</td>
<td>Nawanshahar</td>
<td>100871</td>
</tr>
<tr>
<td>10</td>
<td>Sangrur</td>
<td>370543</td>
<td>20</td>
<td>Faridkot</td>
<td>41799</td>
</tr>
</tbody>
</table>

Source: Compiled from Survey Report of Ministry of Tourism. 2010
Table 3.2 explains the criterion adopted for selecting the districts. It was the number of tourist arrivals, district wise. There are twenty districts in the state, which have been arranged in an descending order according to the total number of tourist arrivals during June 2009 to July 2010. The top five districts were selected having maximum number of tourist arrivals.

3.4 QUESTIONNAIRE DESIGN AND DEVELOPMENT

To collect the primary data a structured questionnaire was used. The questionnaire was divided into two parts. The first part focused on the demographic information and travel characteristics of tourist respondents. The second part focused on the measurement of expectations and perceptions of tourists regarding the level of service quality in Punjab Tourism Industry.

Questionnaire design involves the decision about, which items will best reflect the underlying construct that the researcher intends to measure. The first stage of scale development is the creation of items to assess the construct under examination (Hinkin, 1998). Existing review of literature regarding various aspects of tourism services was considered and accordingly statements were formed related to the expectations and perceptions of tourists. The modified SERVQUAL model consisting of 30 items has been used for assessing tourist’s expectations and perceptions regarding the overall quality of tourism related services. ‘Although SERVQUAL was designed to measure service quality, it provides only a framework or skeleton and thus has had to be adapted and modify to evaluate specific services’ (Parasuraman et al, 1988).
In order to validate and find out the reliability of questionnaire, a pilot study was conducted before processing with actual survey. The preliminary draft of the questionnaire was pre-tested on 25 persons associated to tourism industry, including tourism experts and tourists. Appropriate modifications in contents and format of the questionnaire were then incorporated in the light of experience gained during the pilot study. It helped in improving and finalising the questionnaire. Respondents were asked to rate their level of agreement or disagreement with given statements on a five point Likert scale ranging from ‘strongly disagree’ with scale point 1 to ‘strongly agree’ with scale point 5. ‘Tourists’ perceptions are based on the actual service they receive, while their expectations are based on past experiences and information received’ (Renganathan, 2011).

3.5 DATA COLLECTION

This study is based on primary as well as on secondary data. The primary data was collected by the researcher personally. Convenient sampling technique was used during the course of present research. The respondents were approached at different tourist destinations, accommodation units and tourist information offices in the selected districts. The purpose of the questionnaire was explained to the tourists and they were assured of total confidentiality of the information provided by them. As many as 450 tourists were approached to collect the required data for the study. However, only 300 questionnaires filled by the respondents, 200 domestic tourists and 100 foreign tourists with a response percentage of 66.7 percent were found complete in all respects for the analysis. The secondary data was
collected from various national and international research papers, research articles, Ph.D theses, books, journals and e-sources, published and unpublished reports of Ministry of Tourism, Government of India and Department of Tourism, Government of Punjab.

3.6 STATISTICAL TECHNIQUES USED FOR DATA ANALYSIS

A variety of traditional and advanced statistical techniques have been used to analyze the data in order to make the study more meaningful. A brief description of tools and the formulas used for data analysis is given as under:

1) Average (Mean)

Average is calculated to get a single value that depicts the characteristics of the entire group and to facilitate the comparison between the data. The formula is as under.

\[
\text{Mean } \bar{x} = \frac{\Sigma x}{N}
\]

\(\Sigma = \) Symbol of summation

Where, \(\Sigma x = \) Sum of series of observations

\(N = \) Number of items

2) Standard Deviation (S.D)

To measure the absolute dispersion between the variables, SD is calculated with the help of following formula.

\[
(\sigma) = \sqrt{\frac{\Sigma x^2}{N}}
\]
Where, \( x = (X - \overline{X}) \) \( \bar{X} \) is the mean of the series and \( (X - \overline{X}) \) is the deviation from the mean

\( N = \) Number of items

3) **Coefficient of Variation (C.V)**

To analyze the variations in the various parameters of tourism industry, coefficient of variation has been used which is calculated with the following method.

\[
C.V. = \frac{S.D.}{Mean} \times 100
\]

4) **Compound Growth Rates (CGR)**

In order to assess the existing trend in tourist arrivals and tourists' receipts at international, national and state level, the compound growth rates of different parameters were worked out by applying the regression model in its exponential form as under:

\[
Y = ab^t u
\]

Where,

- \( Y = \) a parameter
- \( a = \) a constant term
- \( t = \) time variable (No. of years under study)
- \( b = \) regression coefficient of time
- \( u = \) a random error term

Log transformation of the above function is:

\[
\ln Y_t = \ln a + t (\ln b) + e
\]

Where,

- \( \ln b = \ln (1 + r) \) and

\[
r = \text{antilog (In b) - 1} \times 100
\]

\[
CGR (\%) = \text{antilog (In b) - 1} \times 100
\]
The compound growth rates were tested to determine their statistical significance with the help of t-test.

5) Trend Equation

To make future projections of tourist arrival on national and state level, trend equations were developed in the form of linear regression as under:

\[ Y = a + bt + u \]

Where, \( Y \) = a performance indicator

\( a = \) a constant term
\( t = \) time variable (No. of years under study)
\( b = \) regression coefficient of time
\( u = \) a random error term

Then the coefficient of \( a \) and \( b \) were used to make future projections.

6) Chi-Square test

In order to see the association between the two way distributions of demographic and travel characteristics of the tourist respondents, Chi-square test was applied by using the following formula:

\[ \chi^2 = \sum \frac{(O - E)^2}{O} \]

Where \( \chi^2 \) = Chi-Square Value
\( O = \) Observed Frequency
\( E = \) Expected Frequency
\( \Sigma = \) Summation
7) **Unpaired t-test**

To compare the two mean values of independent groups, to compare the gaps of domestic and foreign tourist sectors regarding the five dimensions of SERVQUAL scale and to test the significance of difference between various parameters, unpaired t-test was applied as under:

\[
t = \frac{\bar{X}_1 - \bar{X}_2}{SE \left( \bar{X}_1 - \bar{X}_2 \right)}
\]

\[
SE(\bar{X}_1 - \bar{X}_2) = s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}
\]

\[
S = \sqrt{\frac{SD_1^2 (n_1 - 1) + SD_2^2 (n_2 - 1)}{n_1 + n_2 - 2}}
\]

Where

**SE** = standard error of mean difference

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

\( \bar{X}_2 \) = mean value of second group

SD\(_1\) = Standard deviation in first group

SD\(_2\) = Standard deviation in second group

S = Common Standard Deviation

\( n_1 \) = Number of respondents in first group

\( n_2 \) = Number of respondents in second group

8) **Z-test**

In order to compare two proportions of tourist respondents, Z-test i.e. test of proportions was applied as under:

\[
Z = \frac{|P_1 - P_2|}{SE \ of \ (P_1 - P_2)}
\]
\[ SE \ of \ (P_1 - P_2) = \sqrt{pq\left(\frac{1}{N_1} + \frac{1}{N_2}\right)} \]

\[ p = \frac{P_1 n_1 + P_2 n_2}{n_1 + n_2} \]

\[ q = 1 - p \]

Where \( P_1 \) = Proportion of respondents in one group

\( P_2 \) = Proportion of respondents in second group

\( n_1 \) = Total number of respondents in one group

\( n_2 \) = Total number of respondents in second group

9) **Paired t-test**

To order to assess the significance of mean gap (difference between expectations and perceptions of tourist respondents) within the same group, Paired t-test was applied. The formula used is as under

\[ t = \frac{\overline{d} \sqrt{n}}{S} \]

Where \( \overline{d} \) = the mean of the differences

\( S \) = Standard deviation of the difference

Factor analysis is used to identify a smaller number of factors underlying larger number of observed variables. The factor analysis of all independent variables pertaining to the perceptions of tourists on all the five dimensions is conducted in order to reduce the number of items into a manageable number of factors. A varimax rotated principal component analysis is used on 30 items for a sample of 200 domestic tourists and 100
international tourists. To check the reliability of data, Kaiser-Meyer-Olkin (KMO) and Bartlett's Test has been applied. The KMO ranges from 0-1, with higher values indicating greater suitability.

10) **Factor Analysis**

Factor analysis has been employed in the study to assess to respondents' attitude towards quality of various services availed by them during their visit. In order to serve the objective the information collected through several statements has been condensed into important dimensions with the help of factor analysis.

Factor analysis is an interdependence technique in which all variables are simultaneously considered related to each other. It studies the structure of interrelationships among a large number of variables by defining a set of common underlying latent dimensions known as factors. As a result variables within each factor are more highly correlated with variables in that factor than with variables in other factors. This makes it possible to interpret the data from a much smaller number of factors than the original individual variables.

Choice of methods used to extract factors is an important step in factor analysis. In order to summarize most of the original information in a minimum number of factors, the Principal Component Analysis was used. This is a statistical technique that linearly transforms an original set of variables into a substantially smaller set of uncorrelated variables that represents most of the information in the original set of variables. Here linear combinations of variables are used to account for variation of each dimension in a multivariate space. The
variance of the factors is called Eigen values, characteristic root or latent root. Communality is the amount of variance an original variable shares with others. Factor loadings are the correlations between the original variable and the factor. Guidelines exist (Hair et al., 1995) for identifying significant factor loading based on the sample size. Squared factor loadings indicate what percentage of the variance in an original variable is explained by a factor. When the set of variable is large as in this study; the analyst first extracts the largest and best combinations of variables and then proceeds to smaller and less understandable combinations. Hence the number of factors to be extracted becomes an important issue in the absence of any set criterion.

An important step in factor analysis is rotation of factors. Loadings are rotated to make them more interpretable by making the loadings for each factor either large or small, not in between. For rotation, either orthogonal or oblique method can be employed. In this case orthogonal rotation method has been employed; the axes are maintained at 90 degrees so that the resulting factors are uncorrelated. Within orthogonal method either varimax or quartimax method can be employed. The varimax was retained because it is also more realistic because since the theoretically important underlying dimensions are not considered to be uncorrelated with each other. In the present study, orthogonal method along with the varimax method of rotation has been used. The varimax criterion maximizes the sum of the variances of the squared loadings within each column of the loading matrix. This tends to produce some high loadings and some near zero, which is one of the aspects of simple structure.
The final step in this is to name the factors. The naming of factors has been done intuitively; depending on its appropriateness for representing the underlying dimensions of a particular factor. The naming of factors is not typically a scientific process; still some guidelines are recommended (Hair et al, 1995). In view of the above discussion, factor analysis has been used to condense 30 statements into a few uncorrelated factors with the help of principal component analysis with orthogonal rotation along with the varimax option of rotation of factors.

**Mathematical Model for Factor Analysis**

for $i = 1, ..., 200$, the $i^{th}$ respondent's scores are

$$X_{1,i} = \mu_1 \cdot 1_{1 \times 200} + L_{1,1}F_1 + L_{1,2}F_2 + \varepsilon_{1,i}$$

$$X_{n,i} = \mu_n \cdot 1_{1 \times 200} + L_{n,1}F_1 + L_{n,2}F_2 + \varepsilon_{n,i}$$

Where $x_{k,i}$ is the $i^{th}$ respondent's score for the $k^{th}$ statement

$\mu_k$ is the mean of the respondents scores for the $k^{th}$ statement

$n$ is the number of statements (10, 10 and 15)

$F_1$ is the $i^{th}$ respondent's "first factor", $F_2$ is the $i^{th}$ respondent's "second factor", and so on

$\varepsilon_{k,i}$ is the difference between the $i^{th}$ respondent's score in the $k^{th}$ statement and the average score in the $k^{th}$ statement of all the respondents,

In matrix notation, we have

$$X = \mu \cdot 1_{1 \times N} + LF + \varepsilon$$

Where $N$ is No. of respondents

$X$ is a statements $\times$ No. of respondents matrix of observable random variables,
µ is a number of statements × 1 column vector of unobservable constants,
L is a statements × 2 matrix of factor loadings,
F is a 2 × consumer's matrix of unobservable random variables,
ε is a statements × respondents matrix of unobservable random variables.

3.7 LIMITATIONS OF THE STUDY

This study suffers from a number of limitations.

• Though, the geographical location of the study is the whole of Punjab, but for certain reasons the study has been confined to selected districts only, which may not necessarily portray the complete picture of tourism in Punjab.

• The study has focused on tourism services only, but there are many other components of tourism industry, which are imperative to tourist satisfaction. These components are accommodation, restaurant, food services, transportation, shopping, recreation, entertainment, destination competitiveness, local residents' attitude etc. Thus, the current study might be able to provide a comprehensive view of the tourism experience.

• The present study has examined tourists' satisfaction only through service quality. However, satisfaction of tourists also depends upon tourist motivation, travel behavior, trip characteristics, destination image and many other factors. Thus, evaluating tourists'
satisfaction simply based on service quality may be one-sided.

- The present study has employed the SERVQUAL instrument to measure the service quality gaps in tourism industry. However, a number of critics have pointed to its inadequacy. As long as this instrument suffers from such limitations, the present study also cannot claim to be foolproof.

- In the present study, both primary and secondary data have been used. As the primary data is collected through a questionnaire, it may suffer from the basic limitation of recording an unreliable opinion because generally the respondents are reluctant to provide correct information due to their own reasons.

- Similarly, the information collected for the secondary data based parts of the study may suffer from the inherent limitations associated with the secondary data.

- While analyzing the tourism management, only three parameters of management has been taken into consideration. Other areas of management like financial management, Product management have not been studied.