CHAPTER II

RESEARCH DESIGN
This chapter is an attempt to present the research design about the study population, sample design, sampling techniques, collection of data, tools and technique used in the study.

2.1 METHODOLOGY

2.1.1 Study Population

The population of this study consists of the total number of companies in the Bombay Stock Exchange Sensex (BSE 200). Out of which only 30 companies follow Human Resource Accounting. Hence these 30 companies, 16 Public Ltd. companies and 14 Private Ltd. companies have been selected as a sample unit for the study.

2.1.2 Sample Design and Sampling techniques

The present study used convenience sampling method and opinions have been collected from managers, accountants and employees in the respective companies. The companies which follow Human Resource Accounting have been selected as a sample unit for the study. Hence 30 companies have been selected as a sample unit for the study, 16 Public Ltd companies and 14 Private Ltd companies. A pre-test has been conducted through pilot study so as to ensure the relevance and consistency of various items and statements in the questionnaire, inconsistent statements have been altogether dropped and others found appropriate. Based on the East Africat Sample determination and allocation formula, 300 respondents have been selected as a sample size for the study. The formula is as follows:

\[ n = \left( \frac{ZS}{E} \right)^2 \]

Where,

\[ Z = \text{Standardized value corresponding to a confidential level of 95\% i.e. 1.96} \]
S= Sample standard deviation from pilot study is 0.442
E= Acceptable error=5%
\[ n = (1.96*0.442/0.05)^2 = 300.204 \]

Table 2.1 gives clear picture about the sample structure of the study i.e., total number of Public and Private Ltd., companies taken for the study and number of respondents in each Public and Private Ltd., companies in India.

Table 2.1
Sample Structure

<table>
<thead>
<tr>
<th>Public Ltd Companies</th>
<th>Private Ltd Companies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S.No</strong></td>
<td><strong>Company</strong></td>
</tr>
<tr>
<td>1</td>
<td>BHEL</td>
</tr>
<tr>
<td>2</td>
<td>SAIL</td>
</tr>
<tr>
<td>3</td>
<td>NTPC</td>
</tr>
<tr>
<td>4</td>
<td>HUL</td>
</tr>
<tr>
<td>5</td>
<td>OIL</td>
</tr>
<tr>
<td>6</td>
<td>HCL</td>
</tr>
<tr>
<td>7</td>
<td>HPCL</td>
</tr>
<tr>
<td>8</td>
<td>IOC</td>
</tr>
<tr>
<td>9</td>
<td>ONGC</td>
</tr>
<tr>
<td>10</td>
<td>EIL</td>
</tr>
<tr>
<td>11</td>
<td>HZL</td>
</tr>
<tr>
<td>12</td>
<td>MMTC</td>
</tr>
<tr>
<td>13</td>
<td>SBI</td>
</tr>
<tr>
<td>14</td>
<td>NLC</td>
</tr>
<tr>
<td>15</td>
<td>CCI</td>
</tr>
<tr>
<td>16</td>
<td>SPIC</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
</tr>
</tbody>
</table>

2.1.3  **Collection of Data**

The present study has collected both primary and secondary data.

2.1.3.1  **Primary Data**

The primary data are collected from managers, accountants and employees of both Private and Public Ltd companies from the administration of questionnaire. The questionnaire is designed to include set of statements with five point Likert scale, values ranging from 1 to 5 for ‘strongly disagree’ to ‘strongly agree’ to gather opinion from sample respondents about nature and characteristics of HRA, Ability of users of HRA to make wise decisions, perception of management and employees on HRA and contribution of HRA in Public and Private companies in India.

A questionnaire comprising of two parts viz, Part I, Part II. Under Part I personal information data covering 8 questions. It is shown below:

**Part I - Personal Related Variables:** (Name of the concern, Category of respondent, Age, Educational Qualification, Designation, Years of Experience, Years of service in this company)

**Part II of the questionnaire consisting of 78 statements related to Human Resource Accounting.** These 78 statements are classified into four major factors which are presented in the table.
Table 2.2

Details of number of statements for each factor

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Human Resource Accounting Factors</th>
<th>Number of Statements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Factors related to Nature and Characteristics of Human Resource Accounting (HRA)</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>Factors related to the ability of users of Human Resource Accounting (HRA) to make wise decisions</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>Factors related to Human Resource Accounting (HRA) to know about the Perception of Management and Employees.</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>Factors related to Contribution of Human Resource Accounting on Financial statements of Indian companies</td>
<td>8</td>
</tr>
</tbody>
</table>

2.1.3.2 Secondary Data

Secondary data has been collected from the Annual reports of the selected 30 companies from BSE and other necessary information, collected from the Centre for Monitoring Indian Economy (CMIE) reports. The relevant sources have been gleaned from books, journals and magazines.

2.2 Tools and Techniques

The following statistical techniques are used to analyse human resource accounting in Indian companies:

- Descriptive statistics like mean, standard deviation
- Reliability test
- Factor Analysis
- Independent sample t-test
- One way ANOVA
- Correlation Analysis
- One sample Kolmogorov smirnov test
- Regression analysis
- Mantel test

These analyses are calculated using SPSS and R Software. The details of the statistical technique adopted in the present research work are given hereunder:

### 2.2.1 Descriptive statistics

It is the discipline of quantitatively describing the patterns and general trends of a data and summarizes it in a single value. It enables to understand quickly and interpret the set of data that has been collected.

### 2.2.2 Mean

Mean represents the value that the researcher gets after dividing the sum of the observations by the total number of observation taken. This is known as arithmetic mean that can be calculated by using the following formula:

\[
\text{Mean (X)} = \frac{\sum X_i}{n}
\]

Where,

\( (X) = \) Symbol of mean

\( \Sigma X_i = \) Sum of all the observations

\( X_i = X_1+X_2+\ldots+X_n \)

\( n = \) Number of Observations.

### 2.2.3 Standard deviation

Standard deviation represents the measure used to calculate the scattering of values in a given data. The symbol used to represent standard deviation is sigma (\( \sigma \)).
Standard deviation is the square root of variance of the data series. The formula used to calculate standard deviation is as follows:

\[ (\sigma) = \sqrt{\frac{\Sigma (X_i - X)^2}{n}} \]

### 2.2.4 Skewness and Kurtosis

The skewness coefficient, in excess of unity is taken to be fairly extreme. High or low kurtosis value indicates extreme leptokurtic or extreme platykurtic. Generally values for zero skewness and kurtosis at three represent that the observed distribution is normally distributed.

### 2.2.5 Reliability test

Reliability test refers to the degree to which a test is consistent and stable in measuring what it is intended to measure. Most simply put, a test is reliable if it is consistent within itself and across time. Test validity refers to the degree to which the test actually measures what it claims to measure. Reliability is used to describe the overall consistency of a measure. A measure is said to have a high Reliability if it produces similar results under consistent conditions.

Before applying the hypothetical statistical model, it was decided to test the reliability of scale items to identify the nature and characteristics of Human Resource Accounting in Indian Companies. The reliability test is used to choose items of perfect, that is, one should look at each item and eliminate those items that have very low or nearly zero correlation to sum score of all items and removal of which may further increase the value of Cronbach coefficient. Generally, as a rule of thumb any item having correlation to total sum scale value below 0.20 can be removed and remaining items could be used for further analysis. If Cronbach alpha is zero, the scale items could be considered as unreliable measures. Otherwise, if all
items involved in the reliability test are perfectly reliable and measure the same thing (true score) then the coefficient alpha will be equal to 1. As a rule of thumb, the calculated alpha value for a given set of items should be above 0.70 to accept those items as a reliable measurement scale.

2.2.6 Factor Analysis

Factor analysis is a statistical method used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved variables called factors. For example, it is possible that variations in four observed variables mainly reflect the variations in two unobserved variables. Factor analysis searches for such joint variations in response to unobserved latent variables. The observed variables are modeled 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. Computationally this technique is equivalent to low rank approximation of the matrix of observed variables. Factor analysis originated in psychometrics, and is used in behavioural sciences, social sciences, marketing, product management, operations research, and other applied sciences that deal with large quantities of data.

2.2.7 Independent 't' Test

The independent t-test, also called the two sample t-test or student's t-test, is an inferential statistical test that determines whether there is a statistically significant difference between the means in two unrelated groups.

Hypothesis for the independent t-test
The null hypothesis for the independent t-test is that the population means from the two unrelated groups are equal:

\[ H_0: \mu_1 = \mu_2 \]

In most cases, null hypothesis is rejected and accept the alternative hypothesis, which is that the population means are not equal:

\[ H_A: \mu_1 \neq \mu_2 \]

For this, a significance level (alpha) should be set either to reject or accept the alternative hypothesis. Most commonly, this value is set at 0.05.

In order to run an independent t-test, the following is required:

- One independent, categorical variable that has two levels.
- One dependent variable.

**Unrelated groups**

Unrelated groups, also called unpaired groups or independent groups, are groups in which the cases in each group are different. Often the investigation between differences in individuals are done, which means that when comparing two groups, an individual in one group cannot also be a member of the other group and vice versa. An example would be gender - an individual would have to be classified either as male or female - not both.

**2.2.8 Correlation Analysis**

It is often necessary to examine the relationship between two or more financial variables. There are many ways to examine how sets of data are related. Two of the most useful methods are scatter plots and correlation analysis. A scatter plot is a graph that graphically depicts the relationship between the observations for two data series in two dimensions. In contrast to a scatter plot, correlation analysis
expresses the same relationship using a single number. The correlation coefficient is a measure of how two data series are closely related. In particular, the correlation coefficient measures the direction and extent of linear association between two variables. A correlation coefficient can have a maximum value of 1 and a minimum value of -1. The correlation cannot exceed 1 in absolute value. A correlation coefficient greater than 0 indicates a positive linear association between the two variables: When one variable increases (decreases), the other also tends to increase (decrease). A correlation coefficient lesser than 0 indicate a negative linear association between the two variables: When one variable increases (decreases), the other also tends to decrease (increase). A correlation coefficient of 0 indicates no linear relation between the two variables. The correlation coefficient is the covariance of two variables (X and Y) divided by the product of their sample standard deviations (Sx and SY).

The formula for computing the sample correlation coefficient is

\[ r = \frac{\text{Cov}(X, Y)}{S_x \cdot S_Y} \]

Where,

- \( r \) = The sample correlation coefficient
- \( \text{Cov}(X, Y) \) = The sample of covariance X and Y
- \( S_x \) = The sample standard deviation of X
- \( S_Y \) = The sample standard deviation of Y

**Testing the significance of the correlation coefficient**

Significance tests allow analysts to assess whether apparent relationship between the random variables is real or due to chance. If the analyst decides that the relationship is real, they will be inclined to use this information in projections about
the future. To find out whether the relationship is real Test of significance of an observed correlation coefficient is to be applied. This test assumes that both the variables are normally distributed. The following hypotheses are structured:

\( H_0: r_p = 0 \)

\( H_a: r_p \neq 0 \) (two tailed test)

Whether the null hypothesis should be rejected using the ‘r’ can be tested.

A rejection of \( H_0 \) would indicate statistically significant relationship between the variables. Whereas a non rejection of \( H_0 \) would indicate that there is no significant relationship between the variables.

The formula for the t test is

\[ t = r \sqrt{n-2}/ \sqrt{1-r^2} \]

Where

- \( r \) = correlation coefficient
- \( n \) = size of sample
- \( df = n-2 \)

If the calculated value of \( t > t_{0.05} \) for (n-2), df, then the value of \( r \) is significant at 5 % level. If \( t < t_{0.05} \), the data are consistent with the hypothesis of an uncorrelated population.

### 2.2.9 ANOVA

The one-way analysis of variance (ANOVA) is used to determine whether there are any significant differences between the means of three or more independent (unrelated) groups. The ANOVA test is the initial step in identifying factors that are influencing a given data set. After the ANOVA test is performed, the analyst is able to perform further analysis on the systematic factors that are
statistically contributing to the data set's variability. Analysis of Variance (ANOVA) is a statistical method used to test differences between two or more means. The one-way ANOVA compares the means between the groups one is interested in and determines whether any of those means are significantly different from each other. Specifically, it tests the null hypothesis:

$$H_0: \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_k$$

where $\mu$ = group mean and $k$ = number of groups. If, however, the one-way ANOVA returns a significant result, one can accept the alternative hypothesis ($H_A$), which implies that there are at least 2 group means that are significantly different from each other.

### 2.2.10 Mantel Test

The Mantel test is a permutation technique that estimates the resemblance between two proximity matrices computed about the same object. Mantel’s test is an approach that overcomes some of the problems inherent in explaining species environment relationships. Mantel’s test is a regression in which the variables are their distance or dissimilarity matrices summarizing pair wise similarities among sample locations. In mantel test the matrices must be of the same rank, but not necessarily symmetric, though from practice this is often the case. The mantel technique was first introduced as a solution to the epidemiological question where interest is on whether case of diseases that occurred close in space also tends to be close on time, it is explained that multivariate tables of observations are usually condensed into resemblance matrices among any sampling unit of interest computed using proximity measure, in this research the canonical measure is used as a
measure as it is displayed by the DA (distance over objects of group A) and DB (distance over objects of group B). Hence the technique was used to compare matrix of special distances in a generalized regression approach. Since the mantel test has always included any conceivable proximity matrices. Letting $dA_{ij}$ and $dB_{ij}$ represent the distance observational units i and j, derived from the observations for variables A and B, where, $DA=(dA_{ij})$ and $DB=(dB_{ij})$ denote the corresponding $n \times n$ distance matrices. The normalized Mantel statistic, defined as the product – moment coefficient between distance matrices $DA$ and $DB$, is

$$r_M(AB) = \frac{\sum\sum (dA_{ij} - dA)(dB_{ij} - dB)}{\sqrt{\sum\sum (dA_{ij} - dA)^2 \sum\sum (dB_{ij} - dB)^2}}$$

Where $\sum\sum$ denotes the double summation over i and j of $DA$ and $DB$ and $dA$ and $dB$ are means of distances derived from the A and B raw data respectively. The testing procedure is as follows:

1. Considering two symmetric resemblance matrices (similarities) A and B, of size $(n \times n)$, whose rows and columns are corresponding to the same set of objects, Spearman correlation) between the corresponding objects of the upper triangular (or lower-triangular) portions of these matrices, obtaining the mantel correlation (often called the standardized Mantel statistic) $r_M(AB)$, which will be used as the reference value in test.

2. To permute at random the rows and corresponding columns of one of the matrices, say A, obtaining a permuted matrix $A^*$. This procedure is called ‘matrix permutation’.

3. To compute the standardized Mantel statistic $r_M(A^*B)$ between matrices $A^*$ and B which value $r^*_M$ of the test statistic under permutation.
4. Repeat steps 2 and 3, a large number of times to obtain the distribution of $r^*_M$ under permutation: then, add the reference value $r_M (AB)$ to the distribution.

5. For a one tailed test involving the upper tail (i.e. $H_1$ – distances in matrices A and B are positively correlated), calculate the profitability ($p$ – value) as the proportion of values $r^*_M$ greater than or equal to $r_M (AB)$. For a test in the lower tail, the probability is the proportion of values $r^*_M$ smaller than or equal to $r_M (AB)$.

Note that for symmetric distance matrices, only the upper (or lower) triangular portions are used in the calculations while for non symmetric matrices, the upper and lower triangular portions are included. The main diagonal elements need not be included in the calculation, but their inclusion does not change the $p$ value of the test statistic.\textsuperscript{21}

### 2.2.11 One sample Kolmogorov Smirnov Test

The One-Sample Kolmogorov-Smirnov Test procedure compares the observed cumulative distribution function for a variable with a specified theoretical distribution, which may be normal, uniform, Poisson, or exponential. The Kolmogorov-Smirnov $Z$ is computed from the largest difference (in absolute value) between the observed and theoretical cumulative distribution functions. This goodness-of-fit test tests whether the observations could reasonably have come from the specified distribution. Many parametric tests require normally distributed

variables. The one-sample Kolmogorov-Smirnov test can be used to test that a variable (for example, income) is normally distributed.

2.2.12 Regression Analysis

It is a statistical approach to forecast the change in a dependent variable (sales revenue, for example) on the basis of change in one or more independent variables (population and income, for example). It is also known as curve fitting or line fitting because a regression analysis equation can be used in fitting a curve or line to data points, in a manner such that the differences in the distances of data points from the curve or line are minimized. Relationships depicted in a regression analysis are, however, associative only, and any cause-effect (causal) inference is purely subjective. It is a statistical technique used to find relationships between variables for the purpose of predicting future values.

2.3 SUMMARY

The credibility of any research undertaking stands on how scientifically and logically it is conducted. The more scientific a research is, more it is accepted. This chapter reveals the details of all the procedures and processes, followed in arriving at the solution to the stated research problem. The study has availed both primary and secondary data and the collection of each data has been clearly explained in this chapter. The Population, upon which this study is based, is elaborately explained and reported. Sampling design and the formula used to calculate the sample is also adequately covered in this chapter. This chapter also clearly explains the research methodology and the technique used in the data analyses of this study.