Chapter-3

Research Methodology

3.1 INTRODUCTION TO RESEARCH METHODOLOGY

This chapter details out the research methodology for the present study. It explains the research objectives and a suitable methodology to achieve those objectives. The objectives of this study were to analysis the Web Data Design Pattern in data warehousing and business intelligence. It’s involved an exhaustive study of the metadata characteristics and details of the Web Data Design Pattern Illustration in data warehousing projects handled by the respondents. It includes:

(1) Exploration of the risk dimensions, and then comparison of these dimensions across the various characteristics.

(2) Identification and exploration of the organizational climate dimensions those are present in the Indian software companies. This was done by detecting the factors through field survey and comparing them across the demographics and project characteristics.

(3) The moderating effect of demographic characteristics and organizational climate factors on the project specific risk factors was studied through regression analysis.
(4) Assessed the impact of risk factors and organizational climate factors on the success and the three performance constructs of success namely budget, schedule and quality separately.

This was followed by model validation through four case studies involving post-mortem analysis of the projects undertaken by the Indian software companies. The research methodology has to be robust in order to minimize errors in data collection and analysis. Owing to this, various methodologies namely survey, interview (telephonic, structure) has been used.

3.1.1 RESEARCH MOTIVATION

In modern age all reputed and well-suited organizations are under enormous pressure to respond quickly for continuous changing trends in the competitive market. So it is obvious that in order to achieve this challenge, they need rapid access to different varieties of information before a decision can be framed. For making the right choices for organization, it is very much essential to study and analyze the past data and identify most relevant trends. Before perform any trend analysis it is very much necessary to access to entire relevant information, and naturally this information is mainly stored in very large/huge databases. To build a data warehouse is an easiest approach to gain access large volume of data for extract patterns and effective decision making for business environment.

Web Data Design Pattern Illustration in Data warehouse stores historical data and operational data. Operational data is useful to access information and
generate some patterns using data mining techniques. These generated Web Data Design Pattern Illustration can be analyzed to organize and summarize as Web Data Design. Today the Web Data Design Pattern Illustration with quality information at backbone can give the power of withstanding and growth.

3.1.2 OBJECTIVE OF THE RESEARCH

The objective of this research is to do analytical study and computational modeling of statistical methods useful for Web Data Design in data warehouse. How various statistical methods are useful for data Web Data Design Pattern Illustration and how competitive but less expensive computational model can be built? Data warehouse tools available in market are very expensive and too complex. I want to suggest a model for data mining that built with appropriate statistical methods. This has created a special interested in making comparison of different algorithms of data mining with effort to develop experimental paradigms that allow testing the mining algorithms.

For the different problems and applications used Web Data Design pattern in data warehouse and data mining strategies do not follow the same track and gives different pictures in different situations. The variability may cause the implementation complex and success rate not to the anticipated level. This problem is the main reason for failure data warehouse applications in most of organizations. To smoothen the track, efforts are made to propose the
model that keeps the track of data mining and that is expected to cover many different situations.

Today the numbers of organizations are rapidly increasing which are using Web Data Design pattern, data mining applications for Web data design. The model is to be proposed under this research initiative will help the developers for data warehouse solutions and case studies were chosen for data collection. This chapter describes the conduct study, participants of the study, instrumentation done for the study, data collection, and data analysis procedures of the entire study.

3.2 CONDUCT OF STUDY

The pilot study formed the pedestal for the research. It was conducted on 20 software project managers working in various software companies in NCR region. They were asked to list down the various risks that they have faced during the Software Development Lifecycle while executing the project. Through, the pilot survey, a list of 23 project specific risk items were identified and used for the study. Furthermore, the project managers were also asked to identify and rate the organizational climate factors which they perceive were present during the execution of the software projects and based on their perception; a list of 17 items of organizational climate was obtained and used for the survey. Based on the findings of this pilot study, the survey instrument was designed. The detailed description of the instrument is discussed.
3.3 RESEARCH DESIGN

The exploratory and descriptive research design was adopted due to the nature of the study. Exploratory research provides insights into and comprehension of an issue or situation. Exploratory research is a type of research conducted because a Web Data Design problem has not been clearly defined. Exploratory research helps to determine the best research design, data collection method and selection of subjects. While descriptive research, also known as statistical research, describes data and characteristics about the population or phenomenon being studied. Descriptive research answers the questions who, what, where, when and how. Thus, on the basis of the above, the two research designs were appropriate for the present study as it was important to gauge the various project specific risks that impact the software projects and also understand the dynamics of organization’s climate on these software projects.

3.4 INSTRUMENT OF THE STUDY

For effective and flawless Web Data Design, data collection, survey, interview and case study methods were extensively used. Survey method is the most extensively used technique for data collection, especially in behavioral sciences, while interviews are an appropriate method to use when exploring practitioners perspectives due to the qualitative nature of the information. Case study methods are used for an in-depth investigation of a single individual, group, or an event. It provides a systematic way of looking at events,
collecting data, analyzing information, and reporting the results. Thus, these methods have been widely used to extract the most relevant information and help in better analysis of the data.

In order to efficiently use the survey method a questionnaire was developed. The questionnaire was intricately designed to gauge the success rate of the projects undertaken by the Indian software professionals, to tap the project specific risk affecting the software projects and to understand the organizational climate dimensions. The instrument was divided into four parts.

In **Part A**, the respondents were asked to rate the overall success of the project. They were also asked to rate the performance of their project on the basis of the three performances constructs namely budget, schedule and quality. Although there are a number of constructs that define the success of the software project, however in the present study traditional definition of the success has been adopted, i.e. a project is said be to successful when it is delivered within an estimated budget, in predefined schedule and meeting all the quality parameters.

In **Part B and C**, the respondents were asked to rate the impact of risks on the success of their last executed project. Web based Data Design Software risks have been defined in various ways, however for the present study the risk has been defined as the probability -weighted impact of an event on a project.

It must be noted here that only project specific Web Data Design software risks have been taken into consideration for the study. The project specific
risks are those risks that are present in the Software Development Life Cycle and affect the project delivery.

A list of twenty three items was provided to the respondents. All the items were put on a five-point scale ranging from far too much effect to no effect on the success. Respondents indicated no effect with the risk if that threat did not exist in their project or the impact was very trivial. Similarly, the threat having far too much effect indicated that the impact of this risk was very severe and right mitigation technique could not be used to reduce the impact.

In part D, respondents were asked to rate organizational climate factors (seventeen variables) present during the last executed project. The organization climate has been defined as ‘the perceived attributes of an organization and its sub-systems as reflected in the way an organization deals with its members, groups and issues’ for the present research. The primary purpose of asking this question was to identify whether the climate has any effect on the risk factors and the success of the project i.e. does the organization climate help in reducing the risk factors and their impact on the success of the project.

The participants were asked whether clear understanding of roles and responsibilities within the group, full utilization of skills and abilities in the project, on time completion of work tasks, appreciation of work by direct supervisor etc was present in the organization during the execution of the project. A 5 point Likert scale was designed to gauge the responses in part D.
The scale ranged from never to always present. The emphasis was laid on the last executed project rather than the current project keeping in view the fact that the post facto analysis must have been done on these projects and the respondents will be able to answer the questions more precisely and accurately. Moreover, it would be easier for respondents to give information on whether their projects were successfully delivered, keeping in view the risks that the project faced in a given organization climate.

To test the validity of the instrument, a study was done on 100 participants between the months of January 2014 to April 2014. Based on their responses, validity tests were done to check for the validity and usability of the instrument. Cronbach alpha, KMO measure of adequacy and Bartlett’s test of sphericity were conducted. Cronbach alpha was calculated to measure the internal consistency reliability of the instrument. If the value of Cronbach’s alpha is greater than 0.7 then the instrument is considered reliable.

The value of Cronbach alpha came as 0.956 for Part II and III and 0.903 for Part IV; thus, the instrument was considered reliable for the study. Kaiser-Meyer-Olkin test was done to measure the homogeneity of variables and Bartlett's test of sphericity was done to test for the correlation among the variables used. The KMO value for part II and III of the instrument was 0.916, and for part IV was 0.817, both of which are acceptable as a good value.

The Bartlett’s test showed significant results for both the parts and hence the instrument was accepted for further study. Table 3.1
summarizes the entire result viz. Cronbach alpha, KMO test values, and Bartlett’s significance of the instrument. On getting quite meritorious results of the validity, the instrument was floated for data collection.

<table>
<thead>
<tr>
<th></th>
<th>No of items</th>
<th>Cronbach’s Alpha</th>
<th>Kaiser-Meyer-Olkin Measure of Sampling Adequacy</th>
<th>Bartlett's Test of Sphericity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Part II &amp; III</td>
<td>23</td>
<td>0.956</td>
<td>0.916</td>
<td>Approx Chi-Square 5309.252</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DF 253.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sig. .000</td>
</tr>
<tr>
<td>Part IV</td>
<td>17</td>
<td>0.903</td>
<td>0.817</td>
<td>Approx Chi-Square 2945.270</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DF 136.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sig. .000</td>
</tr>
</tbody>
</table>

Table 3.1: Showing Tests of Validity of the Questionnaire

3.5 DATA COLLECTION

The study aimed at employees working in the software and service companies in India. For the data collection 32 companies were selected randomly from the four major IT hubs of India viz. NCR (Delhi, Gurgaon, Noida, Faridabad), Hyderabad, Chennai and Mangalore, eight from each hub through random sampling. A total of 900 questionnaires were sent to these 32 companies with a request to get these filled from the software professionals having an experience of more than 4
years of handling Web based software projects (preferably project Managers and above).

First reminder was sent after 15 days of the first email followed by second reminder with the gap of 15 days. Some of the companies were very helpful and distributed the questionnaire to all the project lead, managers and directors. While for other companies emails and Telephone was extensively used to make them understand the purpose of the research and assure them that the data so provided will be used only for academic research.

The Information was finally gathered through questionnaire and telephonic interviews were also held in order to substantiate the data gathered. Senior project managers and Directors of some of the companies were also contacted individually to build rapport and gauge their understanding about project specific risks and organizational climate. Only 340 filled in questionnaires were received out of which only 300 were found to be fully filled in, the rest 40 were discarded due to incomplete information. Thus, with the unconditional assistance of various associates, data was collected comfortably from the respondents working in the selected software companies.

3.6 DATA ANALYSIS PROCEDURE

Statistical Package for the Social Sciences (SPSS) version 17.0 was religiously used for the statistical analyses. Coding of variables in quantitative research is very critical for better interpretation of results. Age, total experience,
designation, team size, total duration and total value of the project were all coded and were entered in to the computer. The questions and responses were coded and entered in the computer using Microsoft Excel software. Required analysis was done with the aid of Statistical Package for Social Sciences 17.0 Version. Certain statistical methods were applied on the data to get the results which were analyzed. The procedures used for the analysis of the research questions of this study are described below.

3.6.1 PART I: EXPLORING THE DIMENSIONS OF WEB DATA DESIGN SOFTWARE RISK IN SOFTWARE PROJECTS

RESEARCH QUESTION 1.

What are the project specific risk factors that impact the success and the three performance constructs (budget, schedule and quality) of the project? The analysis of this question was done in two parts. First was on the basis of secondary data and the second was on the basis of primary data collected. Identifying factors through secondary data was the beginning of the research study. In order to validate the results and assess them in light of primary data, the second part was carried out. For the secondary data analysis, scoring model was used for ranking the risks.

The risks identified by various researchers were coded depending on the number of times that risk has appeared in various studies and the rank that has been assigned by the respective researcher. Accordingly, the total weighted average score was calculated to get the ranks for the risk factors. Based on
these rankings a list of top ten risks affecting the software projects globally was prepared.

For the primary data analysis, factor analysis was done to extract smaller number of linear combinations out of 23 risk items which are essentially responsible for the project delay or failure. Principal component analysis was used as the method of extraction and varimax was used as the rotation method. Four dimensions of risk were extracted through factor analysis. After factor analysis, means and standard deviations were calculated for each of the dimensions. On the basis of their means, the risk dimensions were ranked.

**RESEARCH QUESTION 2.**

How much do these risk dimensions vary across the demographics and project characteristics? In order to study how these dimensions of risk vary across the personal and project characteristics of the respondents, Duncan’s mean test was used. The tests revealed how these dimensions diverge in case of three levels of designation, experience and age. Duncan’s mean test was also used to reveal how the risk dimensions deviate among the different team sizes, duration of the project and the total value of the project.

**3.6.2 PART II: EXPLORING ORGANIZATIONAL CLIMATE DIMENSIONS PRESENT IN THE INDIAN SOFTWARE COMPINES**

**RESEARCH QUESTION 3.**

What are the organizational climate dimensions that contribute towards the success and the three performance constructs (budget, schedule and quality) of
the software projects? In order to identify the organizational climate dimensions that impact the project’s success, factor analysis was done to extract smaller number of linear combinations out of 17 items under the organization’s climate. Principal component analysis was the method of extraction and varimax was the rotation method. Four factors were extracted whose Given values were more than 1. After factor analysis, means and standard deviations were calculated in order to rank each of the dimensions.

**RESEARCH QUESTION 4.**

How much do the climate dimensions vary across the demographics and project characteristics? The dimensions of organization’s climate were compared across the various personal and project characteristics of the respondents. Again, Duncan’s mean test were done to see if there was any significant difference among the perceived values across age, experience, designation, project team size, project duration and total value of the project.

**3.6.3 PART III: PROPOSITION OF A REGRESSION MODEL**

**RESEARCH QUESTION 5.**

How much do the dimensions of climate and demographic characteristics impact the project specific risk dimension of the software project? Correlations were used to assess the relationship of the organizational climate dimensions and demographic characteristics with the project specific risk factors. The four dimensions of climate along with demographic
characteristics were put for regression with the four dimensions of software risks as dependent variables individually.

**RESEARCH QUESTION 6.**

How much do the project specific risk and organizational climate dimensions affect the success and the three performance constructs (budget, schedule and quality) of the software projects? For assessing the impact of risk and organizational climate dimensions on the overall success of the project and the three performance constructs, stepwise regression was conducted. The dimensions of risk and climate were taken as independent variables while the overall success and the three success constructs were taken as dependent variables respectively. Before this process, correlations of the dimensions with the success of the project and the three performance constructs were computed. It was followed by regression, which gave quite meritorious results.

**3.6.4 PART IV: VALIDATION OF THE PROPOSED MODEL OF WEB DATA DESIGN**

**RESEARCH QUESTION 7.**

How do the risk dimensions and the organizational climate dimensions together affect the success of the software projects in real life? Four case studies involving the real life projects were prepared. The post mortem analysis of two failed projects and two Successful projects was conducted and presented in form of a case. For this, structured interview, were conducted with four senior project
managers with an average experience of 12 years in handling various software projects along with their team members.

3.7 FORMATION OF HYPOTHESES

For testing purpose, some of the above research questions were converted into hypotheses. The dimensions of risk and organization’s climate were used for forming the corresponding hypotheses, each addressing the overall success and the three performance constructs. The demographic characteristics and organizational climate dimensions were also used for forming the hypothesis relating to the various risk dimensions identified using factor analysis. The hypothesis for testing the relation of the organizational climate dimensions and demographic characteristics with software risk dimensions is as follows. The diagrammatic representation of the relationship is shown in figure 3.1.

H1: The demographic characteristics and the organizational climate dimensions affect the project specific Web Data Design risk factors.

![Diagram showing Conceptual Relationship between Organizational Climate Dimensions, Demographics Characteristics, and Project Specific Risk Dimensions in the Software Project](image.png)
The hypotheses for testing the relation of the organizational climate dimensions and software risk dimensions with the overall success and three performance constructs are as follows. The same has been diagrammatically shown in figure 3.2.

**H2**: The organizational climate dimensions and project specific risk factors affect the overall success of the software project.

**H3**: The organization’s climate and project specific risk dimensions affect the budget performance of the software projects.

**H4**: The organization’s climate and project specific risk dimensions affect the schedule performance of the software projects.

**H5**: The organization’s climate and project specific risk dimensions affect the quality performance of the software projects.

**Figure 3.2**: Showing Conceptual Relationship between Organizational Climate Dimensions The Project Specific Risk Dimensions and Success (Overall and the Three Success Performance Constructs) of the Software Project.
3.8 CONCLUDING REMARKS

In this chapter research design has been presented. Use of Web Data Design pattern questionnaire, interview and case study methodology has been extensively used for this exploratory research. The details of research methodology, questionnaire design, its validation and administration are discussed. The research questions and the formulation of hypothesis are also highlighted.