CHAPTER 3

RESEARCH METHODOLOGY

3.1 INTRODUCTION

This chapter presents the methodology adopted for conducting the study. The aim of this study is to understand the overall state of BI&A, the maturity levels of BI&A capability and dimensions which influence the maturity of BI&A in organizations, as well as to assess the effectiveness of BI&A of organizations in India. This study also aims to identify critical success factors which influence the maturity of BI&A capability. An effort has been made to meet the objectives of this research and bridge the gaps identified from literature review, as discussed in the previous chapter, using various research methods.

This chapter has been organized as follows: The section 3.2 describes the research methodology adopted starting from literature review, identification of critical success factors by Expert Panel, questionnaire development, data analysis and qualitative analysis for the case organizations. Table 3.1 shows the objectives of the research derived from the gaps found as well as the method used to address the objective.

Section 3.3 describes the step-wise process of consolidation of 108 dimensions to identify critical success factors by the Expert Panel. Section 3.4 discusses the development of the tools used for data collection – the questionnaire for quantitative analysis and the interview questions for the case organizations. Section 3.5 describes the sampling procedure and the profile of unit of sampling – the individual respondents and the unit of analysis – the organizations.

Section 3.6 presents the data collection and analysis which includes factor analysis for combination of relevant and meaningful measurement items for each factor, cluster analysis to group the organizations into clusters and thematic analysis to analyse the qualitative data

obtained from the interviews with case organizations. Section 3.7 presents the concluding remarks.

3.2 RESEARCH METHODOLOGY ADOPTED

As is clear from the research objectives outlined and mentioned in Chapter 1, an attempt has been made to understand the overall state of BI&A, the maturity levels of BI&A capability and dimensions which influence the maturity of BI&A in organizations, as well as to assess the effectiveness of BI&A of organizations in India and identify critical success factors which influence the maturity of BI&A capability. Various research methods have been used to address the gaps and objectives as shown in Table 3.1.

Table 3.1 Gaps, Objectives and Research methods used

S.No	Gaps in the literature	Objectives	Research methods used
G1	There are twenty-nine maturity models for BI&A originated from research and practice, yet there is no single model which is a standard like the Capability Maturity Model (CMM) for software development.	O1: To study the dimensions and maturity levels needed to build analytics capability in organizations from the existing maturity models of BI&A	• Literature Review
G2	There are no clear guidelines for managers to make a decision on which of these models to use. There is not enough documentation on how to select the right model for the organization.		
G3	Some studies have focused only on assessing the business value of BI&A and not on understanding what is needed to build the BI&A capability in organizations.	O2: To identify critical success factors which influence the maturity of BI&A capability of organizations in India	• Literature review • Expert Panel opinion
G4	Large number of dimensions identified from the maturity models making it difficult for a manager to select the critical ones to focus on.		
G5	There is no empirical study in India, which is helpful in determining BI&A capability maturity of organizations in India.	O3: To determine the maturity level of Business Intelligence & Analytics capability (BI&A) of organizations in India	• Questionnaire based study • Factor analysis • Descriptive analytics • k-means clustering analysis • Case study method using in-depth interviews

S.No	Gaps in the literature	Objectives	Research methods used
G6	There is insufficient empirical research for organizations in India about where they may find BI&A value for the organization. Where is BI&A found to be effective? Which are the topmost functions or areas of BI&A usage?	O4: To assess the effectiveness of BI&A of organizations in India	• Questionnaire based study • Case study method using in-depth interviews • Thematic analysis for interviews

The research process followed in this study is shown in Figure 3.1. This involved extensive literature review from journals, conference papers, reports and online sources like white papers, blogs and websites to study the BI&A practise, how BI&A capability maturity was measured and what were the dimensions for measurement. It was found that the BI&A capability maturity can be measured using a maturity model. Twenty-nine maturity models for BI&A were studied, from which we found 108 dimensions which were then studied. Further, with help of an Expert Panel, these dimensions were analysed and consolidated to identify six critical success factors which influence and measure BI&A capability maturity.

A questionnaire was developed to assess the maturity of each of the six factors. This questionnaire was pre-tested and modified based on suggestions from a group of business executives who were working in the BI&A domain. A pilot test was conducted for the questionnaire after which it was administered to 183 organizations in India, from which 145 organizations gave a completed response which could be used for analysis.

Factor analysis was used to combine relevant and meaningful measurement items for each factor. Using the factor analysis scores, k-means cluster analysis was performed to group the organizations based on the maturity of the six factors. The optimum number of clusters found was six. These clusters were further analysed to understand their characteristics. One case organization was taken from each cluster to gain deeper insights through in-depth interviews. Finally, from the results and findings, conclusions were made about the state of BI&A capability maturity, usage and effectiveness of BI&A, for organizations in India.

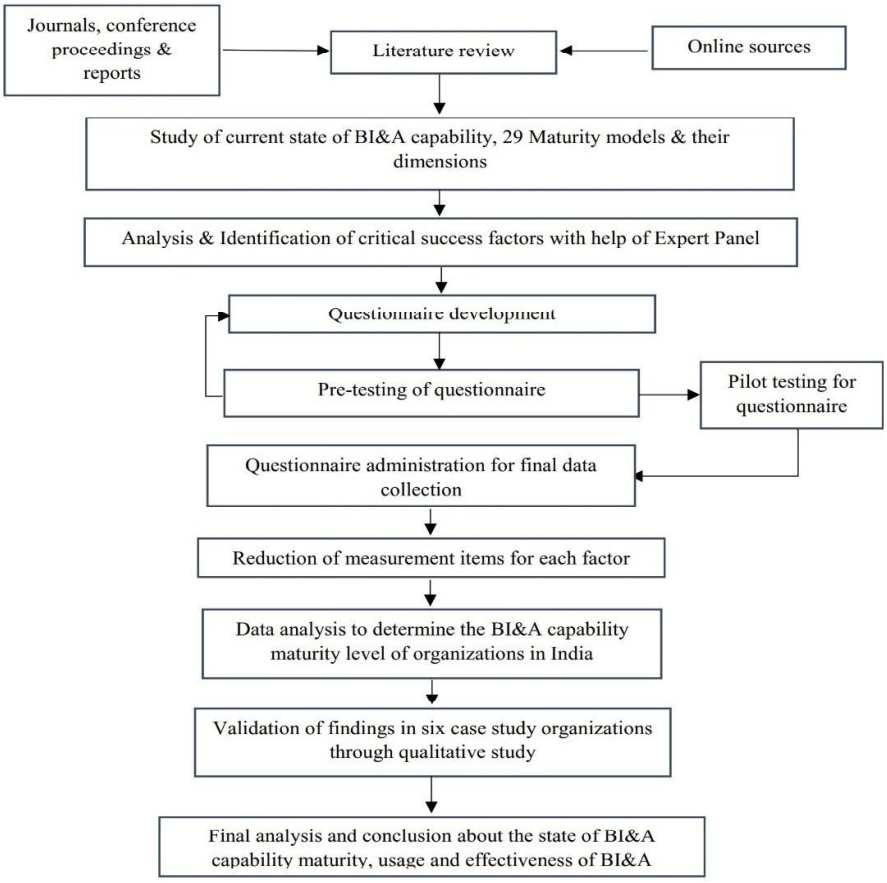

Figure 3.1 Overview of the Research Process

3.3 IDENTIFICATION OF CRITICAL SUCCESS FACTORS

There were 29 BI&A maturity models studied in literature review, from which a total of 108 dimensions were identified. This has been described in Chapter 2. From these 108, some were duplicates, some were found to be synonyms where dimensions had different names but similar description for example: 'technology' and 'data & infrastructure'. There were dimensions with descriptions indicating a common theme for example: 'Organizational practices' and 'organization culture'. There were dimensions with related areas like data architecture, metadata management, data management, master data management, warehousing architecture.

To resolve these issues, this inventory of 108 dimensions was shared with an Expert Panel comprising of a group of academicians and practitioners. There were total of twelve experts

with four academicians and eight practitioners who studied and consolidated these dimensions using the following steps:

Step 1- Dimensions with same name, for example, process or people were removed. This reduced the 108 dimensions to 67 dimensions.

Step 2 - The dimensions which had different names but similar definitions or description in literature review indicating a common theme for example: Strategy, BI Strategic Alignment, were identified and duplicates were removed. After this step the 67 dimensions were reduced to 42 dimensions.

Step 3- Finally the dimensions with related areas covered were clubbed together to make a factor, for example: Data Management covered the areas of data quality, data architecture, metadata management, data management, master data management, warehousing architecture. Each Expert clubbed these into different set of factors ranging from 3 to 10. The total union of factors from all the Experts was 20 as seen in Table 3.2.

Step 4 –From these 20, there were six factors which were selected by majority of the experts, as seen in the Table 2.4. These were identified to be the critical success factors.

These six factors include people, process, technology, data management, organization culture and strategic alignment with BI&A. As seen in Table 3.2, all 12 experts identified data management and organization culture as critical factors, 11 experts identified people and technology as critical factor, 10 identified process as a critical factor and 7 identified strategic alignment with BI&A as a critical success factor. The rest has lower consensus of less than or equal to 4 experts and hence were discarded as 'critical'.

Table 3.2 Identification of critical factors by Expert Panel members

		Expert-1	Expert -2	Expert-3	Expert -4	Expert-5	Expert -6	Expert-7	Expert -8	Expert-9	Expert -10	Expert-11	Expert -12	Total
S.No	**No of factors → Factor Name**	**8**	**7**	**9**	**9**	**8**	**10**	**6**	**3**	**8**	**6**	**8**	**6**	
1	Data/ Data Management	Y	Y	Y	y	Y	y	Y	y	y	Y	Y	y	12
2	Culture	Y	Y	Y	y	Y	y	y	y	y	Y	Y	y	12
3	People	Y	Y	No	y	Y	y	y	y	y	Y	Y	y	11
4	Technology / Infrastructure	Y	Y	Y	y	Y	no	y	y	y	Y	Y	y	11
5	Process	Y	Y	Y	y	Y	y	y	No	no	Y	y	y	10
6	Strategic alignment with BI&A	No	no	Y	y	No	no	no	y	y	Y	y	y	7
7	Impact	Y	Y	no	y	No	no	no	no	y	No	no	no	4
8	Leadership	Y	no	no	no	No	y	no	y	no	No	y	no	4
9	Information Architecture	No	no	no	y		no	no	y	y	No	no	no	3
10	Costs & Benefits	Y	no	no	no	No	y	no	No	no	No	no	no	2
11	Knowledge Management / Info Quality	No	Y	y			no				No	y	no	3
12	Visualization aid to decision making	No	Y	no			no							1
13	BI Environment			y			no							1
14	Analytical Environment			y			y							2
15	Outcomes/business performance					Y	y							2
16	Structure					Y								1
17	Governance					Y	y							2
18	Applications						y			y				2
19	Enterprise Penetration& Position						y							1
20	Strategy							y						1

The critical success factors have been named as follows: 'Data Management', 'Enterprise Process', 'People skills', 'Organizational culture', 'Strategic alignment with BI&A', and 'Infrastructure & Technology'. An acronym "DEPOSIT" has been proposed for these critical success factors using the first letter of each factor. The typology and measurement items in previous studies along with the six factors is shown in Table 3.3. These measurement items were used to develop the questionnaire from these six factors. The description of these factors as found and consolidated from literature review has been given below.

Table 3.3 Typology & Measurement items for the six critical success factors

Typology with references	Measurement items	Critical Factor
Architecture (Watson et al., 2001); (Dinter, 2012), Data (Watson et al., 2001); (Cates, Gill, & Zeituny, 2005); (Davenport and Harris, 2007), data management (Eckerson, 2004); (Halper and Krishnan, 2014), Warehousing architecture (Tan et al., 2011), Master Data Management (Tan et al., 2011); (Chuah and Wong, 2012), Information quality(Tan et al., 2011); (Lukman et al., 2011); (Chuah and Wong, 2012), Data Warehousing (Chuah and Wong, 2012), Metadata management (Sen et al., 2012); (Chuah and Wong, 2012), Technology – technical architecture, data management, information (Dinter, 2012), Data – Management & Analytics (Comuzzi and Patel, 2016), Technology – Information Management (Comuzzi and Patel, 2016), Security and compliance for analytical assets (Grossman, 2018), Data & infrastructure (List and Burciaga, 2014)	• Good quality and consistent data • Strong data orientation • Central data repositories • Integrated, accurate, common data in a central warehouse/data lake • Custodian of data • Data management is executive concern • Quality of daily operational data. • Data access • Data architecture type • Master data management • Role of data steward role for master data management • Reliability perception of reports generated by the MIS/BI&A team	Data Management
BI Portfolio management (Williams et al., 2004), Partnership between business units and IT (Williams et al., 2004), Business Proficiency (Shaaban et al., 2012), Organization (Eckerson, 2004); change management (Chuah and Wong, 2012); Enterprise (Davenport and Harris, 2007); Deployment - practices (Lahrmann, 2011), use of BI – Organizational use (Lahrmann, 2011), Governance (Cosic et al., 2012); (Halper and Krishnan, 2014), (Comuzzi and Patel, 2016); Organization-structure, processes, profitability strategy (Dinter, 2012), analytics capability (List and Burciaga, 2014); Implementation (White and	• Width of BI use in organization • Usage in everyday decisions • BI&A facilitating team across enterprise • Enterprise wide management of resources • Defined processes for implementing change • Standardization of processes • Integration of processes	Enterprise Processes

Typology with references	Measurement items	Critical Factor
Oestreich, 2017); Analytic governance structure (Grossman, 2018)		
Impact on user skills & jobs (Watson et al., 2001), DW staff (Watson et al., 2001), DW users (Watson et al., 2001), People (Cates et al., 2005); (Hostmann et al., 2006); (Sacu and Spruit, 2010);(G. Lahrmann, Marx, Winter, & Wortmann, 2011); (Cosic et al., 2012); (Chuah and Wong, 2012) , analyst (Davenport and Harris, 2007), , Organization – People (Comuzzi and Patel, 2016), People, Skills & Organization (White and Oestreich, 2017)	• Awareness & acceptance to learn new skills • Level of skills • Encouragement by top management • Recognition & appreciation for analytical skills • Continued training • Awareness of big data potential • Skills to make use of Big Data • Recruitment & hiring of analytical minded employees. • Cultivation of analytical amateurs • Professional analysts are managed as a strategic workforce • Highly capable analysts are explicitly recruited, developed and deployed • Data analytics activities	People skills
Information & analysis usage culture (Williams et al., 2004); (Popovic et al., 2010), decision process engineering culture (Williams et al., 2004), continuous process improvement culture (Williams et al., 2004); (Popovic et al., 2010), analytics (Eckerson, 2004); (Tan et al., 2011); (Halper and Krishnan, 2014), organization (Eckerson, 2004); (Shaaban et al., 2012); (Raber, 2012); (Halper and Krishnan, 2014), leadership (Davenport and Harris, 2007), culture (Lahrmann et al., 2011); (Cosic et al., 2012), risk & reward (Lahrmann et al., 2011), analytical (Chuah and Wong, 2012), organizational culture (Chuah and Wong, 2012), Social system (Raber, 2012), organization – culture (Comuzzi and Patel, 2016), Organizational practices and culture (List and Burciaga, 2014)	• Familiarity with using information, analytical frameworks and quantitative analysis • Norms towards systematic use of gathering, analyzing and disseminating data • Leadership support • Developing analytics capabilities as top priority • Big data as a trusted capability for decision making • Practices to enable effective use of analytics	Organizational Culture
BI strategic alignment (Williams et al., 2004), Targets (Davenport and Harris, 2007), Business enablement (Shaaban et al., 2012), Business perspectives (Lukman et al., 2011), strategy & program management (Shaaban et al., 2012), Strategic alignment – strategy (Comuzzi and Patel, 2016), Data & Analytics Vision & Strategy (White and Oestreich, 2017), developing analytical strategy (Grossman, 2018).	• Overall strategy and vision • Cohesion & alignment between business, IT and BI&A strategy • Engagement with business stakeholders • Strategic targets & Opportunities based upon available resources • Big data initiatives sponsored by top mgmt. • Use of Big data for decision-making	Strategic alignment with BI&A

Typology with references	Measurement items	Critical Factor
	• Purpose of BI&A	
Stability of production environment (Watson et al., 2001), Applications (Watson et al., 2001), BI & DW Technical readiness (Williams et al., 2004), BI portfolio management (Williams et al., 2004), metrics and technology (Hostmann et al., 2006), infrastructure (Eckerson, 2004); (Lahrmann et al., 2011); (Chuah and Wong, 2012); (Halper and Krishnan, 2014), Deployment (Lahrmann, 2011); technology (Lahrmann et al., 2011); (Sacu and Spruit, 2010); (Shaaban et al., 2012); (Lukman et al., 2011); (Cosic et al., 2012); (Comuzzi and Patel, 2016), information technology (Cates et al., 2005); (Shaaban et al., 2012), technical systems (Raber, 2012), technical architecture (Dinter, 2012), penetration level (Dinter, 2012), Technology & Solutions (White and Oestreich, 2017), Analytic infrastructure (Grossman, 2018)	• State of technology architecture to support real time data warehouse / data lake • Methods which are working for managing and implementing BI that creates value • BI applications usage • Use of devices for delivery of analytics • Technology & architecture for big data environment • There are standard physical facilities for analytics • Use of advanced BI&A tools • Vendor management skills for data and analytics platforms	Infrastructure & Technology

Data Management

Data Management describes the processes and technologies required for data capture, storage and retrieval for analysis (Gandomi and Haider, 2015). It deals with warehousing architecture, data quality and consistency, data accessibility, security and privacy of data. It evaluates how the processing and storage of data is managed for structure and unstructured data and how is the data integrated across business functions. It also evaluates the reliability perception of reports generated by BI&A. As mentioned in her study by (Veena Tewari Nandi, 2012), Database Management & Analytics give organizations the leeway to be more responsive and reduce costs associated with conducting analysis on their data by a significant amount.

Enterprise Processes

Enterprise processes evaluate how BI&A opportunities are managed at enterprise level for different functions to create business value. This requires a fulfilling partnership between business and IT, effective approaches to implement change in the business, supporting

solutions, effective control of big data capability management and managing the use of BI&A resources within an organization. It also evaluates the accountabilities for aligning business analytics initiatives with organizational objectives.

As mentioned by (Harriott, 2013), "A centralized analytics team reporting directly to the CEO helps the organization examine the data with an unbiased view rather than individual departments aligning BI&A to their own particular agendas".

"To ensure successful adoption of any technology implementation, the organization is required to spread awareness of the technology's usefulness, ease of use and promote organizational changes to enhance communication and learning" (Khan and Brock, 2017). This dimension evaluates the regular use of BI&A across the organization, it evaluates whether there is a facilitating team, infrastructure and other resources for BI&A across the organization. It evaluates the level of standardization and integration of processes. In a study by (Arnott et al., 2017) it was found that Enterprise BI&A systems are an effective support for operational and management decisions.

Organizational culture

Organization culture evaluates how the organization uses analytics and how advanced it is, how widespread it is, how the business strategy, leadership, skills and investments support a successful analytics program.

This dimension assesses the way people in the organization think about and recognize BI&A as an important and trusted capability for an organization, what norms the organization follows in using information, analytical frameworks, quantitative analysis and structured decision processes as well as the practices enabled for effective use of analytics. Is the top management playing a significant role in moving their organizations towards analytical decision-making? It has been experienced that if top management and leadership supports

analytical initiatives, the organization is more likely to succeed. This may be because they have the power to influence the people in their organization, power to deploy investments, people and time to build effective analytical caapbilities (Davenport and Harris, 2010), (Ellingsworth, 2012).

The success of implementation of a BI&A program rests upon leadership to promote performance management and allocate resources for an appropriate architecture, skills and technology as suggested by (Ghosh and Scott, 2011) whatever maybe the industry. Hence leadership plays a huge role in supporting a successful B&A program. (Holsapple et al., 2014) suggests that awareness and commitment to an organization's strategy and an analytics friendly culture are key factors for BI&A to actually work in an organization. The challenges for usage of BI&A vary from one organization to another so understanding the organization culture is very important for successful BI&A implementation (Thamir and Poulis, 2015).

Strategic alignment with BI&A

Strategic alignment with BI&A refers to consistency between business strategy, business organization and processes, IT strategy, IT organization and processes and IT infrastructure.

It refers to how well aligned are the big data strategies considered by top management in the definition of the organization's strategy and whether big data is used for decision making. It describes the advancing and evolving types of business needs and problems that are solved with BI&A solutions and how the organization is using BI&A for decision making. It describes the engagement with business stakeholders and the strategic targets and opportunities based upon available resources as described in Table 2.5.

People skills

People skills refer to level of existing skills, the recruitment, training, development and assessment of all those individuals within an organization who use BI&A as part of their job

function and are involved in the use of information. This dimension evaluates the extent to which employees within an organization are aware of the potential of emerging technologies in BI&A and/or knowledgeable about it. It evaluates whether there is awareness and acceptance to learn new skills including those to do with big data. It evaluates whether there is recognition and appreciation for analytical skills, whether highly capable analysts are explicitly recruited, developed and deployed. (Božič and Dimovski, 2019b), in their study mention that people skills with domain and data knowledge are crucial for successfully delivering value from BI&A use.

Infrastructure & Technology

Infrastructure & Technology describes the rapidly advancing nature of the information solutions an organization adopts to service its various business needs and technology required to acquire, manage and extract knowledge from the evolving nature of data effectively. It describes the established processes for maintaining and expanding the warehouse (ETL). This factor refers to the increasing and mature use of data and analytics platforms, integration, infrastructure and data as a service (DaaS). It refers to vast range of BI&A applications which can improve the functional performance of an organization in the various business units, including applications that aid with driving revenue growth, optimize costs and profits.

This factor describes the state of technology architecture to support real time data warehouse and data lake, proven methods for managing & deploying BI&A that creates value, use of devices for delivery of analytics, what are the BI&A applications used in organization, whether there is a technology and architecture for big data environments, how advanced are the BI&A tools and what are the available vendor management skills as data and analytics are managed through platforms now in most organizations.

These six critical factors were used for developing a questionnaire to understand the state of BI&A and the BI&A maturity level of organizations in India. The development of the questionnaire is described in the next section.

3.4 DEVELOPMENT OF THE TOOL

Data collection has been done in this study using two tools – questionnaire and interviews. These tools have been designed based on the research objectives outlined and mentioned in Chapter 1.

3.4.1 Questionnaire development

A well designed questionnaire includes a standard set of questions and it can be successfully used to collect information from a large sample in a reasonably short time period (Chawla and Sondhi, 2015). As this study required the information to be gathered from a large number of organizations, the questionnaire seemed to be the most cost effective and efficient way to collect information about the state of BI&A in organizations.

The questionnaire was developed based on the six critical factors which have been described in Chapter 2. Some of the earlier survey instruments from (Popovic et al., 2010), (Lahrmann, 2011), (Lukman et al., 2011), (Tan et al., 2011), (Dinter, 2012), (Olszak, 2013), (Raber et al., 2013) and (White and Oestreich, 2017) were referred to for better understanding.

The questionnaire has eight sections with a total of 43 questions (see Appendix 1). A five point Likert scale was used where response from 1 - 5 was in terms of "Strongly disagree", "Disagree", "Neutral/Not sure", "Agree" and "Strongly agree". The questions were developed to elicit responses where 1 depicted low maturity and 5 depicted high maturity of a factor.

The first section was designed for understanding the state of BI&A, the usage and effectiveness seen in the respondent's organization. The usage and benefits of BI&A found in prior research were used as the list of responses to the questions "Which functions in your organization use BI&A for driving decisions?" (Table 3.4) and "Where do you see the effectiveness and benefits of BI&A?" (Table 3.5).

Table 3.4 Sources for Responses to BI&A usage questions

S.No	Responses	Author
1	Marketing, Sales, Finance, Production	(Olszak, 2013)
2	Supply chain, Supplier Relationship, Customer Relationship, Human Resource management	(Popovic et al., 2010)
3	Inventory management, procurement, Supplier relationship	(Raber et al., 2013)
4	Projects, Information Technology, Risk Management	Response as suggested by Experts during pre-testing

Table 3.5 Sources for Responses to the effectiveness of BI&A

S.No	Responses	Author
1	Better access to data, Better informed decision making, Improved efficiency of internal processes, New way of doing business, Reduced operational costs, Transparency of information	(Olszak, 2013)
2	Reduction of lost sales, increased employee productivity, \Reduced operational costs, Improved efficiency of internal processes	(Raber et al., 2013)
3	Improved customer service, Enhanced profit margin, Improved competitive advantage, Increased Return on Investment (ROI)	(Elbashir et al., 2008)

The next six sections were developed for each of the six factors. The second section had questions regarding the factor 'Strategic alignment with BI&A' (SA), the third section had questions regarding the factor 'Data Management' (DM), the fourth section was on questions regarding 'Enterprise Processes' (EP), fifth section had 'Organizational culture' (OC), sixth section had questions regarding 'People skills' (PS) and seventh section was regarding 'Infrastructure & Technology' (IT). Each of these sections had one direct question to capture the respondent's perceived maturity score for the factor based on the scale from 1 to 5. The last section had demographic details of the respondent and the organization.

The questionnaire was pre-tested with a group of six executives working in BI&A practice and one academician. There were six mid and senior level respondents who were selected for their knowledge about BI&A practise and would understand the objective of the questionnaire. The average experience of these respondents was 18 years and the designations were as follows: Head BI&A, Head – Data Analytics Consulting, Associate Professor – IT, GM – Information Systems (IS), Chief Manager – IS and Product Lead supply. In the pre-test, the respondents focussed on questionnaire instructions, document ease of use, content and wording of questions. The objective of the questionnaire was to determine the level of BI&A capability maturity in organizations. Their focus points were:

- Are the instructions clear and concise?
- Are all the questions simple to understand?
- Are all questions exhaustive or is there anything missing for a factor?

Based on the suggestions that came from the six respondents, the changes made to the questionnaire are summarized as follows:

- Some questions were re-worded to make them more specific and simpler to understand
- Some questions were broken up into two questions to make it simpler to answer
- For some questions an option to enter a qualitative response was added
- Some questions which were overlapping and related, were combined together
- Some questions were made into a grid for easy administration

The revised questionnaire was uploaded on Survey-Monkey and was then pilot tested by six business users from the industry. The respondents were business users working in the BI&A practice with an average work experience of 3 years. The Pilot test was conducted to check for user friendliness, ease of understanding and to check time taken to complete the survey.

The overall feedback from the users was that the questionnaire was user friendly and comprehensive. The average time taken to complete the survey was found to be 20 minutes. There were some minor changes required in the navigation within the questionnaire which were incorporated before administering to the larger sample.

3.4.2 Interviews for Case Study method

Case study method was selected to understand one organization from each cluster in depth. This method is used to examine an occurrence in its natural setting to collect information from individuals, groups of people, or organizations. Case studies offer in-depth understanding of current situations within their organizational context (Aberdeen, 2013), (Yin: Case study research: Design and methods (Vol. 5) fifth edition, SAGE publications, 2003), (Yin, 2012). This methodology provides better explanations and understanding of the examined situation which may otherwise be lost in using other quantitative methods (Miles and Huberman, 1994).

The qualitative case study method was used in this study to know more about the characteristics of organizations in each of the clusters. The organization chosen was based on convenience sampling largely depending on the availability of the respondent.

We chose in-depth interviews as the method of gathering information from one case organization in each cluster, as semi-structured, in-depth interviews are considered to be the most effective method of gathering information as these are flexible and accessible (Božič and Dimovski, 2019a). This kind of an interview has a more defined format and only the broad areas to be investigated are formulated (Chawla and Sondhi, 2015).

The interview questions for this study were semi-structured and exploratory in nature. The questions brought out insights for understanding the state of BI&A, the functions it is used for and where is the effectiveness of BI&A seen.

3.5 SAMPLING PROCEDURE

A simple random sampling procedure has been followed here. The unit of analysis here is the organization and the unit of sampling is the individual respondent.

The organizations from which data was collected were selected based on their size. We have used Gartner's definition for size of organization. As per Gartner, the metric for size is number of employees in the organization. Gartner is a globally accepted research and consulting company, with a presence in India. They conduct research with organizations in India for business and technology. As per Gartner, small businesses are defined as organizations which have employees fewer than 100. It is observed that they often face different IT challenges and their IT resources are highly constrained. It may then be difficult to find any BI&A initiatives in these organizations. Whereas the midsize enterprises are those organizations which have 100-999 employees, and large enterprises are the ones with more than 1000 employees. It is observed that these enterprises would have enough resources and operations which could be capturing and processing data with a good scope to use BI&A. The medium to large size organizations are more able to and likely to invest in different IT technologies with related employee training (Elbashir et al., 2013), (Chawla and Sondhi, 2015). Hence for this research, the midsize to large organizations have been selected for the sample.

As per CMIE Prowess, there are about 10,000 midsize to large size organizations in India based on size of employees as described above. About 2% (200 organizations) of these were taken as sample in the study. From this, the researcher was able to obtain contact for respondents in 183 organizations. These organizations were selected to represent a diverse spread of different segments of the industry in both the manufacturing sector and services sector. This gave a good breadth of understanding of state of BI&A in India across sectors. The respondent organizations were from various industries such as financial services,

banking, insurance, FMCG, oil & gas, e-commerce, construction & cement, healthcare, manufacturing, IT services & consulting, telecom, paints manufacturing, entertainment & media, aviation, chemicals, retail, agriculture, communications, textiles, IT infrastructure and transport. The names of the organizations have not been revealed due to confidentiality request from the respondents.

The respondents had to be selected on the basis of their understanding of the domain and BI&A practice. Hence respondents with a minimum experience of 3 years in the Information Technology or BI&A practice were selected. They were largely business users using BI&A to make business decisions and/or a business analyst. The complete profile of the respondents is seen in Table 4.1 in Chapter 4 - Analysis, Results & Discussion.

The questionnaire was sent to 183 organizations from which 145 organizations had given a complete response hence, about 79% rate of response. The respondents were contacted telephonically and the questionnaire was explained to them. The questionnaire was then administered online sharing the link for the Survey monkey tool. The questionnaire was active for a period of three months after which data was downloaded from survey monkey and prepared for analysis.

3.6 DATA COLLECTION AND ANALYSIS

The data collected from respondents in 145 organizations, through the questionnaire was both quantitative and qualitative. The questionnaire had both close ended and open ended questions. Next steps taken were as follows: incomplete records in survey monkey were removed, data was downloaded from survey monkey and the data file was cleaned. The missing values were replaced with the mode of all the responses for the same factor.

To check the reliability of the scales for this research, Cronbach alpha was used (Chawla and Sondhi, 2015). The values were found to be at an acceptable level, ranging from 0.79 to 0.94 for each of the six factors, indicating that further analysis could be carried out on the data.

Next, we had to combine the relevant and meaningful measurement items for each factor and drop the ones with a weak relationship. Kaiser-Meyer-Olkin (KMO) Test is a measure of how suitable the data is for Factor Analysis. The test assesses the sampling adequacy for each variable in the model as well as for the complete model. The KMO test for all the six factors are seen in the Table 3.6. Values for all six factors are > 0.80 and hence the data is appropriate to proceed with factor analysis.

Table 3.6 KMO measure of sampling adequacy

Factor	KMO measure of sampling adequacy
Strategic Alignment with BI&A	0.82
Data Management	0.84
Enterprise Processes	0.90
Organizational culture	0.83
People skills	0.91
Infrastructure & Technology	0.94

3.6.1 Exploratory Factor analysis

This study used factor analysis for combination of relevant and meaningful measurement items for each factor. Factor analysis is commonly used to reduce measurement items into a smaller set to save time and bring out easier interpretations (Yong and Pierce, 2013). It is useful for studies that involve a large number of variables or measurement items from questionnaires which may be reduced to a smaller set, to get at an underlying concept, and to help with interpretations (Rummel, 1970). (Dubey and Gunasekaran, 2015) have used exploratory factor analysis (EFA) to check the reliability and validity of measurement items for each factor. As explained by (Zainol et al., 2014), EFA is used to confirm the validity of variables or measurement items. As mentioned by (Salmasi et al., 2016), EFA is good for

detecting "misfit" measurement items. It provides a grouping of measurement items based on strong correlations.

This study used SPSS to perform factor analysis on the measurement items for each of the six factors using principle component analysis extraction with Varimax rotation for each factor in isolation. Varimax minimizes the number of variables that have high loadings on each factor and works to make small loadings even smaller.

The factor loading for a variable is an assessment of how much the variable contributes to the factor. High factor loading scores indicate that the dimensions of the factors are better accounted for by the variables. It should be 0.40 or greater as anything lower would suggest a weak relationship between the variables or measurement items (Tabachnick et al., 2019). Hence the items which have a factor loading < 0.40 have been eliminated. The signs of the factor loadings show the direction of the correlation and do not affect the interpretation of the magnitude of the factor loading or the number of factors to retain (Kline, 1994).

Table 3.7 Exploratory Factor Analysis Results

Factor	Item Label	Cronbach alpha	Factor Loading
Strategic Alignment with BI&A (SA)	SA_vision	0.85	0.75
	SA_cohesion		0.72
	SA_engage		0.75
	SA_resource		0.31
	SA_BDmgmt		0.56
	SA_BDuse		0.51
	SA_purpose		0.23
Data Management (DM)	DM_descr1	0.80	0.51
	DM_descr2		0.58
	DM_descr3		0.32
	DM_descr4		0.57
	DM_descr5		0.27
	DM_descr6		0.38
	DM_Qual		0.01
	DM_access		0.002
	DM_arch1		0.07
	DM_arch2		0.03
	DM_arch3		0.52
	DM_arch4		0.33
	DM_arch5		0.24

Factor	Item Label	Cronbach alpha	Factor Loading
	DM_mdm1		0.48
	DM_mdm2		0.50
	DM_mdm3		0.53
	DM_reports		0.33
Enterprise Processes (EP)	EP_descr1	0.92	0.64
	EP_descr2		0.63
	EP_descr3		0.57
	EP_descr4		0.70
	EP_descr5		0.74
	EP_descr6		0.63
	EP_descr7		0.66
	EP_stds		0.55
Organizational culture (OC)	OC_descr1	0.86	0.50
	OC_descr2		0.55
	OC_descr3		0.65
	OC_descr4		0.62
	OC_descr5		0.54
	OC_descr6		0.67
People skills (PS)	PS_descr1	0.93	0.58
	PS_descr2		0.70
	PS_descr3		0.61
	PS_descr4		0.45
	PS_descr5		0.66
	PS_descr6		0.48
	PS_descr7		0.59
	PS_descr8		0.53
	PS_mgmt1		0.47
	PS_mgmt2		0.62
	PS_mgmt3		0.60
	PS_activity		0.45
Infrastructure & Technology (TI)	TI_state1	0.94	0.73
	TI_state2		0.81
	TI_state3		0.77
	TI_state4		0.55
	TI_state5		0.67
	TI_state6		0.74
	TI_state7		0.75
	TI_state8		0.67

As can be seen from Table 3.7, there were two measurement items, "SA_resource" and "SA_purpose" in the factor "Strategic alignment with BI&A" with a factor loading < 0.4 and % of variance 54.6. These were removed and factor analysis done once again. The % of variance was then found to be 67.75.

There were ten measurement items in the factor "Data Management" with a factor loading < 0.4 and loading percent of variance 33.29. Statistically this suggests a weak relationship and hence after removal of these ten items, the total variance found was 57.1. The extraction method used in SPSS was Principal Component Analysis for fixed factor number =1. Hence the total variance explained is for one component. The total variance explained using the principal component extraction method for a fixed factor = 1 is given in Table 3.8. The factor values were generated for each of the six factors for the data sample with the loadings percent of variance. These factor values were further used as data points for the cluster analysis.

Table 3.8 Loading variance (%) in Factor Analysis

Factor	Loading variance (%)
Strategic Alignment with BI&A	67.75
Data Management	57.19
Enterprise Processes	64.13
Organizational culture	58.99
People skills	56.08
Infrastructure & Technology	71.09

3.6.2 K-means clustering analysis

The scores from the factor analysis were used to perform K-means clustering analysis to group the organizations into different clusters based on the maturity of the six factors. The k-means algorithm has been chosen as it is one of the most commonly used and simple algorithms. It has been used by (Lukman et al., 2011) to study the key practices which influence BI&A maturity in Slovenian organizations. K-means clustering as explained by (Bowler and Datar, 2018) is an unsupervised learning technique used to group things together by similarity. The basic objective of applying the k-means algorithm to the data was to group together organizations having similarity in the maturity of the six critical factors which were described in Chapter 2. The six critical factors were Data Management (DM), Enterprise

Processes (EP), People Skills (PS), Organizational Culture (OC), Strategic Alignment with BI&A (SA) and Infrastructure & Technology (IT).

K-means clustering was selected as the clustering method because it is represented by its centroid value which is a mean of points within a cluster whereas Hierarchical clustering is a method which builds a hierarchy of clusters (Rai and Singh, 2010). Also k-means clustering generates a specific number of disjoint, flat clusters whereas hierarchical clustering generates a hierarchy and not a partition of objects (Kaushik and Mathur, 2014). In hierarchical clustering, data points which are closer to each other get combined in one cluster. This cluster further combines with another cluster which is close to it and so on until there is only one cluster. As a result, many a times, two data points which are significantly apart, but have a close neighbour, also gets combined in one cluster. This approach was not found suitable for this study. Here it was required to group the organizations into similar levels of maturity of the six factors, hence, k-means clustering was used for this study.

The goal of the k-means algorithm was to partition the sample into k disjoint clusters so that they are homogenous amongst themselves based on the measurements used. The k-means algorithm begins with partitioning the data set into k clusters. Subsequent steps modify the partition to reduce the sum of the distances of each record from its cluster centroid (Patel, 2018).

K-means clustering technique was performed using the values created from factor analysis, for k = 2 to k=10 clusters using SPSS. This was done one step at a time to find the best suitable value of k. This has been presented in Table 3.8 to 3.14, with the visual representation of the final cluster centres after the iterations, shown in Figure 3.3 to 3.10. The visual representation makes it easy to observe the varying centroid values of the six factors in the clusters.

Table 3.9 Final Cluster centres for k=2

Factor	1	2
Strategic Alignment with BI&A	0.56456	-0.75579
Data Management	0.5872	-0.7861
Enterprise Processes	0.60989	-0.81647
Organizational Culture	0.61953	-0.82937
People Skills	0.61106	-0.81803
Technology & Infrastructure	0.59897	-0.80185

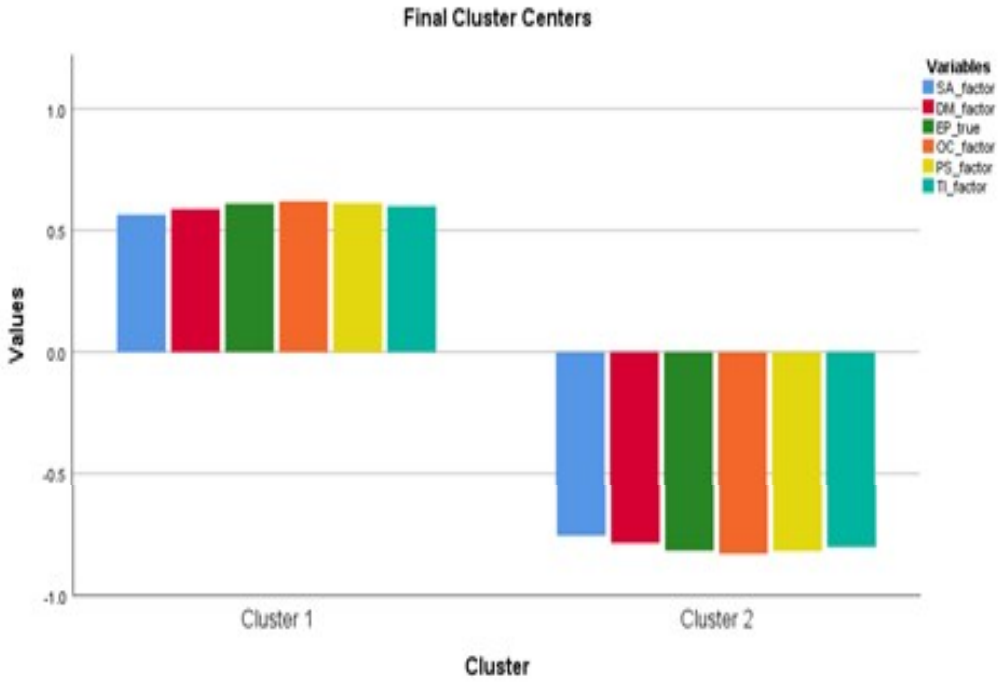

Figure 3.2 Final Cluster centres k=2

Table 3.9 shows the final cluster centres for k=2. The Figure 3.2 shows the visual representation of the final cluster centres for k=2. It is clearly observed that all six factors in Cluster 1 have positive centroid values while those in Cluster 2 have negative centroid values. The clusters are very clearly distinct from each other. All the respondent organizations have been divided into two clusters with 83 organizations in Cluster 1 which have positive centroid values for six factors and 62 organizations in Cluster 2 which have negative centroid values for six factors.

Table 3.10 Final cluster centres k=3

Factor	1	2	3
Strategic Alignment with BI&A	0.90073	-1.08399	0.00443
Data Management	0.98126	-0.93873	-0.12728
Enterprise Processes	1.06096	-1.3009	0.01835
Organizational Culture	1.08342	-1.13536	-0.08658
People Skills	1.04375	-1.09774	-0.08125
Technology & Infrastructure	1.07389	-1.1653	-0.06404

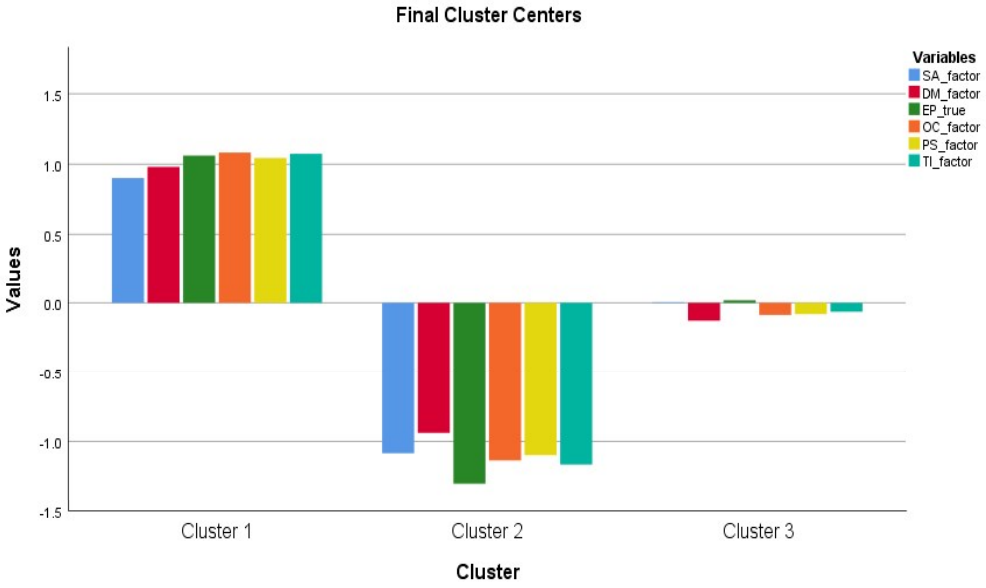

Figure 3.3 Final Cluster centres k=3

Table 3.10 shows the final cluster centres for k=3. The Figure 3.3 shows the visual representation of the final cluster centres for k=3. It is clearly observed that all six factors in Cluster 1 have positive centroid values while Cluster 2 has larger negative centroid values as compared to those in cluster 3 which have values closer to 0. There are 43 organizations in Cluster 1, 36 organizations in Cluster 2 and 66 organizations in Cluster 3. It is observed that the centroid values for each factor of the most positive cluster, in this case Cluster 1, are higher than the centroid values of the respective factors in the most positive cluster (Cluster 1) for k=2.

Table 3.11 Final cluster centres k=4

Factor	1	2	3	4
Strategic Alignment with BI&A	1.33163	0.28990	-0.31240	-1.30337
Data Management	1.16818	0.47990	-0.59930	-0.99859
Enterprise Processes	1.30857	0.42987	-0.39144	-1.44235
Organizational Culture	1.27908	0.45516	-0.43504	-1.38332
People Skills	1.42361	0.36728	-0.38420	-1.42293
Technology & Infrastructure	1.50439	0.31646	-0.34012	-1.47317

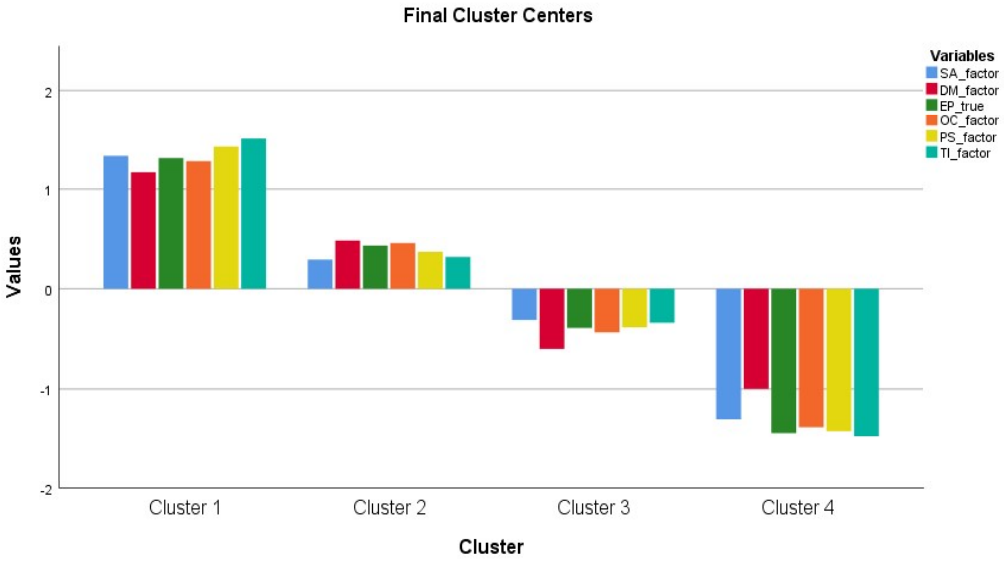

Figure 3.4 Final Cluster Centres k=4

Table 3.11 shows the final cluster centres for k=4. The Figure 3.4 shows the visual representation of the final cluster centres for k=4. It is clearly observed that all six factors in Cluster 1 have positive centroid values while those in Cluster 2 also have positive centroid values but less than those in cluster 1. Cluster 3 and 4 have negative centroid values with Cluster 3 being less negative than Cluster 4. Number of organizations in the clusters are as follows: Cluster 1 - 22, Cluster 2 – 53, Cluster 3 - 47 and Cluster 4 – 23. Cluster 1 is the most positive cluster and the centroid values of each factor here are higher than the centroid values of the most positive cluster (Cluster 1) in k=3. Also the number of organizations in the most positive cluster are reducing hence the cluster is becoming more homogenous with regard to maturity of six factors.

Table 3.12 Final cluster centres k=5

Factor	1	2	3	4	5
Strategic Alignment with BI&A	0.52248	-1.76213	-0.68327	1.45833	0.10424
Data Management	0.74898	-1.17541	-0.79707	1.35081	-0.04135
Enterprise Processes	0.82612	-1.81406	-0.80399	1.38813	0.03520
Organizational Culture	0.80939	-1.75942	-0.74411	1.44800	-0.03638
People Skills	0.56655	-1.94091	-0.69207	1.73188	0.02664
Technology & Infrastructure	0.65464	-1.83688	-0.72196	1.56698	0.02610

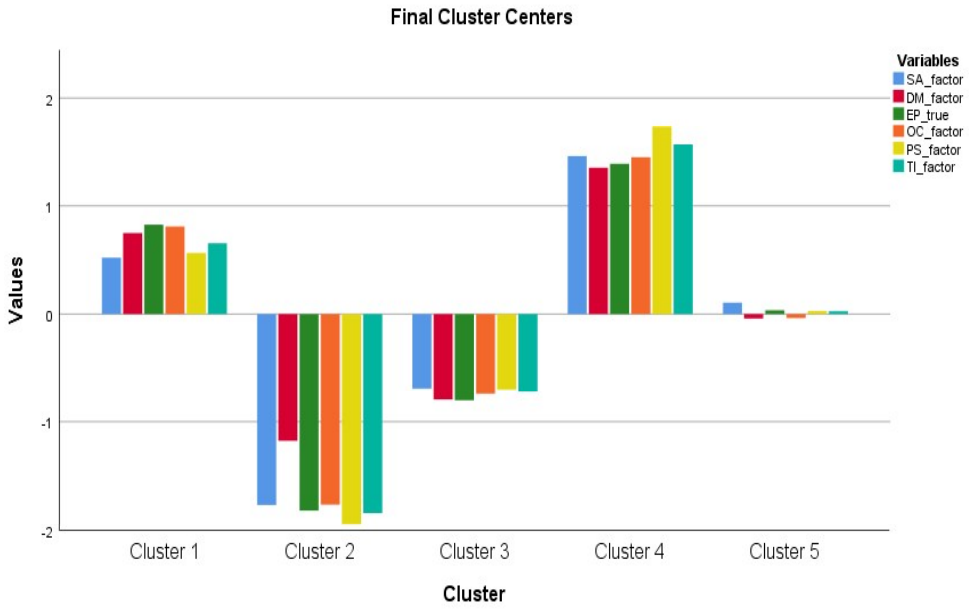

Figure 3.5 Final Cluster centres k=5

Table 3.12 shows the final cluster centres for k=5. The Figure 3.5 shows the visual representation of the final cluster centres for k=5. It is clearly observed that the cluster with the most positive centroid values is Cluster 4 and the cluster with highest negative centroid values is Cluster 2. Number of organizations in the clusters are as follows: Cluster 1 - 32, Cluster 2 – 10, Cluster 3 – 40, Cluster 4 – 16 and Cluster 5 - 47. Cluster 4 is the most positive cluster and the centroid values of each factor here are higher than the centroid values of the most positive cluster (Cluster 1) in k=4. Also the number of organizations in the most positive cluster are reducing hence the cluster is becoming more homogenous with regard to maturity of six factors.

Table 3.13 Final cluster centres k=6

Factor	1	2	3	4	5	6
Strategic Alignment with BI&A	-2.07385	0.53531	-0.46066	0.19363	1.45833	-1.07936
Data Management	-1.93340	0.72365	-0.83602	0.14284	1.35081	-0.75124
Enterprise Processes	-2.15736	0.85643	-0.33190	0.07819	1.38813	-1.33243
Organizational Culture	-2.89056	0.86367	-0.44086	0.04932	1.44800	-1.12767
People Skills	-2.42041	0.63546	-0.39819	0.03952	1.73188	-1.13182
Technology & Infrastructure	-2.33749	0.72629	-0.39704	0.04804	1.56698	-1.15654

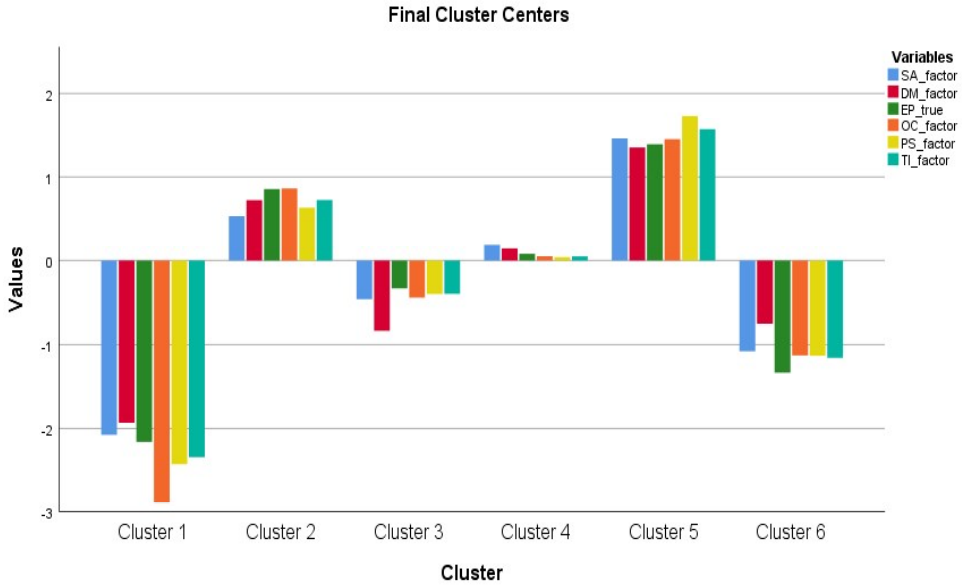

Figure 3.6 Final Cluster Centres k=6

Table 3.13 shows the final cluster centres for k=6. The Figure 3.6 shows the visual representation of the final cluster centres for k=6. It is clearly observed that the cluster with the most positive centroid values is Cluster 5 and the cluster with highest negative centroid values is Cluster 1. Number of organizations in the clusters are as follows: Cluster 1 - 3, Cluster 2 – 29, Cluster 3 – 28, Cluster 4 – 43, Cluster 5 – 16 and Cluster 6 - 26. Interestingly, the centroid values of each factor, in the most positive cluster here (Cluster 5), are equal to the centroid values of the most positive cluster (Cluster 4) in k=5. Also the number of organizations in the most positive cluster are the same as in k=5 hence the cluster seems to have stabilized with regard to maturity of six factors.

Table 3.14 Final cluster centres k=7

Factor	1	2	3	4	5	6	7
Strategic Alignment with BI&A	-0.34548	-0.94111	0.01840	-1.05512	1.45833	-1.97672	0.82768
Data Management	-1.03435	0.92341	0.03029	-0.71973	1.35081	-1.83477	0.60947
Enterprise Processes	-0.45412	0.91149	0.01451	-1.32905	1.38813	-1.97226	0.73922
Organizational Culture	-0.29032	0.77320	-0.13464	-1.11574	1.44800	-2.52440	0.73437
People Skills	-0.48393	0.77161	-0.02114	-1.11782	1.73188	-2.18574	0.50513
Technology & Infrastructure	-0.53432	0.48603	0.02564	-1.11328	1.56698	-2.31263	0.64575

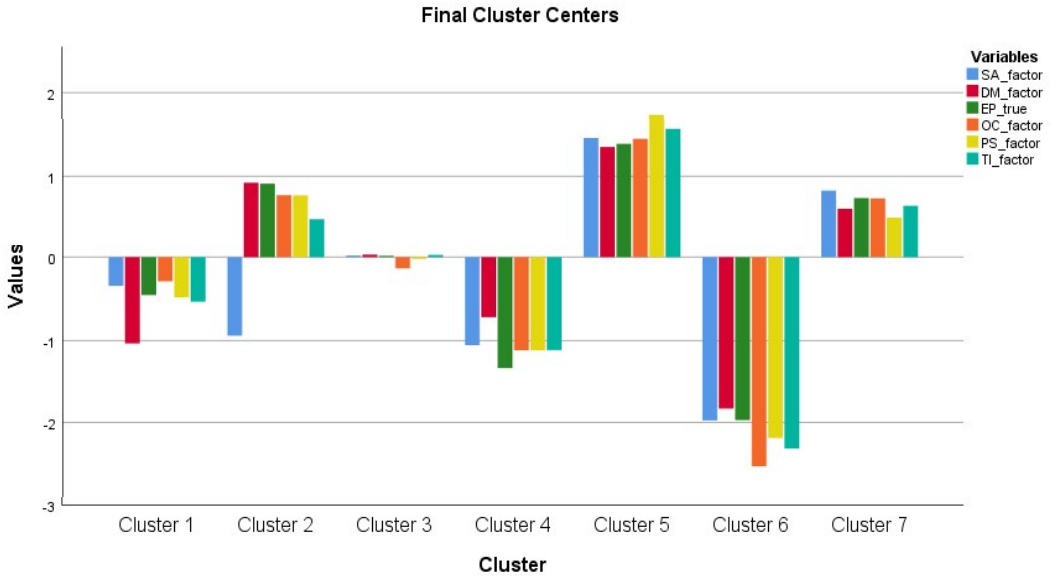

Figure 3.7 Final Cluster centres k=7

Table 3.14 shows the final cluster centres for k=7. The Figure 3.7 shows the visual representation of the final cluster centres for k=7. It is clearly observed that the cluster with the most positive centroid values is Cluster 5 and the cluster with highest negative centroid values is Cluster 6. Number of organizations in the clusters are as follows: Cluster 1 - 21, Cluster 2 – 7, Cluster 3 – 43, Cluster 4 – 25, Cluster 5 – 16, Cluster 6 – 4 and Cluster 7 - 29. Interestingly, the centroid values of each factor, in the most positive cluster here (Cluster 5), are equal to the centroid values of the most positive cluster (Cluster 5) in k=6. It is observed that the number of organizations in the most positive cluster are the same as in k=6, hence the cluster seems to have stabilized with regard to maturity of six factors.

Table 3.15 Final cluster centres k=8

Factor	1	2	3	4	5	6	7	8
Strategic Alignment with BI&A	-1.23019	-0.34746	1.48442	0.85242	-2.07385	-0.87904	0.25251	-0.14377
Data Management	-0.07910	0.96234	1.35883	0.50167	-1.93340	-1.27049	-0.35674	-0.30900
Enterprise Processes	-1.32153	0.72698	1.51191	0.55094	-2.15736	-1.12133	-0.11679	-0.01790
Organizational Culture	-1.12137	0.57131	1.48848	0.72180	-2.89056	-0.91562	-0.69232	0.03039
People Skills	-1.24224	0.30581	1.73793	0.63010	-2.42041	-0.78686	-0.42572	-0.07673
Technology & Infrastructure	-1.10646	0.31798	1.55020	0.73360	-2.33749	-0.95066	0.01281	-0.23909

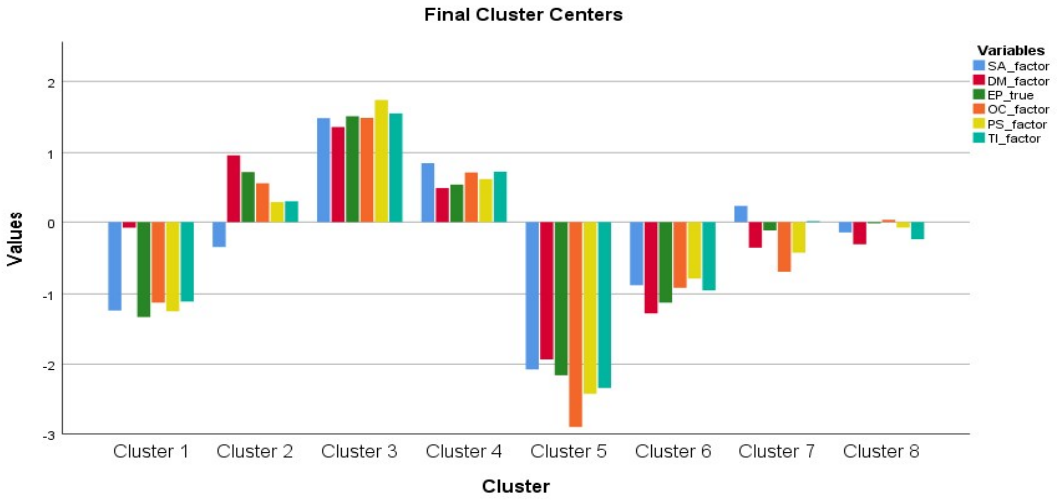

Figure 3.8 Results of k=8

Table 3.15 shows the final cluster centres for k=8. The Figure 3.8 shows the visual representation of the final cluster centres for k=8. It is clearly observed that the cluster with the most positive centroid values is Cluster 3 and the cluster with highest negative centroid values is Cluster 5. Number of organizations in the clusters are as follows: Cluster 1 - 13, Cluster 2 – 15, Cluster 3 – 15, Cluster 4 – 29, Cluster 5 – 3, Cluster 6 – 21, Cluster 7 – 15 and Cluster 8 - 34. Cluster 3 is the most positive cluster and the centroid values of each factor here are marginally higher than the centroid values of the most positive cluster (Cluster 5) in k=7. The number of organizations in the most positive cluster are 15 whereas for k=7, the number of organizations in most positive cluster = 16. Hence the most positive cluster seems to have stabilized with regard to maturity of six factors.

Table 3.16 Final Cluster Centres k=9

Factor	1	2	3	4	5	6	7	8	9
Strategic Alignment with BI&A	-0.80267	-1.38657	-0.29205	-0.77916	0.81259	-2.29159	-1.30877	0.24652	1.48442
Data Management	-0.22238	1.54513	-1.22852	-0.38289	0.66343	-2.59302	-0.73421	0.14452	1.35883
Enterprise Processes	0.13640	1.18186	-0.64875	-1.42998	0.77050	-2.19865	-1.33287	0.08391	1.51191
Organizational Culture	-0.51426	1.14174	-0.48678	-0.79005	0.81610	-3.08097	-1.27213	0.08680	1.48848
People Skills	-0.19747	0.75441	-0.48457	-0.64921	0.65343	-2.49124	-1.45759	0.04615	1.73793
Technology & Infrastructure	-0.16175	0.74891	-0.55513	-0.58765	0.82225	-2.54160	-1.53705	0.05466	1.55020

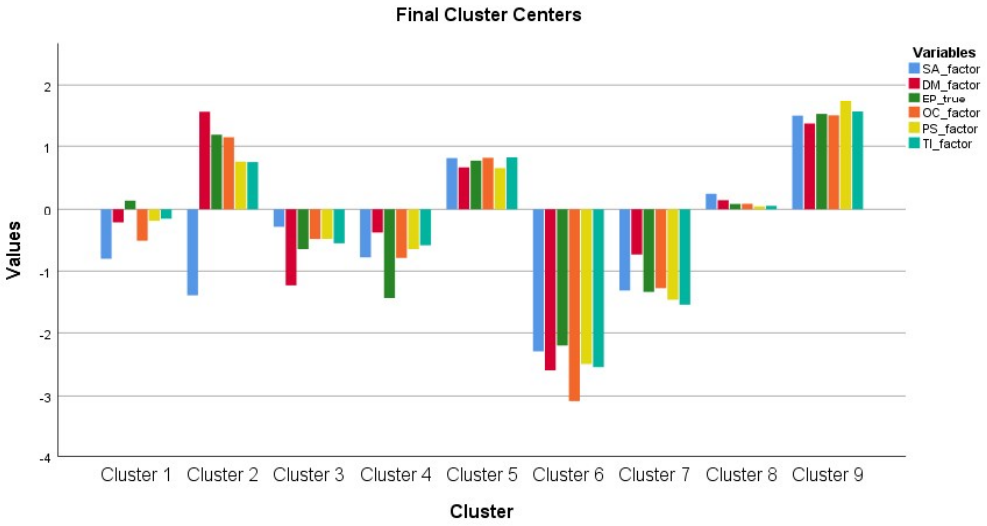

Figure 3.9 Results of k=9

Table 3.16 shows the final cluster centres for k=9. The Figure 3.9 shows the visual representation of the final cluster centres for k=9. It is clearly observed that the cluster with the most positive centroid values is Cluster 9 and the cluster with highest negative centroid values is Cluster 6. Number of organizations in the clusters are as follows: Cluster 1 - 13, Cluster 2 – 3, Cluster 3 – 20, Cluster 4 – 8, Cluster 5 – 26, Cluster 6 – 2, Cluster 7 – 17, Cluster 8 – 41 and Cluster 9 - 15. Cluster 9 is the most positive cluster and the centroid values of each factor here are equal to the centroid values of the most positive cluster (Cluster 3) in k=8. The number of organizations in the most positive cluster are 15 which is equal to the

number of organizations in most positive cluster for k=8. It seems the most positive cluster is stable and strong with regard to maturity of six factors.

Table 3.17 Final Clusters Centres k=10

Factor	1	2	3	4	5	6	7	8	9	10
Strategic Alignment with BI&A	-0.78144	-1.03519	0.78418	-1.06593	-2.29159	0.11395	0.78076	-1.63288	1.65312	-0.35821
Data Management	-0.34078	1.02203	0.81571	-1.27488	-2.59302	-0.00278	0.61633	-0.01893	1.36059	-1.10201
Enterprise Processes	-0.99553	0.93186	0.82858	-1.17602	-2.19865	0.00851	0.64742	-1.65830	1.62156	-0.38063
Organizational Culture	-0.92359	0.69482	0.91022	-1.24557	-3.08097	-0.07625	0.67294	-1.36271	1.62161	-0.18800
People Skills	-0.59779	0.76402	1.08929	-1.14088	-2.49124	-0.02440	0.44769	-1.81259	1.70840	-0.46328
Technology & Infrastructure	-0.42862	0.69110	1.46522	-1.29473	-2.54160	0.01346	0.32132	-1.59745	1.54385	-0.57752

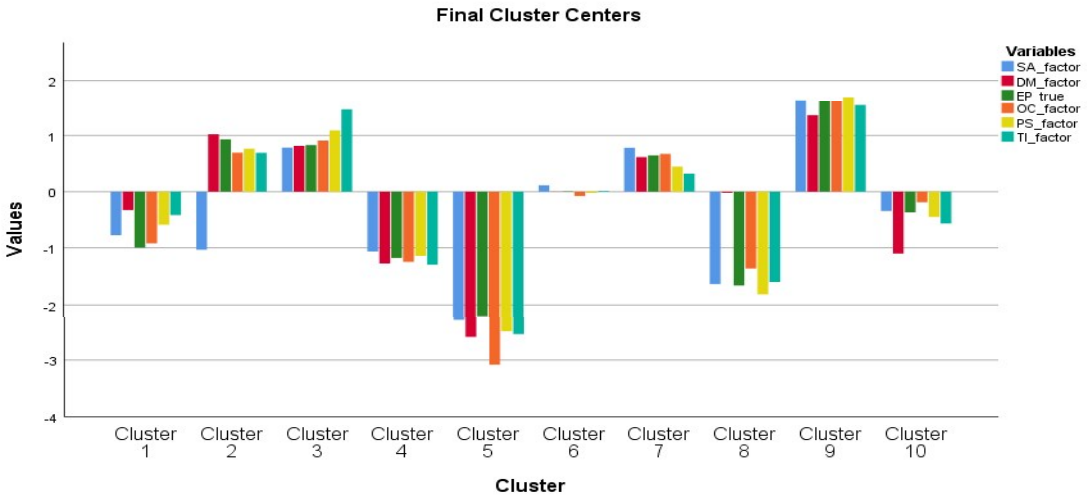

Figure 3.10 Results of k=10

Table 3.17 shows the final cluster centres for k=10. The Figure 3.10 shows the visual representation of the final cluster centres for k=10. It is clearly observed that the cluster with the most positive centroid values is Cluster 9 and the cluster with highest negative centroid values is Cluster 5. Number of organizations in the clusters are as follows: Cluster 1 - 13, Cluster 2 – 6, Cluster 3 – 12, Cluster 4 – 14, Cluster 5 – 2, Cluster 6 – 40, Cluster 7 – 23, Cluster 8 – 6, Cluster 9 – 12 and Cluster 10 - 17. There is no pattern seen here for the centroid values for the most positive cluster in comparison to the clusters for the earlier k.

It was observed that it is not meaningful to have more than ten clusters in a sample size of 145 organizations and hence the clustering technique was concluded at k=10.

Finding optimum value of k

Once the results from k=2 to k=10 were obtained, the next step was to find the optimum number of clusters, that is the best value of 'k'. There are several methods which help in determining the best value of 'k', like silhouette value method as described by (Bowler and Datar, 2018) or the "volume metric" method explained by (Kane, 2012). We can also use the Silhouette Distance method and the Calinski-Harabasz (CH) Index method (Wang and Xu, 2019) to find optimum number of clusters. The Elbow method (Bholowalia and Kumar, 2014) was found to give appropriate results and was hence used in this study to evaluate the optimum cluster number.

The Elbow method looks at the percentage of variance explained as a function of the number of clusters. This method is based on the idea that one should choose that number of clusters k, so that k+1 does not give a better model of the data (Bholowalia and Kumar, 2014). The ratio calculated here was equal to intra-cluster distance divided by inter-cluster distance as seen in Table 3.18. The intra-cluster distance was calculated as average distance between all data points in the cluster whereas the inter-cluster distance was calculated as average distance between clusters. These were calculated for all k = 2 to k= 10 based on results obtained from SPSS.

An 'elbow chart' is a line chart depicting the decline in cluster heterogeneity as we add more clusters (Patel, 2018).

Table 3.18 Ratio used for Elbow method

Clusters	Intra-cluster	Inter-cluster	Ratio=Intra/Inter
2	1.63	3.43	0.47
3	1.38	3.52	0.39
4	1.3	3.63	0.36
5	1.21	3.88	0.31
6	1.16	4.14	0.28
7	1.13	3.83	0.30
8	1.14	3.7	0.31
9	1.06	3.89	0.27
10	1.05	3.92	0.27

Based on the graph in Figure 3.12 where x-axis has number of clusters and y-axis has the ratio, the 'elbow' indicates the optimum number of clusters as illustrated in Figure-3.11. Clearly from the figure, one can see that before k=6 there is a sharp drop in the ratio and hence on the line graph, whereas from k=6 the ratio eases off into a plateau. Therefore, as explained by (Kodinarya and Makwana, 2013), we can conclude that the optimum number of clusters here is k = 6.

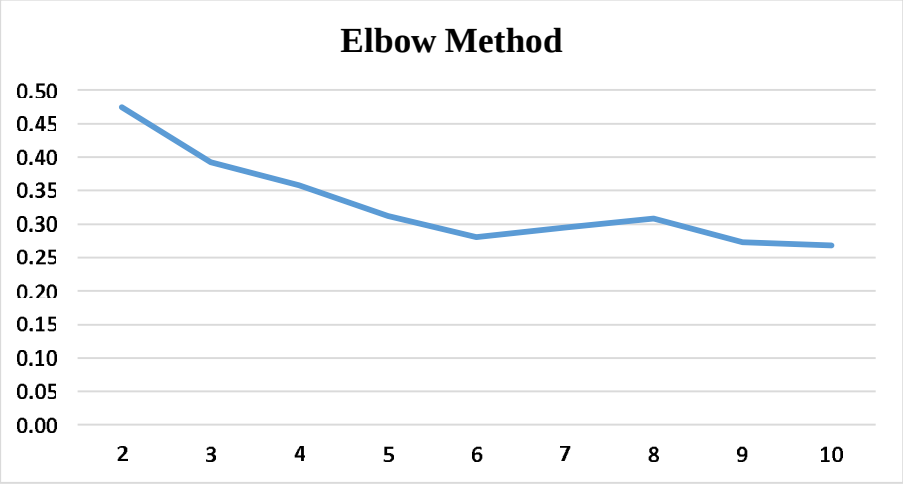

Figure 3.11 Elbow method for identifying optimum cluster number

Referring to the data for six clusters, each organization was a member of any one cluster which was either 1,2,3,4,5, or 6. The six clusters indicated the different levels of maturity for the six factors. Hence the organizations were divided into six non-overlapping groups. There

were well separated clusters. It was found that an organization in a cluster is closer (or more similar) to every other organization in that cluster than to any other organization which is not in the cluster (Rai and Singh, 2010).

Next, each cluster was studied in detail to understand the characteristics based on the six factors. The clusters were then reorganized based on the cluster centroid values for the six factors. There was an interesting pattern found in various characteristics based on the ascending centroid values of factors across the six clusters. This has been discussed in Chapter 4.

3.6.3 Case Method Analysis

The qualitative data collected from the in-depth interviews was analysed using thematic analysis with the help of a tool called NVIVO 12 Plus. The data collection from the in-depth interviews was based on the underlying assumption that findings would emerge from the qualitative data collected.

The qualitative data from the in-depth interviews for each of the case organizations was analysed using an inductive approach with the Thematic analysis method. As mentioned by (Thomas, 2006), inductive approach is a method commonly used to find concepts, themes or any other categories from the data that has been collected, to satisfy the emergent nature of a qualitative study. The thematic analysis method involves searching through a dataset to identify repetitive characteristics emerging out of the data. In this study, the dataset comprised of the interviews taken. A step by step phase wise iterative process similar to the one suggested by (Braun and Clarke, 2006) was followed for thematic analysis with the following phases:

- Familiarization with the data
- Generating line by line coding
- Identifying themes
- Reviewing, defining and naming themes
- Making sense out of the analysis

Familiarization with the data

The interviews were recorded and transcribed. The process of transcription proved to be helpful and created a thorough understanding and familiarization of the data. These were then read repeatedly to identify any patterns emerging from the six case organizations.

Generating line by line coding

Line by line coding was done using the tool NVIVO 12 plus. Each case was read line by line to identify codes. This gave an initial list of ideas about the respondent's views about the state of BI&A in his organization. As each case was being read, the codes were revisited, modified and deleted to arrive at more accurate codes which gave clarity in the data.

Identifying themes

Once all the data was coded, a broader level of themes was identified. Several codes were sorted and merged into a theme. This gave an idea about the relationship between codes, themes and various levels of themes. It was observed that there were different levels for the themes that emerged. For example, for the theme 'Overall alignment with business strategy and goals', some respondents had mentioned 'high' alignment while some had mentioned 'low' or 'yet developing' alignment.

Reviewing, defining and naming themes

Next, the themes were reviewed and finalised. Meaningful and shorter names were given to these emerging themes which identified the essence of what the theme is all about. This

generated a list of themes after the third iteration of coding. The list of these themes is seen in Chapter 5 - Table 5.2.

Making sense out of the analysis

It was observed that there were themes which had the highest number of references and were coded in the highest number of interviews. These themes were repeated again and again in most interviews. They emerged as the key characteristics as shown in Chapter 5 - Table 5.3. The interviews for the six case organizations clearly showed a pattern in the key characteristics identified, based on the maturity of six factors. They validated the results obtained from the quantitative analysis.

From the findings obtained from the quantitative analysis, Cement Industry was observed to be having highest BI&A capability maturity. Hence to validate these findings, an in-depth interview was also conducted with a respondent from a large Cement organization to understand the state of BI&A, usage and effectiveness of BI&A of the organization. This was the seventh case organization.

The results of these interviews have been discussed in Chapter 4.

3.7 CONCLUDING REMARKS

In this chapter we discussed the development of the tool, the sampling procedure and the various research methods used to address the research objectives. The next chapter discusses the analysis, results and discussions thereof.