Gordon Institute of Business Science
University of Pretoria

Investigating the Relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making

Zaheer Dhoodhat

21083615

A research project submitted to the Gordon Institute of Business Science, University of Pretoria, in partial fulfilment of the requirements for the degree of Master of Business Administration.

7 November 2018

Abstract

Big Data Analytics can be a means of extracting value and generating competitive advantage. However, a challenge in achieving this is that key specialised capabilities are required. Data-Driven Decision-Making, which depends on data and analytics insights, has been shown to have a positive impact on firm performance. This study investigates the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making through a questionnaire based, quantitative study which surveyed managers and analytics professionals across several industries. The study found that Big Data Analytics Infrastructure Flexibility, Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise have significant, positive correlations with Data-Driven Decision-Making. Furthermore, this study also found that Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise are significant predictors of Data-Driven Decision-Making. The empirical support of these relationships provides a contribution to literature while also providing practical implications for business looking to create value through Big Data Analytics.

Keywords

Big Data, Big Data Analytics, Big Data Analytics Capabilities, Data-Driven Decision-Making

Declaration

I declare that this research project is my own work. It is submitted in partial fulfilment of the requirements for the degree of Master of Business Administration at the Gordon Institute of Business Science, University of Pretoria. It has not been submitted before for any degree or examination in any other University. I further declare that I have obtained the necessary authorisation and consent to carry out this research.

____

Zaheer Dhoodhat

07 November 2018

Contents

List of Figures

List of Tables

1 Introduction

This chapter provides an introductory overview and background to this study, which deals with the investigation of the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making. This is done through a discussion of the background of the research, an outline of the research problem, the theoretical and business significance of the research as well as the scope and purpose of the research.

1.1 Background to the research problem

In a continuously advancing technological age and increasingly connected world there has been a significant growth in the amount of data being generated and stored. To illustrate this, it has been reported that in 2010, businesses produced and stored approximately seven exabytes of data (which are in the order of billions of gigabytes)(Matthias, Fouweather, Gregory, & Vernon, 2017). Furthermore, there are estimates that this number will grow to 40 exabytes by 2020 (Matthias et al., 2017). This is due to the number of sources generating data, such as sensors, wearable technology, social media and enterprise systems, continuously increasing (George, Osinga, Lavie, & Scott, 2016). The voluminous amounts of data available to organisations provide an opportunity for generating value through its processing and analysis, thereby gaining answers to key questions in the organisation.

The term Big Data (BD) has been coined as having a "moving definition" due to the progression of its definition, which is based on its attributes (Sheng, Amankwah-Amoah, & Wang, 2017). The definition of Big Data has evolved from the 3 Vs of Velocity, Variety and Volume to comprise additional Vs which include the Veracity, Variability and Value of data (Gandomi & Haider, 2015). The term Big Data can, to some extent, be seen as a misnomer since the size (volume) of the data is just one of the attributes which characterise it. A more in-depth discussion of the attributes of Big Data is provided in Section 2.2. The evolution of the definition of Big Data is testament to the continued relevance and importance of the field in societal as well as business contexts and is an active area of research (Sheng et al., 2017).

In order to handle the growing amount of data, technology dealing with the storage, processing, analysis as well as visualisation of Big Data has been developing at a significant rate (Chen, Preston, & Swink, 2015). An example of one technology platform, Apache Hadoop, which offers distributed storage and computing capability provides organisations with the storage as well as computing power to deal with Big Data (Chen et al., 2015; Grover, Chiang, Liang, & Zhang, 2018; Gupta & George, 2016). Grover et al. (2018) highlight that large organisations such as Walmart and Deutsche Bank have invested in this technology with one of Walmart's use cases being analysis of every click on their websites and product recommendations based on purchase patterns. To provide a sense of scale, Walmart was collecting data from one million customers every hour and consolidating data from 10 websites on this platform (Grover et al., 2018).

Analytics refers to the process of generating insights, from data, by making use of statistical, quantitative, predictive and cognitive models (Kiron, Prentice, & Ferguson, 2014b; Vidgen, Shaw, & Grant, 2017). Further to this Big Data Analytics (BDA), which is a field related to business intelligence and data science, is regarded as a discipline of extracting value from the big data through its various attributes or Vs (Velocity, Variety, Volume, Veracity, Variability and Value) (Günther, Rezazade Mehrizi, Huysman, & Feldberg, 2017). This is done through the management, processing and analysis of the data (Côrte-Real, Oliveira, & Ruivo, 2017; Wamba et al., 2017). A key additional part to the definition of analytics is that it is used, amongst other things, to drive decisions (Kiron et al., 2014b; Vidgen et al., 2017). A means of generating value is by ensuring that decisions made within the organisation are made based on facts or evidence which can be gleaned from the data (Dremel, Herterich, Wulf, Waizmann, & Brenner, 2017; Sheng et al., 2017).

Data-Driven Decision-Making (DDDM) relates to the use of data and data-based insights to support decision-making (Brynjolfsson & McElheran, 2016a; Cao, Duan, & Li, 2015). Kiron, Prentice and Ferguson (2014a) highlight that decisions based on experience or intuition can be enhanced through the use of data and analytics. Furthermore, Sharma, Mithas and Kankanhalli (2014) posit that organisational value creation through Big Data Analytics lies in the improved decision-making as a result of being data-driven. This is further supported by Popovič, Hackney, Coelho and Jaklič (2012) who mention that the gathering of data by decision-makers will have limited effect if not used to inform decisions. It has also been shown that Data-Driven Decision-Making has links to superior firm performance (Brynjolfsson, Hitt, & Kim, 2011; Brynjolfsson & McElheran, 2016b). This provides an indication that gaining an understanding of Big Data Analytics and Data-Driven Decision-Making is valuable.

Dremel et al. (2017) provide a case study of how car manufacturer Audi evolved into an organisation leveraging Big Data Analytics and adopting evidence-based or Data-Driven decision-making from an organisation previously using limited analytics (mainly for marketing) and making decisions based on intuition or experience. This process saw Audi transform from having limited application of analytics in the use of marketing data to moving towards providing Analytics-as-a-Service. Furthermore, the transformation saw Audi develop their Big Data Analytics Capabilities to the point where car data is incorporated and advanced analytics methods employed in the design and development of digital services for customers. Additionally, the evolution toward data-driven decision-making has seen decision-makers increasingly rely on data and analytics for daily operational decisions as well as in the optimisation of digital services (Dremel et al., 2017). This case study illustrates the significant positive effect that Big Data Analytics and Data-Driven Decision-Making could have on an organisation and hence suggests that a study investigating Big Data Analytics Capabilities and Data-Driven Decision-Making is warranted.

In order to employ Data-Driven Decision Making, employees within organisations need to have the capabilities or skills of dealing with and extracting useful information from the data (Pigni, Piccoli, & Watson, 2016). One of these capabilities being Big Data Analytics which has been found to be a challenge in business (Sivarajah, Kamal, Irani, & Weerakkody, 2017). Furthermore, there are Big Data Analytics infrastructural requirements that are necessary in order for employees to employ Data-Driven Decision-Making (Pigni et al., 2016; Raguseo, 2018). Additional capabilities that are necessary include business knowledge and management capabilities (Gupta & George, 2016; Pigni et al., 2016; Wamba et al., 2017). Therefore, this study will deal with investigating the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making.

1.2 Research Problem

In order to effectively espouse Data-Driven Decision-Making, Big Data Analytics capabilities are required. As alluded to in Section 1.1, Big Data makes use of advanced, and continuously changing, technology and thus personnel require a specialised set of skills in order to access or extract the information necessary for making decisions (Pigni et al., 2016; Ransbotham, Kiron, & Prentice, 2015b). However, these skills need to be supported by other factors such as Big Data Analytics infrastructure flexibility which ensures that the Big Data Analytics infrastructure has the characteristics that are required to provide an effective Big Data Analytics function (Gupta & George, 2016; Pigni et al., 2016; Wamba et al., 2017). An additional factor or capability that is required, in addition to expertise and infrastructure, is that of Big Data Analytics management capabilities which covers aspects of the ability to manage Big Data Analytics activities as well as governance related aspects (Wamba et al., 2017).

Figure 1 shows a high-level research model that illustrates the (second order) constructs of Big Data Analytics Capabilities which are Big Data Analytics Infrastructure Flexibility, Big Data Analytics Personnel Expertise and Big Data Analytics Management Capabilities each of which further comprise first order dimensions (Wamba et al., 2017). The complete research model showing all the dimensions of the constructs is presented in Chapters 2 and 3. A more detailed description of each of the constructs is also provided in Chapter 2.

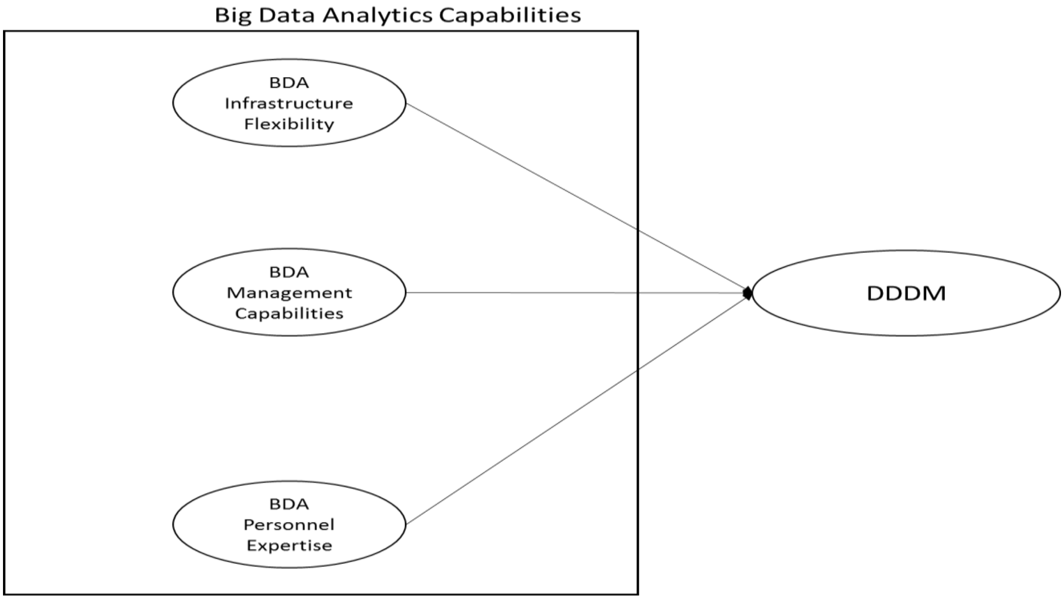


Figure 1: Simplified Research Model: Adapted from Wamba et al. (2017)

The research problem that is to be addressed in this study is to investigate the relationship between Big Data Analytics Capabilities (which encapsulates the constructs shown in Figure 1) and Data-Driven Decision-Making.

1.3 Significance of the Research

This section provides a description of the significance and relevance of this study from the perspective of the business as well as theoretical need.

1.3.1 Business Need

In their study Kiron et al. (2014) found that 87% of managers felt that they needed to increase their usage of analytics. However, Ransbotham, Kiron and Prentice (2015a) assert that organisations have a higher likelihood of generating a competitive advantage by combining analytical skills and business knowledge. This means that in order to create value from analytics, analytical skills should be linked to an outcome. This research aims to explain this outcome in the sense of Data-Driven Decision-Making whereby the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making is investigated. Brynjolfsson, Hitt and Kim (2011) show the link between Data-Driven Decision-Making and superior firm performance. Additionally, as alluded to in Section 1.1, Dremel et al. (2017) describe Audi deriving business value from the transformation of the organisation to using Big Data Analytics and employing Data-Driven Decision-Making. This rouses several questions from a business perspective, one of which being the association between the Big Data Analytics Capabilities and Data-Driven Decision-Making.

In a white paper, published by World Economic Forum in collaboration with Accenture, on the digital transformation amongst enterprises, several key aspects of digital businesses are highlighted for business leaders to take heed of (World Economic Forum, 2016). A number of key factors highlighted in the paper relate to Big Data Analytics and decision-making which underscores the business and management relevance of this study. An excerpt from one of the questions in the paper relating to the operations environment states “How are you empowering employees through digital channels to enable faster decision making…” (World Economic Forum, 2016, p. 4). Further to this, relating to business models, the paper raises the question “Do you emphasise decisions informed by solid analytics?” (World Economic Forum, 2016, p. 4). Other aspects relating to leveraging data and analytics are also included in the paper. This brief description, of the digital aspects that business leaders need to deal with, illustrates the pertinence of this study in the business environment, particularly in this age of digital revolution (World Economic Forum, 2016).

Goes (2014), highlights a very important challenge in business which states that “executives of most corporations and midmarket companies are struggling with understanding and deciding what to do” in the sphere of Big Data (p. iii). In light of the above there is business value in understanding the relationships between the Big Data Analytics Capabilities constructs and Data-Driven Decision-Making in that business leaders will be galvanised with a lens through which to analyse their organisations. Through this insight, business leaders will be empowered to effectively gauge the Big Data Analytics capabilities of employees in their organisations as well as the extent of Data-Driven Decision-Making and take appropriate action. This could be upskilling employees in certain technical skills, improving business knowledge, providing Big Data Analytics infrastructure with the required characteristics and reviewing business processes.

1.3.2 Theoretical Need

George, Haas and Pentland (2014) noted that the exploration of creating value from Big Data has been mainly driven from practice. They further assert that limited “management scholarship” has been published relating to how to confront the challenges or provide new theories in this field (George et al., 2014, p. 321). This is further corroborated by Sheng et al. (2017) who highlight that technological studies dominated while research relating to Big Data and its business consequences fell behind. However, they are of the view that research relating to Big Data and management, with a specific reference to the data-driven concept, is still anticipated. Günther, Rezazade Mehrizi, Huysman and Feldberg (2017) call for additional research into the implications of Big Data use in organisational contexts. The recommendations from these studies indicate that there is a need for theoretical studies relating to the challenges and use of Big Data from a business perspective. They also describe the need for investigating or studying the business consequences (which could include creating value) of Big Data.

To further highlight the theoretical need to this study, Janssen, van der Voort and Wahyudi (2017) assert that research on decision-making through making use of Big Data and Big Data Analytics is limited. Sivarajah et al. (2017), in their study investigating challenges in Big Data and Big Data Analytics through a systematic literature review, highlight that there is a need to understand Big Data Analytics through quantitative or survey based studies since there is a gap in literature in this respect. This directly relates to this study since it aims at gaining and understanding of the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making through a quantitative study. Furthermore, this is reinforced in Wamba, Akter, Edwards, Chopin, and Gnanzou (2015) who mention the need for more research providing explanatory theories to several topics, including those in the Big Data domain relating to the decision-making process. The above discussion demonstrates the need and relevance of this study from an academic and research perspective.

1.4 Scope of Research

This study will be focussed on providing a description of the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making by managers and analytics professionals with decision-making responsibilities. The study will be focussed particularly within technical environments such as information technology, software engineering and analytics across several industries.

1.5 Research Purpose

The purpose of the research is to describe, by means of empirical support, the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making among managers and analytics professionals in technical environments.

1.6 Conclusion

This chapter started by providing background to the research problem. This was then followed by a description of the research problem to be addressed in this study, the related business and theoretical need as well as the scope of this research.

In Chapter 2, a literature review is provided, which is followed by a description of the research proposition and hypotheses in Chapter 3. Chapter 4 covers the research design and methodology which is succeeded by a presentation of the results in Chapter 5. Chapter 6 provides a discussion of the results after which Chapter 7 provides a conclusion to this study.

2 Literature Review

This chapter presents a concise review of theory and literature relevant to this study. As mentioned in Section 1.5, the purpose of this study is to investigate the relationship between Big Data Analytics Capabilities and Data-Drive Decision-Making. Therefore, the theory discussed in this section relates to the decision-making process which provides a basis for understanding of decision-making. This is relevant since it is related to Data-Driven Decision-Making. Further to the theory, literature relating decision-making to management practise and Data-Driven Decision-Making is discussed.

Before discussing at Big Data Analytics Capabilities, it is important to gain some familiarity with the underlying concepts of Big Data and Big Data Analytics as well as the work done in these areas. Therefore, literature relating to Big Data, Big Data Analytics as well as Big Data Analytics Capabilities is presented in this section. Further to this the benefits and constraints of Big Data Analytics and Data-Driven Decision-Making are discussed. This section ends with a discussion of the research gap which this study addresses, based on literature.

2.1 Introduction to Decision-Making

Decision-making has been referred to as a set of mental or cognitive processes that the decision-maker goes through when identifying, selecting and making a choice from possible alternatives (Intezari & Pauleen, 2017). Further to this, decision theory is the study of the underlying reasoning behind choices and is focussed on issues relating to decision-making methodologies (E. Borgonovo, Cappelli, Maccheroni, & Marinacci, 2018).

In outlining the decision-making process Adair (2013) proposes four steps, which excludes the step of the implementation of the choice. The proposed generic steps are defining the objective, collecting relevant information, generating feasible options and finally making the decision (Adair, 2013). Other variants of this process outline additional specific steps such as analysing the consequences of generated options (E. Borgonovo et al., 2018). However, this can be viewed as being part of the step in generating feasible options proposed by Adair (2013).

This decision-making process, which is based on classical decision theory, makes fundamental assumptions which have been challenged in work by Simon (1982) (as cited in Winkler, Kuklinski and Moser (2015)). The first of these assumptions is that the decision-maker has access to complete information. This is not always the case and thus affects the decision-making process. Secondly, in generating the options or alternatives, there is a limit to this since not all alternatives are known and thus an incomplete view of options results. Thirdly, in making a choice, there is an assumption that the decision-maker is completely rational which was challenged by the bounded-rationality argument relating to decision-making (Simon, 1979, 1982). This extends the previous assumptions by adding that objectives are not always well-defined and that further unknowns relating to probabilities of events taking place in the future add to the complexity of decision-making (Frisk & Bannister, 2017; Winkler et al., 2015).

Building on classical decision theory, there has been a significant amount of research done on the effects and influences of psychological and cognitive processes (e.g. biases, ambiguity) as well as interventions (e.g. skills development, training) on decision-making (Ashby, 2017; Emanuele Borgonovo & Marinacci, 2015; Borrero & Henao, 2017; Donovan, Güss, & Naslund, 2015). This highlights the complexity of the decision-making process and some of the many variables that could influence the decisions made by people. A few examples of these studies which reveal some of the influences on decision-making will be presented in the paragraph to follow.

An example of a study on a factor influencing decision-making was done by Donovan, Güss and Naslund (2015), who investigated the effect of self-reflection on dynamic decision-making. The study showed that there was an improvement in the decision-making after undergoing self-reflection training and undertaking self-reflection exercises. A further example is a study by Borrero and Henao (2017) which investigated how cognitive biases, such as confirmation bias, affect decision-making. This study showed results consistent with the bounded-rationality argument alluded to previously in that individuals, depending on their individual preferences and rationality, could be influenced by their biases when making decisions. However, it should be noted that this was done in a student setting, which has its limitations, and future studies in a management environment have been suggested. In another study, Ashby (2017) indicates that an important trait in predicting decision-making skill is numeracy. In addition to this Ashby (2017) has shown that numeric ability increases the amount of information that is sought by individuals in making decisions and that these individuals are more likely to make more consistent decisions. These studies bring to the fore that although the decision-making process can be summarised into a few steps, decision-making is not a static concept and there are many factors that influence the final decision.

Frisk, Lindgren and Mathiassen (2014) propose that decision-making should be a creative process with adaptiveness whereby evidence (or support), from varied sources, is collected and interpreted in a recursive manner such that several options can be investigated and evaluated. It is noted here that support for decision-making may possibly be gained from data which could be structured and unstructured. This ties in with the second step of the decision-making process alluded to earlier, where relevant information is collected.

2.2 Decision-Making in Organisations

In organisations, there are different types of decisions and different levels at which decisions are made (Ransbotham, Kiron, & Prentice, 2016; Shivakumar, 2014). Decision-making in the business context can be considered one of the key duties of employees with a certain level of responsibility within an organisation. It has been noted that there is an undeniable link between management and decision-making. This link is so evident that decision-making can be considered central to what managers do, even to the extent that they are considered synonymous (Intezari & Pauleen, 2017). The level of decisions and the consequences thereof vary with the level of the decision-maker. The consequences of key strategic decisions can have significant impact on businesses' profitability or even viability (Azar, 2014).

Shivakumar (2014) provides a framework of the types of decisions made in businesses which classifies decisions into strategic, neo-strategic, tactical and operational. The framework classifies decisions based on the level of their influence on two dimensions, which are commitment and scope. This is shown in Figure 2 with commitment on the horizontal and scope on the vertical axis. An overview of the two dimensions used for the categorisation is that, commitment refers to the degree to which a firm is required to commit resources such as finances, amongst others, and scope refers to the degree to which a firm is required to change the scope of its offerings and activities. Shivakumar (2014) asserts that strategic decisions attempt to address unprecedented problems which have indeterminate outcomes and serious consequences while at the other end of the spectrum, in this framework, are operational decisions which aim to address well formulated and structured problems that have established methods of being addressed. Strategic decisions, according to Shivakumar (2014), significantly changes the company's commitment and scope, as shown in Figure 2. In the bottom left quadrant of

Figure 2, with the attributes of significant commitment changes and insignificant scope changes are, what the author terms, tactical decisions. These types of decisions are characterised by problems that are clearly understood while the solutions are often unknown. The paper goes into a significant amount of detail into the categorisation of decisions with various illustrative examples.

Figure 2: Categorising Decisions (Shivakumar, 2014, p. 87)

In this study, any type of decision ranging from operational to strategic qualifies for consideration since the area of interest is Data-Driven Decision-Making while not focussing on the type of decision.

The subsections to follow discuss Big Data and Big Data Analytics, which, in the context of decision-making, could be viewed as inputs into the second and third steps of the decision-making process which relate to collecting information and generating alternatives or options.

2.3 Big Data

Braun, Kuljanin and DeShon (2018) posit that finding a specific definition of Big Data (BD) is “highly elusive” (p. 635) . Wamba, Akter, Edwards, Chopin and Gnanzou (2015) provide a sample list of definitions of Big Data mentioned in literature. This bears testament to the statement by Braun et al. (2018) that finding a specific definition of Big Data is a challenge. These definitions generally encapsulate the nature of Big Data as being outside the traditional means of storage and analysis (Wamba et al., 2015). Big

Data definitions found in literature are regularly being reviewed and revised, as alluded to in Chapter 1 (Sheng et al., 2017).

In addition to the definitions provided by Wamba et al. (2015), Big Data is regularly defined in terms of data attributes which are described by words that begin with the letter V. The three Vs (3Vs), which lay the foundation for further definitions of Big Data, are volume, velocity and variety (Erevelles, Fukawa, & Swayne, 2016; George et al., 2016; Sheng et al., 2017). Volume refers to the quantity of the data which can be viewed as the amount of storage required or the amount of records contained in the data (Wamba et al., 2015). Apart from other data, Walmart has been reported to be generating 2.5 petabytes of data relating to consumers (Erevelles et al., 2016; Sivarajah et al., 2017). To provide some context of the size of this data, 1 petabyte is equivalent to 1000000 gigabytes or 1000 terabytes. Velocity refers to the rate at which data is generated, which varies according to data sources from slower sources to real-time streaming data (Erevelles et al., 2016; Sheng et al., 2017). Variety refers to the various types of data which includes structured and unstructured data that is generated from various sources (Wamba et al., 2015). Data sources include enterprise systems, social media, web data and Internet of Things (IoT) devices (Janssen et al., 2017; Sivarajah et al., 2017). The 3Vs definition of Big Data has been mentioned to serve as a differentiator between Big Data and large sets of data (Erevelles et al., 2016).

Goes (2014) makes mention of the 4Vs definition of Big Data which introduces the attribute of veracity of data. Veracity in the definition of Big Data refers to the quality of data in terms of completeness, consistency and accuracy (Janssen et al., 2017; Sivarajah et al., 2017). Further to this, a 5Vs definition of Big Data which incorporates the value attribute of data has been used in literature (Sheng et al., 2017; Wamba et al., 2015). Value refers to the usefulness and relevance of the data that is collected and stored such that benefit can be extracted from the data (Erevelles et al., 2016; Sheng et al., 2017; Wamba et al., 2015). There have been further additions to the definition via further attributes (Vs) such as variability of the meaning of the data as well as visualisation which relates to presenting Big Data in a comprehensible manner (Gandomi & Haider, 2015; Sivarajah et al., 2017).

Braun, Kuljanin and DeShon (2018) offer an argument which challenges the V attributes of Big Data such as volume and variety since they argue that the context has a significant impact. They offer an alternate definition of Big Data which states that Big Data comprises data sets in which manual assessment is considered impracticable. However, the author considers the 5Vs definition, adopted by Wamba et al. (2015), as a sufficient conceptualisation of the Big Data construct for this study. Additionally, Erevelles et al. (2016) state that the 5Vs are essential for extracting insights from Big Data.

Big Data is an active and growing field of research and application (Günther et al., 2017; Sheng et al., 2017). By its definition, Big Data comes with large volumes of data at high velocity in a variety of types from several sources. These sources could be within the organisation or external and include structured as well as unstructured data. Thus, Big Data provides businesses with a rich source from which to draw useful information. This is also reflected in the value attribute of Big Data previously alluded to. Progressing from the captured raw data to decision-making requires following a process of several steps. Various processes defining these steps can be found in literature, but they have the same fundamental building blocks (Janssen et al., 2017). An example of steps in the data progression process, listed by Janssen et al. (2017) are problem definition, data searching, data entity resolution and answering the query or solving the problem (decision-making). The problem definition and data searching steps of the process are self-explanatory. The data entity resolution step in the process includes aspects of pre-processing and transformation of the data through creating links between related data entities (Ayat, Akbarinia, Afsarmanesh, & Valduriez, 2014).

When contrasting the data progression process steps to those of the decision-making process listed discussed in Section 2.1, one is able to clearly see common elements. An example being the first steps of defining the objective in decision-making and problem definition in the data progression process or the information collection and data searching steps. This provides an indication that the decision-making process and the process of resolving a query or problem resolution from data (which could be Big Data) are aligned (Janssen et al., 2017; Sheng et al., 2017).

2.4 Big Data Analytics

Due to the complex characteristics and attributes of Big Data, mentioned previously, dealing with Big Data is not trivial (Goes, 2014; Janssen et al., 2017; Sivarajah et al., 2017). As such, specific skills, capabilities and technologies are required in order to manage Big Data and furthermore derive value from Big Data in progressing along the steps in the process to making a decision from raw-data (Ransbotham et al., 2015b; Sivarajah et al., 2017). In their study, Ransbotham et al. (2015b) highlight the lack of analytics skills and ability to deal with Big Data as a key challenge to generating value from Big Data.

Analytics is described, by Goes (2014), as services that generate knowledge and intelligence to support decision-making through the use of complex techniques and vast data sources. The definition of Big Data Analytics builds upon this definition by aligning analytics with Big Data. The Big Data Analytics definition states that Big Data Analytics is a means of realising value from Big Data through the application of various analytical techniques and processes (Günther et al., 2017). Further to this, Côrte-Real, Oliveira and Ruivo (2017) supplement the aforementioned definition of Big Data Analytics by adding the concept of new technologies and architectures enabling discovery and analysis to extract value in an economical manner.

A fundamental point of Big Data Analytics is that it is a means of supporting and driving decision-making (Chen et al., 2015; Liberatore, Pollack-Johnson, & Clain, 2017; Ransbotham et al., 2016). Grover, Chiang, Liang and Zhang (2018) noted that this (support of decision-making by Big Data Analytics) can be achieved through accessibility to data and analytics models that enhance human decision-making. They further allude to supporting models being integrated into business processes. It is also noted that Big Data Analytics can also be a means of driving actions (Bumblauskas, Nold, Bumblauskas, & Igou, 2017; Vidgen et al., 2017). However, it is highlighted by Ransbotham, Kiron and Prentice (2015a) that translating from analytics to actions is still found to be a challenge amongst managers.

2.5 Big Data Analytics Capabilities

In order to gain value from Big Data and Big Data Analytics, described in Sections 2.2 and 2.4 respectively, there are capabilities that are required of organisations and their employees (Akter, Wamba, Gunasekaran, Dubey, & Childe, 2016; Kiron et al., 2014b; Kowalczyk & Buxmann, 2015; Wamba et al., 2017). Big Data Analytics Capabilities is based on the foundations of resource-based theory (RBT) or resource-based view (RBV) of organisations which is derived from theory in strategic management (Akter et al., 2016; Wamba et al., 2017). Resource-based view of organisations provides the view that through establishing key capabilities and making use of key physical resources, organisations can generate a competitive advantage (Gunasekaran et al., 2017; Kwon, Lee, & Shin, 2014). Akter et al. (2016) define Big Data Analytics Capabilities as “the competence to provide business insights using data management, infrastructure (technology) and talent (personnel) capability to transform business into a competitive force” (p. 114). This definition with the dimensions of data management, technology and talent or skills is also reflected in the study by Kiron et al. (2014b). Although Kiron et al. (2014b) additionally emphasise the analytics culture as a key capability, they also highlight that the culture is built upon the data management, infrastructure and talent and thus would not exist without those capabilities.

Big Data Analytics Capabilities has been modelled by a third order model as shown in Figure 3 (Akter et al., 2016; Wamba et al., 2017). In this model of Big Data Analytics Capabilities there are eleven first order constructs (BDA technical knowledge, technological management, business knowledge, relational knowledge, coordination, control, planning, investment, connectivity, compatibility, modularity) and three second order constructs (BDA Personnel Expertise, Management Capabilities, Infrastructure Flexibility) leading to the third order construct of Big Data Analytics Capabilities.

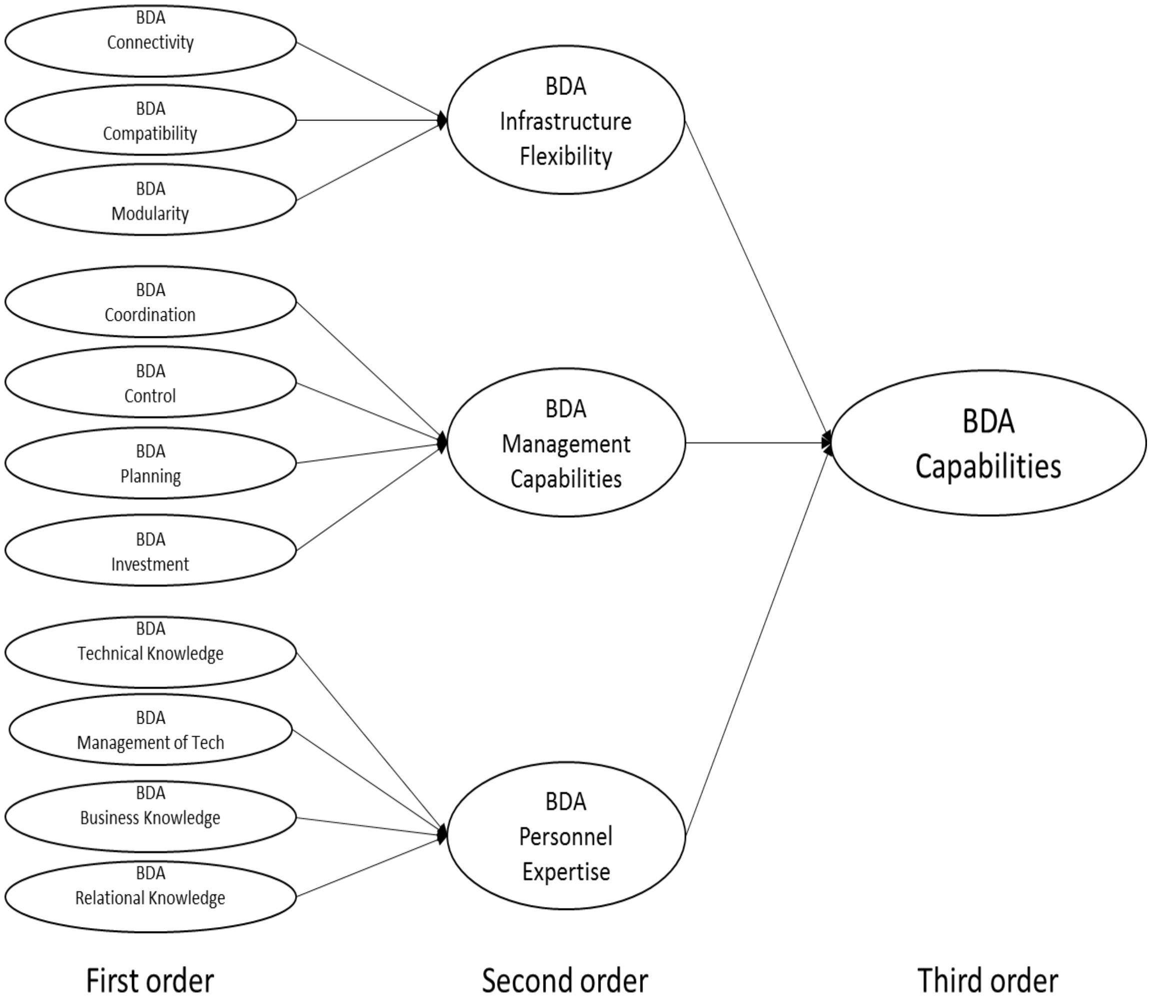


Figure 3: Big Data Analytics Capabilities Model (Wamba et al., 2017, p. 358)

The Big Data Analytics Infrastructure Flexibility, second order construct, relates to the connectivity, compatibility and modularity characteristics of the Big Data Analytics infrastructure (Wamba et al., 2017). This is related to the nature of the infrastructure in enabling users, such as analytics professionals, to rapidly connect to data sources and effectively develop and deploy analytics solutions or output (Akter et al., 2016). Related to this construct Kache and Seuring (2017) as well as Liberatore et al. (2017), note that inadequate infrastructure is a considerable hindrance to adoption of Big Data Analytics. The connectivity and compatibility dimensions (first order constructs) measure the accessibility of data and the sharing of the derived analytics insights and the compatibility of the Big Data Analytics Infrastructure Flexibility in interfacing with other systems, respectively. Liberatore et al. (2017) make use of categories of a similar nature relating to infrastructure which include access and flexibility while Ramanathan, Philpott, Duan and Cao (2017) allude to the incompatibility of infrastructure being an obstruction to the adoption of business analytics. The modularity dimension of Big Data Analytics Infrastructure Flexibility provides insight into the ability or ease of modifying and reusing components of the Big Data Analytics platform or infrastructure (Akter et al., 2016).

Big Data Analytics Management Capability, refers to the capability of personnel in managing the Big Data Analytics resources from the technology perspective which comprises coordination, investment, control and planning (Akter et al., 2016; Wamba et al., 2017). The aspect or first order construct of coordination deals with the ability to coordinate Big Data Analytics activities across departments or other structures in organisations. The control construct explores the management of resources and initiatives as well as regulating and improving business processes relating to Big Data Analytics. Liberatore et al. (2017) complements this dimension with the perspective of the governance around Big Data Analytics. The first order construct of Big Data Analytics planning capabilities refers to aspects of Big Data Analytics such as the design of Big Data Analytics for adaptability in an evolving industry and for planning or strategising for effective utilisation of Big Data Analytics (Wamba et al., 2017).

The final second order construct in Figure 3, Big Data Analytics Personnel Expertise, refers to the competence and knowledge of personnel in a variety of aspects comprising technical, technology management, business and relational. Technical knowledge or expertise (first order construct) explore aspects such as programming skills, data management abilities, predictive modelling knowledge and distributed computing capabilities (Cao et al., 2015; Wamba et al., 2017). Business understanding refers to the knowledge of personnel regarding business policies, business functions and understanding business problems and developing solutions (Wamba et al., 2017). Relating to technical and business knowledge in Big Data Analytics, Big Data Analytics skills as well as business understanding have been cited in Ransbotham et al. (2015b) as a pathway to gaining a competitive edge for organisations. The term "Relational capabilities" refers to personnel ability to navigate and manage relationships in the business environment which include aspects such as managing client relationships and working in collaborative environments and is explored through the relational knowledge, first order, construct (Wamba et al., 2017). Kiron et al. (2014b) assert that the development of capabilities in the management of data was not keeping up with the need and thus affecting the use of data in decision-making or data-driven decision-making. This shows the importance of the Big Data Analytics management capability and its relationship to Data-Driven Decision-Making.

Further to this there are other aspects of Big Data Analytics infrastructure such as system quality (not included in the provided model of Big Data Analytics Capabilities) which entails, inter alia, integration and adaptability of the Big Data Infrastructure (Ji-fan Ren, Fosso Wamba, Akter, Dubey, & Childe, 2017). Ji-fan Ren et al. (2017) have tested and shown the positive influence of Big Data Analytics infrastructure in terms of Big Data Analytics system quality on the business value gained from Big Data Analytics. Rejikumar et al. (2018) noted adequate Big Data Analytics technology and access to the infrastructure as influencing factors in the intention to adopt Data-Driven Decision-Making. These findings emphasise the importance of Big Data Analytics infrastructure.

Janssen et al. (2017) highlight that Big Data Analytics Capabilities contribute to influencing Big Data decision quality. Further to the mentioned capabilities, there are other organisational factors which influence Big Data Analytics as well as Data-Driven Decision-Making which include, inter alia, organisational culture, system quality, information quality, management support (Ji-fan Ren et al., 2017; Kiron et al., 2014b; Ransbotham et al., 2016). In light of this discussion on Big Data Analytics Capabilities, the following section will build on this and provide a discussion on Data-Driven Decision-Making.

2.6 (Big) Data-Driven Decision-Making

Data-Driven Decision-Making (DDDM) refers to the availability and use of data to support the decision-making process (Brynjolfsson & McElheran, 2016b; Rejikumar et al., 2018). Kowalczyk and Buxmann (2015) highlight the importance of supporting and enabling data-centric decisions through business intelligence and analytics, which is a related field of Big Data Analytics (Côrte-Real et al., 2017). Ransbotham, Kiron and Prentice (2016) outline that data analytics is used for various types of decisions in business from operational to strategic. Their research found that the type of decisions influenced by the application of analytics changed with the analytical maturity of the organisation. Thus, analytically advanced organisations used analytics in strategic decisions while analytically challenged or immature organisation focused mostly on cost reduction in the application on analytics.

Data-Driven Decision-Making has received research attention in the fields of education as well as healthcare. Consistent with the construct of analytics expertise mentioned in Section 2.5, Viera and Freer (2015) reported that, the lack of, skills and training were a barrier to Data-Driven Decision-Making in the United States education system. In their study of the factors influencing decision-making quality using Big Data, Janssen et al. (2017) identified several influencing factors which included Big Data quality and the "quality" of the decision-maker. Decision-maker quality in this context refers to the characteristics and interactions of the decision-makers in terms of understanding the data, insights and the implications of the decisions. Congruent with this, Shah, Horne and Capella (2012) contest that having good data alone does not provide any guarantee of good decisions. In addition to the factors mentioned by Janssen et al. (2017), they add that employees with a set of specific characteristics are best equipped for Data-Driven Decision-Making. These characteristics include the ability to balance judgment and analysis, analytic skills competency and willingness to question opinions.

Rejikumar et al. (2018) have investigated the perceptions of managers relating to the intention of Data-Driven Decision-Making adoption. Their study was based on the technology acceptance model and was not approached from a Big Data Analytics Capabilities perspective as is the case in this study. However, although the study was based on managers' perception, this study highlighted that infrastructure (one of the second order constructs in Figure 3) is an influencing factor in managers' intention to adopt Data-Driven Decision-Making.

An important point about Data-Driven Decision-Making in a business and management context is that it is not prescribed that Big Data Analytics and sophisticated algorithms take the control away from the decision-makers by automatically making decisions. However, it is viewed that the decision-maker take advantage of the available data and analytics to support, inform and influence the decision (Goes, 2014; Horita, de Albuquerque, Marchezini, & Mendiondo, 2017; Kowalczyk & Buxmann, 2015).

2.7 Benefits of Big Data Analytics and Data-Driven Decision-Making

Big Data and Big Data Analytics can contribute value to organisations through the application of Data-Driven Decision-Making (Günther et al., 2017; Sheng et al., 2017; Weill & Woerner, 2015). Ramanathan, Philpott, Duan and Cao (2017) have found the benefit of business analytics adoption in impacting the performance of the business. In this study, conducted in the retail sector in the U.K., it was found that businesses could readily identify the benefits accrued in specific functions of operations such as supply chain and marketing. An example of the benefit cited in the marketing context was understanding the effectiveness of marketing campaigns and understanding customer product preferences. From their oft-cited study amongst U.S. manufacturing companies, Brynjolfsson and McElheran (2016a) report that the use of Data-Driven Decision-Making is associated with increased productivity as well as better performance. Further to this, organisations gaining competitive advantage through data analytics have been mentioned to be achieved through improved customer relationships, optimised operations and organisational agility (Côrte-Real et al., 2017; Sheng et al., 2017).

In addition to this, Kiron, Prentice and Ferguson (2014a) assert that decision-making that relied on experience can be enhanced through the use of analytics. This assertion was premised on their finding of an increasing interest and inclination in analytics from decision-makers. They further posit that an additional consequence to the adoption of analytics and Data-Driven Decision-Making is the behavioural shift in organisations. This behavioural shift is in terms of increased innovation using analytics and more collaboration amongst business partners with the view to achieving strategic objectives. Thus, analytics has become an instrument of sorts for collaboration, with partners and stakeholders, by organisations who employ analytics to increase innovation. Furthermore, the way in which organisations, which adopt analytics, generate their strategies are influenced by the insights gleaned from the analytics. This can be related to the earlier discussion on the types of decisions (as proposed by Shivakumar (2014) that can be made in organisations which was discussed in Section 2.1. Kiron et al. (2014a) cite examples from various industries such as media, brewery and health that used analytics to influence their strategies, albeit in their initial phases. This shows that the benefits of applying analytics and Data-Driven Decision-Making are not one dimensional and could have positive knock-on effects in various aspects of businesses.

Weill and Woerner (2015) illustrate the benefits through the application of Data-Driven Decision-Making in practise by Procter and Gamble (P&G). In an effort to keep and grow their power in the market as well as address other business challenges P&G have, amongst other initiatives, introduced support environments for real-time decision-making globally. Managers and executives use the real-time data which is displayed in what is described as "planetarium-like" rooms. These rooms provide an environment where decision makers have access to the necessary data and analysis outputs in order to evaluate, deliberate and make decisions. Further to this, the impact of the decisions on the business are tracked and decision makers are able to decide whether they want to change direction. This level of control and insights on businesses allows decision makers to gain deep understanding of the dynamics of the business which could provide a competitive advantage. The advantage over competitors can manifest through superior and quicker decisions by making use of data and analytics (Kiron et al., 2014a). Additionally, Weill and Woerner (2015) have highlighted that companies need to emphasise Data-Driven Decision-Making (they refer to evidence-based decision-making). They stress that a culture of using evidence as the basis for decisions be fostered and encouraged as opposed to instinct or gut-feel.

In a small business setting, Arunachalam and Kumar (2018) suggest that even Small and Medium Enterprise (SMEs) which are assumed to be data-poor have generated business value through Data-Driven Decision-Making. Their study focussed on using data analytics methods, which included Self-Organising Maps, for gaining insights into profitable consumer segments for use in Data-Driven Decision-Making in business practise. This shows that the benefit of the application of Data-Driven Decision-Making can be accrued by organisations of various sizes.

The examples provided above illustrate that the benefits of Big Data Analytics and Data-Driven Decision-Making include a positive impact in operational departments, enhancing intuition based decisions, track the effects of decisions and alter course if required as well as to organisations. This provides some level of confidence that this study into the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making will offer business value.

2.8 Constraints to Big Data Analytics and Data-Driven Decision-Making

Although there are numerous benefits to Big Data Analytics and Data-Driven Decision-Making, as discussed, there are also constraints. Ransbotham, Kiron and Prentice (2015a) have noted that just producing of analytics is not enough, analytics needs to be consumed for value to be generated. Shah, Horne and Capella (2012) argue that unless employees are able to utilise the data (and insights from analytics) then investing in analytics could be futile. This is an important point since it should be noted that Big Data Analytics is a means to an end and without managers and other decision makers using the information and insights generated through Big Data Analytics limited value will be generated.

Further to this Martin and Golsby-Smith (2017) challenge the use of a scientific approach, such as Data-Driven Decision-Making, in a management context. They suggest that using a scientific approach to decision-making could be an obstacle, particularly for innovation and strategic decision-making. However, in their proposal to how decisions should be made Martin and Golsby-Smith (2017) distinguish between two groups of decision types. For the one group of decisions, which is when there are obstacles or barriers (such as physics constraints) that limit the amount of creativity that can be employed, they promote the use of Big Data Analytics and the scientific approach. In the second decision-making scenario, which is what they refer to as "can" situations (these are situations when there are less physical limitations), the authors suggest a more imaginative and experimental approach with the science and analytics being used as tools for evidence. A parallel can be drawn between this suggestion and the discussion in Section 2.7 where Kiron et al. (2014a) point out that the use of analytics be used to enhance decision-making that previously relied on experience or intuition. Erevelles et al. (2016) provide a related proposition which relates to uncovering hidden insights from Big Data from an inductive approach.

The study by Ashby (2017) notes that although being an important skill, there is a limit to the benefit that numerical skills provide in that searching for information comes at the price of time spent. If this time was taken as a cost, it could be argued that the numerical skills led to deteriorated performance since not having the numeric skills and hence not searching for additional information would not have incurred the “time cost”. However, the contrary view is that the benefit of gaining the additional information, despite the time spent, outweighed the cost.

In their study within the Nordic banking environment Persson and Ryals (2014) found that decisions relating to customer relationships were largely done using heuristics from management rather than on analytics. This is an important insight which highlights that even in the banking sector which is a data-rich environment with customer relationship management (CRM) systems in place which has the required customer relationship data, decisions were based on heuristics. This phenomenon could be related to the insights provided by Ransbotham, Kiron, and Prentice (2016) which suggests that managers in less analytically advanced organisations give less weight to analytics than intuition.

2.9 Research Gap

In order to identify the research gap of investigating the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making, in terms of literature, the following process was followed. The constructs of Big Data Analytics Capabilities and Data-Driven Decision-Making were identified from previous studies in literature. From this the links between the constructs were analysed and the research gap identified. This section provides a description of the process in identifying the research gap and arriving at the research model.

In order to draw the link between Big Data Analytics and Data-Driven Decision-Making, Vidgen et al. (2017) mention that business analytics is concerned with “the extensive use of data, statistical and quantitative analysis, predictive models and fact-based management to drive decisions and actions” (p. 1). This provides a link between business analytics use and decision-making which is also endorsed by studies such as Kiron et al. (2014a) which argue that Big Data Analytics can enhance decision-making. Building on this, the research model provided in Cao, Duan and Li (2015) shows an indirect link between business analytics and Data-Driven Decision-Making. Additionally, Matthias et al. (2017) found that the “collection, processing and utilisation of data to inform decision-making” (p. 50) is a key factor while Wamba et al. (2015) assert that an important aspect in gaining return on investment from Big Data is that employees (at all levels) make use of data in their decision-making process. This provides support that there is link between Big Data Analytics and Data-Driven Decision-Making.

Chen et al. (2015) have undertaken a study between Big Data Analytics and value creation in a supply chain environment while Wamba et al. (2017) have studied and shown the relationship between Big Data Analytics Business Capabilities and firm performance. These studies provide a relationship between Big Data Analytics and firm performance which takes into account the Big Data Analytics Capabilities. Furthermore, Brynjolfsson and McElheran (2016a, 2016b) have illustrated the link between Data-Driven Decision-Making adoption and firm performance.

From the above discussion, we have shown that literature has illustrated a link between Big Data Analytics Capabilities and firm performance and further to this a positive relationship between Data-Driven Decision-Making and firm performance has been shown. However, the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making has not been investigated, to the best of the author's knowledge. Therefore, the contribution of this study is the investigation of the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making. This contribution will facilitate and enable further research in this field through providing insight into an untested relationship.

To support the theoretical need for the proposed study, Sheng et al. (2017) suggest that research into the use of Big Data (structured and unstructured) in informing decisions within organisations be undertaken. Additionally, Sharma, Mithas and Kankanhalli (2014) propose that improved firm performance is achieved as a result of superior decision-making through the application of analytics and further state that richer analysis experimentation in this area is required. Furthermore, Janssen et al. (2017) states that research into the use of Big Data for decision-making is limited notwithstanding its significance. In addition to this, Sivarajah et al. (2017) suggest that there is a need to undertake quantitative studies using surveys to enhance the empirical body of knowledge in the Big Data and Big Data Analytics sphere. Although out of the scope of the proposed study, Intezari and Pauleen (2017) emphasise the relevance of Data-Driven Decision-Making by noting that there is a need for research in Data-Driven Decision-Making, while also considering the incorporation of other theories such as Wise Decision-Making.

3 Research Hypotheses and Proposition

This chapter outlines and provides the rationale for the research hypotheses and proposition investigated in this study, to meet the objective of investigating the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making. These research hypotheses and proposition are based on the literature discussed in Chapter 2 which provides an indication of the applicability and substantive nature of this study.

3.1 Research Overview

The third order model of Big Data Analytics Capabilities employed by Akter et al. (2016) and Wamba et al. (2017b), described in Section 2.5, forms the basis of the research model employed in this study. The third order construct, Big Data Analytics Capabilities, and its antecedent first and second order constructs forms the independent variable. The dependent variable investigated in this study is Data-Driven Decision-Making. This is illustrated in Figure 4. In this study, three research hypotheses which relate the second order constructs of Big Data Analytics Capabilities to the dependent variable, Data-Driven Decision-Making, are tested. Secondly, a research proposition relating Big Data Analytics Capabilities as a predictor of Data-Driven Decision-Making is investigated.

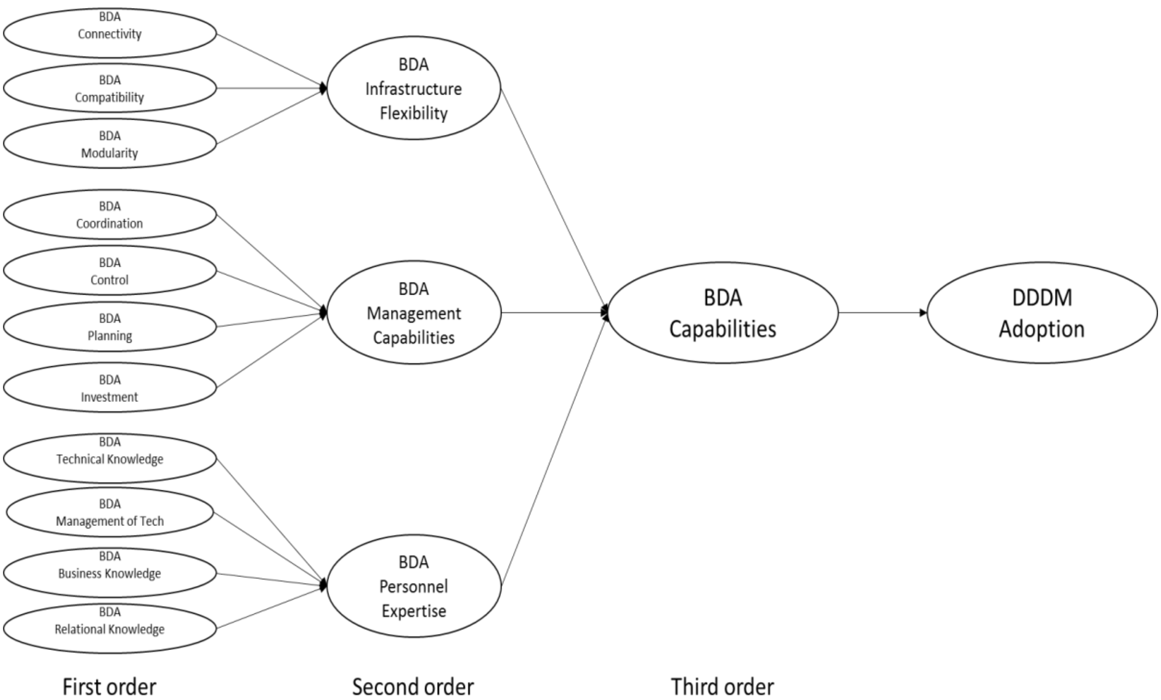


Figure 4: Research Model

3.2 Research Hypotheses

As discussed in Section 2.5, as well as shown in Figure 4, the second order constructs comprising Big Data Analytics Capabilities are Big Data Analytics Expertise, Big Data Analytics Management Capabilities and Big Data Analytics Infrastructure Flexibility. The research hypotheses are based on these second order constructs in the research model relating to Data-Driven Decision-Making.

Characteristics of Big Data Analytics Infrastructure, such as modularity and flexibility, are key in influencing decisions and decision quality (Janssen et al., 2017). Literature also highlights that Big Data Analytics Infrastructure is a challenge and influences Data-Driven Decision-Making through providing the ability to process and provide information to decision makers (Kache & Seuring, 2017). Therefore, the first hypothesis to be tested is:

H1: Big Data Analytics Infrastructure Flexibility is positively related to Data-Driven Decision-Making.

Kache and Seuring (2017) highlight the important opportunities provided by the first order constructs of Big Data Analytics Management Capabilities which include coordination and control as per the research model. These are referred to as integration and collaboration and governance and compliance by Kache and Seuring (2017). Therefore, the second hypothesis is:

H2: Big Data Analytics Management Capabilities and Data-Driven Decision-Making are positively related.

Literature clearly mentions the need for expanding the skills in managers that need to use analytics for decision-making (Ransbotham et al., 2015b, 2016). Further to this, skills and data management capabilities, which form part of Big Data Analytics Personnel Expertise, are mentioned to be opportunities for improvement in supporting decision-making through the use of Big Data Analytics (Kache & Seuring, 2017; Kiron et al., 2014b). Therefore, the third hypothesis is:

H3: Big Data Analytics Personnel Expertise and Data-Driven Decision-Making are positively related.

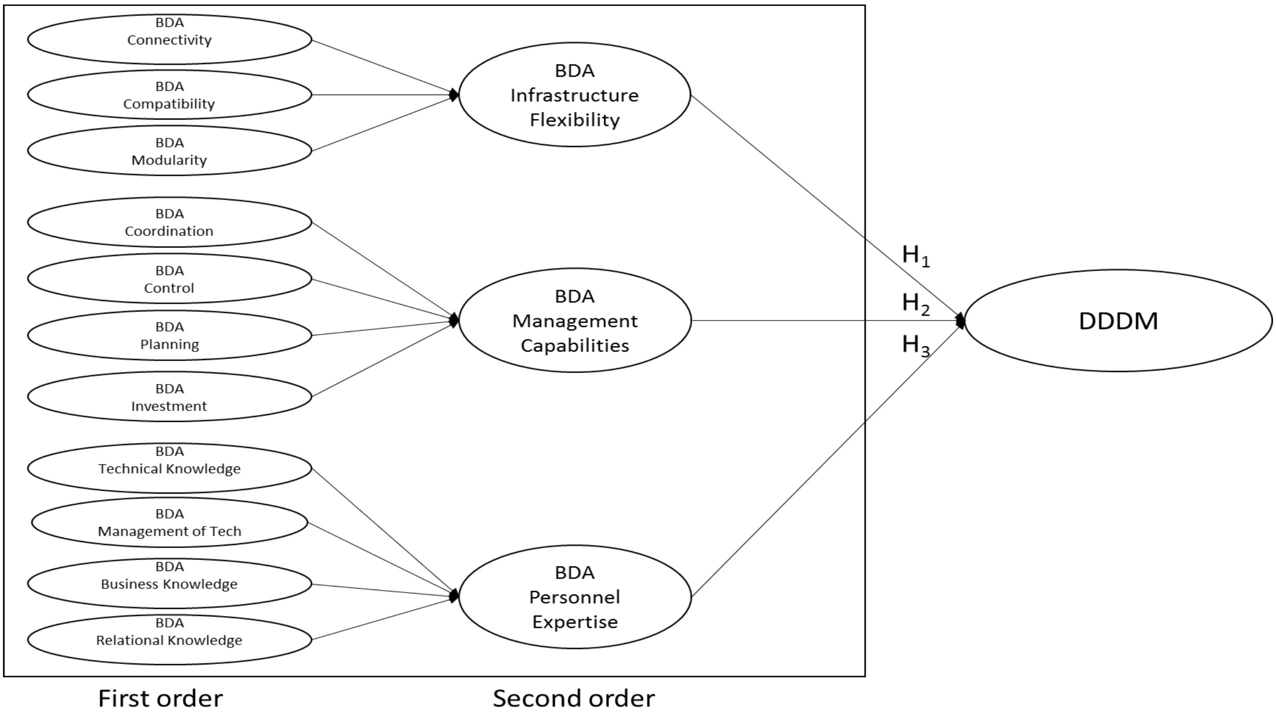


Figure 5: Research Model with Hypotheses

3.3 Research Proposition

The third order model for Big Data Analytics Capabilities was described in Section 2.5. The proposition which is derived from the research model as well as the discussion provided in Section 2.9, shown in Figure 6, is as follows:

P1: Big Data Analytics Capabilities are a predictor of Data-Driven Decision-Making.

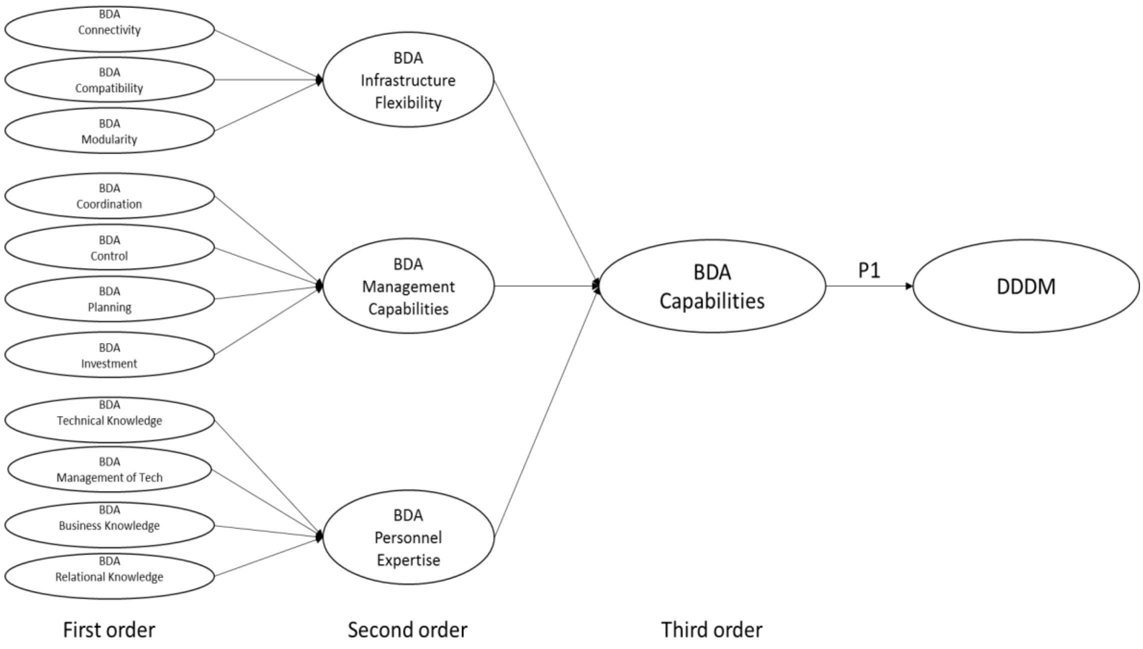


Figure 6: Research Model with Proposition

4 Research Methodology

This chapter outlines the methodology and design of the research conducted in this study. This includes a description of the methodological choices as well as the rationale for the selection. The methodological choices (such as research philosophy and research purpose), the population, unit of analysis, sampling, measurement instruments, data gathering process, analysis approach as well as the limitations are discussed in this chapter.

4.1 Choice of Methodology

This section describes the methodological choices made for the execution of this study.

4.1.1 Philosophy

The research philosophy employed in this study was positivism. The positivist research philosophy follows where the research aims to test and describe the relationship between observable and measurable variables (Bernroider & Schmöllerl, 2013; Saunders & Lewis, 2012). This philosophy is aligned with the objectives of this study since it aims to measure the constructs and then use statistical analysis to extrapolate relationships and associations between the independent and dependent variables. These variables are Big Data Analytics Capabilities and Data-Driven Decision-Making respectively. Studies by Wamba et al. (2017) as well as Bernroider and Schmöllerl (2013) are examples of research where positivism was employed.

4.1.2 Approach

A deductive approach was employed in this study. A deductive research approach is conducted to test theory as opposed to building theory in inductive research approaches. Deductive reasoning has been described as the process of going from a general principle to deriving a conclusion on a specific instance through the use of a logical process (Zikmund, Babin, Carr, & Griffin, 2010).

In this study, a conceptual model was developed from literature from which a research model was distilled and then tested using data. The hypotheses in this study were logically derived from existing literature such as Wamba et al. (2017), Janssen et al. (2017) and Brynjolfsson et al. (2011) which led to the deductive approach being adopted.

4.1.3 Methodology

This study made use of a mono-method. A mono-method adopts the use of a single data collection method and corresponding analysis procedure (Saunders, Lewis, & Thornhill, 2008). This method was applicable for this study since the collection of data was only done through one method. This method, being a questionnaire, was also undertaken by Bernroider and Schmöllerl (2013), in their study of Information Technology decision-making and decision-making satisfaction. Another example of a mono-method being applied, is the study by Gunasekaran et al. (2017) of organisational and supply chain performance in light of Big Data and Predictive Analytics.

4.1.4 Purpose of research design

The purpose of the research design is to provide a descriptive study. A descriptive study is one which provides a description of the characteristics of entities or concepts (Zikmund et al., 2010). The purpose of this study is to describe the relationship between Big Data Analytics Capabilities constructs and Data-Driven Decision-Making. It is for this reason that the research design was chosen to be descriptive.

4.1.5 Strategy

A survey with structured questions, which limits the number of responses, was adopted in this study (Zikmund et al., 2010). The survey strategy is typical of a descriptive study which is the purpose of this research (Zikmund et al., 2010).

In order to gather data to measure the constructs of Big Data Analytics Capabilities and Data-Driven Decision-Making, a survey strategy was adopted. This strategy facilitated the targeting of a sufficiently sizeable population of relevant respondents for the collection of data, due to it being able to be easily distributed. Furthermore, the structured nature of the survey lends itself to the measurement of defined constructs. This aligned with the positivist philosophy as well as the deductive approach of this study as it allowed for testing of theory through a logical process to reach conclusions. Studies such Brynjolfsson and McElheran (2016b), Ransbotham, Kiron, and Prentice (2015a) and Akter et al. (2016) have made use of survey strategies.

4.1.6 Time horizon

This study was cross-sectional research which means that it was conducted at a period in time. This type of time horizon has also been described as a snapshot or conducted at a point in time (Saunders & Lewis, 2012; Zikmund et al., 2010). This means that the study will describe the variables and explain the relationships between the variables at a point in time as was similarly done by Ji-fan Ren et al. (2017) and Chen et al. (2015), which is adequate for this study. Rindfleisch, Malter, Ganesan and Moorman (2008), in their comparison of cross-sectional and longitudinal research found that cross-sectional research is adequate under certain conditions which include having a well-designed survey. The choice of a cross-sectional time horizon is also due to the time constraints on the research project which does not allow for longitudinal study. The survey in this study was based on surveys used in prior research and was also pre-tested with experts to ensure the survey questions were sufficiently clear and were relevant to the study as well as to check if the time estimate was correct. The use of predefined constructs as we as the pre-test assisted in ensuring that a well-designed survey was used as per the recommendation by Rindfleisch et al. (2008).

4.1.7 Techniques and procedures

This study made use of self-administered or self-completed questionnaires using the Internet as a medium i.e. Internet-Mediated Questionnaires (Saunders et al., 2008). Self-administered questionnaires are where each of the respondents reads and answers the same set of pre-defined questions and thus takes the responsibility for responding to the questionnaire (Zikmund et al., 2010). This is done without an interviewer being present.

The self-administered questionnaire technique was selected for this study since the constructs and measurement items within the Big Data Analytics Capabilities and Data-Driven Decision-Making variables have been defined in literature such as Wamba et al. (2017) and Cao et al. (2015). Thus, a structured pre-defined questionnaire was constructed and distributed to respondents using the internet as the medium. The questionnaire is further discussed in Section 4.5 which covers the measurement instrument. The questionnaire can be found in Appendix A.

4.2 Population

The population for this study was managers and analytics professionals with decision-making responsibility in technical environments which include data engineering and software development. The reason for selecting this population was that the aim of this study is to gauge Data-Driven Decision-Making and the possible relationships that Big Data Analytics Business Capabilities constructs have in predicting Data-Driven Decision-Making. As such, this study aimed to get responses from professional individuals who have decision-making authority since they are able to provide applicable and informed responses which will enhance the value of the information gleaned from the study. Furthermore, these individuals are expected to have sufficient education on some level of Big Data Analytics or Business Analytics which will assist in getting suitable responses.

The selected population is not industry specific which can be viewed as advantageous in that insights and learnings from this study could be applied across various industries. Studies such as Kiron et al. (2014a), Ransbotham et al. (2016) and Cao et al. (2015)

have also used managers and analytics professionals as their population. Kiron et al. (2014a) and Ransbotham et al. (2016) had target populations across a wide range of industries and organisations of various sizes across the globe while Cao et al. (2015) focused on medium to large sized UK based organisations.

4.3 Unit of Analysis

The unit of analysis for this study was individual managers and analytics professionals. This is because the study aims to describe Data-Driven Decision-Making and its relationship with Big Data Analytics Capabilities, through the responses from managers and analytics professionals.

4.4 Sampling Method and Size

The sampling method for this study was purposive since the study required a sample of participants who would have some knowledge of Big Data, Data Analytics or Business Analytics. This was done by selecting participants who considered themselves as analytics professionals and managers in technical environments such as data, software engineering or similar. Purposive sampling is a non-probability sampling method where judgement is used by the researcher to select sample members based on some premises or reasons (Saunders & Lewis, 2012; Zikmund et al., 2010). Judgement will be exercised in the classification of the environment or department that the respondent is in, such as analytics or software engineering, rather than the industry. These departments are not industry specific and can be found across various industries.

Further to this, a snowball sampling method was used in order reach a larger portion of the population. However, care was taken to ensure that this did not detract from getting responses from the intended population by providing instructions for distribution of the survey as well as having a suitable filter question(s).

One of the guidelines provided relating to the application of snowball sampling was that of eliminating suspicious cases (Marcus, Weigelt, Hergert, Gurt, & Gelléri, 2017). Cases in which all questions had the same answer were considered suspicious in this study. A high (greater than 50%) nonresponse rate by a respondent was also considered unfit for inclusion in the data for analysis (this is discussed further in Chapter 5). The questionnaire contained a filter question(s) to ensure that the participants are valid members of the target sample i.e. they are decision makers with some knowledge of Big Data, Data Analytics, Business Analytics or Big Data Analytics.

Although there are differing opinions, to obtain the minimum size of the sample the guideline by (Pallant, 2010) which states that $N > 50 + 8m$ was used (where N is the sample size and m is the number of independent variables). This decision was made since this calculation is related to the use of multiple regression which is one the analysis methods to be used in this study (as discussed in Section 4.7). Using the secondary constructs of Big Data Analytics Capabilities (Big Data Analytics Infrastructure Flexibility, Big Data Analytics Management Capability and Big Data Analytics Personnel Expertise) as the independent variables ($m = 3$) a sample size of greater than 74 is required. Section 5.2 provides a description of the sample obtained.

4.5 Measurement Instrument

The measurement instrument that was used for this study was a self-administered questionnaire. The questions in the questionnaire were based on survey questions used in previous studies and revised for individual measurement. Wamba et al. (2017) provide the measurement items used to measure Big Data Analytics Capabilities in their study. Further to this, studies such as Akter, Wamba, Gunasekaran, Dubey, and Childe (2016) and Liberatore et al. (2017) provide measurement items for similar constructs relating to the study and were thus considered for inclusion in the questionnaire. An example of a question relating to compatibility of Big Data Analytics Infrastructure Flexibility is "Software applications can be easily used across multiple analytics platforms" (Wamba et al., 2017, p. 360). A second example of a question included in the questionnaire, which relates to personnel (technical) expertise, is "I am very capable in the areas of data management and maintenance" (Wamba et al., 2017, p. 360). The Data-Driven Decision-Making construct questions were adopted from Cao et al. (2015) and U.S. Census Bureau (2015). An example of a Data-Driven Decision-Making question included in the survey which relates to product and service development states "We use data-based insight for the creation of new services/products" (Cao et al., 2015, p. 387). An older version (2010) of management and organisation practises which was developed the U.S. Census Bureau (under the U.S. Department of Economics and Statistics Administration) was used by Brynjolfsson and McElheran (2016b, 2016a).

The questionnaire made use of a seven-point Likert scale. Likert scales are a well-established method for measuring the extent to which respondents agree or disagree (positive and negative) with a statement and, although they have limitations, are commonly used in research (Lantz, 2013; Zikmund et al., 2010). The questionnaire is provided in Appendix A, with measurement instruments from literature (Akter et al., 2016; Cao et al., 2015; U.S. Census Bureau, 2015; Wamba et al., 2017).

Validity is defined as the "accuracy of a measure or the extent to which a score truthfully represents a concept" (Zikmund et al., 2010, p. 307). In order to measure a construct such as Big Data Analytics Capabilities it was necessary to ensure that questions included within the questionnaire actually measure the construct. Construct validity was ensured by making use of high quality literature and cross checking against various sources. As was done in Wamba et al. (2017) and Akter et al. (2016) the content validity of the survey was conducted by experienced analytics experts in the field. Convergent and discriminant validity was addressed to ensure that constructs that should be related are related and those that should not be related are not. This was done by means of conducting bivariate correlation analysis between each construct and construct total (Hair, Black, Babin, & Anderson, 2010; Kinnear & Gray, 2012). Confirmatory Factor Analysis (CFA) was also employed to test validity since the research model was based on constructs from Wamba et al. (2017) which were previously tested for validity. Studies by Kock and Gemünden (2016) as well Chavez, Yu, Jacobs and Feng (2017) also employed CFA. CFA is further discussed in Chapter 5.

Reliability is described as a measure of the internal consistency of a measure which means that it provides an indication of the consistency in the findings that result from the construct (Saunders et al., 2008; Zikmund et al., 2010). The first means of ensuring reliability included careful design of the questions, logical and clear layout of the questionnaire, clear explanation of the purpose of the questionnaire, careful administration as well as pilot testing (Saunders et al., 2008; Zikmund et al., 2010).

The questionnaire was pre-tested through a pilot study which was sent to eight experts who are seasoned professionals in the fields of Big Data and Big Data Analytics was carried out to ensure the reliability and validity of constructs, although only four completed the study in time. Feedback relating to the survey questionnaire was elicited through a feedback template which is provided in Appendix B. The feedback questionnaire asked questions relating to the length of the questionnaire, clarity of the questions, structure and format of the questionnaire and, most importantly, whether the experts felt that the questionnaire omits any factors of Big Data Analytics Capabilities and Data-Driven Decision-Making. The feedback received was positive in general. However, there was some ambiguity identified in two of the questions. The first ambiguity was addressed through adding of a word to provide additional clarity. The second question with ambiguity was analysed for risk of omission of key component of survey and was consequently removed due to the question being considered too ambiguous for inclusion.

Additionally, Cronbach's Alpha, which is a measure of reliability and internal consistency, was calculated to test the reliability of constructs. Wamba et al. (2017) note that the Cronbach's Alpha values for the Big Data Analytics Capabilities constructs in their study all exceeded 0.7 which provides confidence into the reliability of the instrument. Chavez, Yu, Jacobs, and Feng (2017) as well Côrte-Real et al. (2017) also used Cronbach's Alpha to address reliability in their study. Construct validity and reliability are further discussed in Sections 5.3 and 5.4.

4.6 Data Gathering Process

The data was collected by means of a self-administered questionnaire using an internet-mediated questionnaire. The questionnaire was in electronic format available through the internet. The link to the online survey was distributed through electronic communication tools such as email and instant messaging platforms. Studies such as Ransbotham, Kiron, and Prentice (2016) and Côrte-Real et al. (2017) have also employed a similar data gathering process making use of an internet-mediated self-administered questionnaire.

The data collection process proved to be a challenge in terms of obtaining suitable (i.e. target sample) responses to the questionnaire from a sufficiently large sample. In an attempt to attain the maximum number of responses from the target sample the researcher sent 80 direct messages (via instant messaging and email) to members of the target population known to the researcher. This included Chief Data Officer of a large financial institutions, senior data scientists at large telecommunications operators as well as senior managers of analytics at insurance providers and banks, all of whom have large Big Data and data analytics teams which are potential respondents.

Further to this members in the researcher's network undertaking postgraduate studies provided an opportunity to gain access to the member of the classes (which are part of the target population). The link to the survey was sent to two MSc in Big Data classes and a post-graduate diploma in data analytics which provided reach to a further 160 potential respondents. Furthermore, the researcher approached an internal Big Data workgroup within the researcher's organisation which comprised 55 members. In addition to this the researcher posted the link to the questionnaire on five LinkedIn groups with a combined total of 844 361 members. A summary of the data gathering efforts and potential respondent numbers are summarised in Appendix C.

It was important that netiquette was observed so that participants were not offended or any of their rights infringed upon (Saunders et al., 2008). This also related to ethics whereby the rights of participants were protected. To facilitate this, the proposed study was submitted to the Research Ethics Committee (REC) for ethical clearance approval prior to conducting the study. The ethical clearance letter obtained for this study from the Research Ethics committee is provided in Appendix E.

As mentioned is Section 4.5, a pilot study was conducted with seasoned experts in the fields of Big Data and Big Data Analytics to test validity and reliability of the questions and to highlight any issues that needed to be addressed prior to distributing the questionnaire. As also alluded to in Section 4.5, the feedback from the pilot study was taken into account to ensure that the data gathering process continues with minimal disruptions and that valid and reliable data was gathered for the study.

4.7 Analysis Approach

The analysis approach needed to align with the purpose of the research design which was previously described as descriptive in Section 4.1.4. Thus, the analysis undertaken needed to provide the information necessary to provide a description of the relationships between the Big Data Analytics Capabilities and Data-Driven Decision-making variables.

The data was analysed using statistical techniques using the IBM SPSS analytics platform. Firstly, descriptive analysis was done to provide descriptive statistics such as central tendency, dispersion and trends (Saunders & Lewis, 2012; Zikmund et al., 2010). Measures of central tendency include the mean, median and mode while dispersion can be described through the standard deviation. The descriptive statistics assisted in gaining insight into the data by simplifying the data into manageable and concise information. These were graphically represented which assisted in becoming familiar with the data.

Pearson’s correlation is a parametric statistical test which analyses the association between variables (Pallant, 2010). The non-parametric equivalent is the Spearman Rank Order correlation which is used when the data does not fulfil the criteria for Pearson correlation (Pallant, 2010). Pearson’s and Spearman’s correlation analysis were conducted, depending on the fulfilment of the criteria for Pearson’s correlation, between each of the constructs of Big Data Analytics Capabilities (Big Data Analytics Infrastructure Flexibility, Big Data Analytics Personnel Expertise, Big Data Analytics Management Capability) and the Data-Driven Decision-Making. This revealed the level of significance, the strength as well as the type (positive or negative) of the relationships between the sets of variables. Table 1 provides a guideline for describing the strength of correlations recommended by Cohen (1988, pp. 79–81) (as cited in Pallant (2010)). The correlation analysis is a key measure since it links directly to the research hypotheses for this study. The results of the analysis is provided is Chapter 5 and the discussion of these results are presented in Chapter 6. Kock and Gemünden (2016) have also utilised correlation analysis to analyse the relationships of the variables in their study relating to the antecedents of decision-making quality.

Table 1: Description of Correlation Strength (Pallant, 2010, p. 134)

Description	Coefficient Value
Small	.10 - .29
Medium	.30 - .49
Large	.50 – 1.0

Furthermore, multiple linear regression analysis was conducted which is used when there are several independent variables and one dependent variable, as is the case in this study. Multiple regression is a commonly used multivariate technique in research since it allows for the testing of the prediction of a dependent variable by the independent variables (Hair et al., 2010). Furthermore, multiple regression analysis allows for the explanation of the relationship between the independent and dependent variables through the analysis of the attributes of the regression coefficients in the generated (linear) model (Hair et al., 2010). Further to the regression coefficients in the model, the significance of the modelled relationships was provided with the multiple linear regression outputs. The generated outputs provided insight into synthesising an explanation of the relationships between the variables which is further discussed in Chapters 5 and 6. Brynjolfsson and McElheran (2016a) conducted similar analyses in their study relating to the adoption of Data-Driven Decision-Making in the U.S. manufacturing industry.

4.8 Limitations

Various limitations relating to the study were considered. The first of these being the accessibility of respondents to the questionnaire. Significant effort was taken to ensure that a large enough sample of respondents were obtained through personal and professional networks and professional online groups.

A further limitation in the research methodology of the study was that of using non-random sampling. This may mean that the sample is not necessarily adequately representative of the population and thus the results may not be generalised to the population.

Although steps were taken to reduce the effects of these, further limitations relate to participant error and participant bias in completing the questionnaires. Careful attention was paid to the questions and their design; however, this may still be a limitation in the study.

Furthermore, this study was a cross-sectional study and thus captured a snapshot of the environment at a point in time. The Big Data Analytics and business environment is constantly evolving and thus the results of the study may only be relevant for a limited period of time.

It should also be noted that this study aimed to investigate the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making through the use of correlation and multiple regression analysis. The existence of significant relationships between variables does not imply causality and thus should not be interpreted as such in this study.

5 Results

Following on from the methodology description in Chapter 4, this chapter provides the results of the statistical analysis conducted on the data collected from the survey. This was done in in order to address the research hypotheses and proposition that were defined in Chapter 3.

The chapter begins with a description of the data preparation carried out and a description of the sample based on the demographic data. This is then followed by the validity and reliability checks of the constructs, factor analysis and testing of the assumptions for the correlation and regression tests. This chapter then moves on to providing the results from the statistical analysis of each of the three research hypotheses as well as the proposition as laid out in Chapter 3.

5.1 Data Preparation

As mentioned in Section 4.6 the data was collected by means of a self-administered questionnaire in electronic format through an online survey tool (Google Forms). The collected data was then exported as a Microsoft Excel file which was then imported into IBM SPSS which was the statistical analysis platform used to conduct the data analysis. All responses to the questionnaire were collected and stored in text format and thus required preparation before being ready for analysis. The following subsections describe the data preparation steps that were carried out.

5.1.1 Coding

In preparation for analysis, the data (formatted as text) was coded into numerical values. The collected data consisted of categorical data for the demographics section and ordinal data for the sections with questions relating to measuring the constructs. The seven-point Likert scale responses were coded as shown in Table 2. A full outline of the codes used (code book) for the entire data set is provided in Appendix D.1.

Table 2: Likert Scale Coding

Likert Scale	Coded Value
Strongly disagree	1
Disagree	2
Moderately disagree	3
Neutral	4
Moderately agree	5
Agree	6
Strongly agree	7

5.1.2 Missing Values

A total of 95 responses were received for the survey. However, two responses were excluded due to more than 50% of questions being not answered, thus reducing the number of usable responses from 95 to 93.

Figure 7 provides a plot of the number of nonresponses per question from respondents excluding the two which were removed from the. This shows that questions 14 (CP2) and 17 (MOD2) had the largest number of nonresponses with 10 and 7 missing values respectively. Appendix D.9, which provides descriptive statistics per question, may be used as a reference for questions and question codes

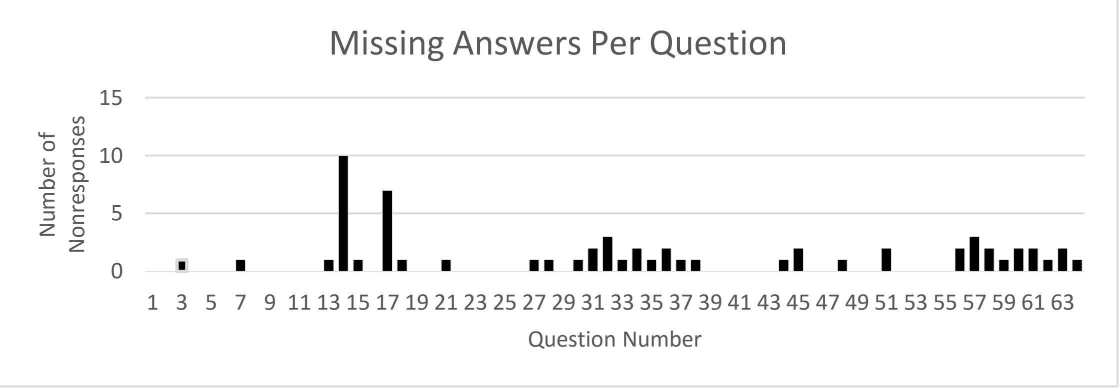


Figure 7: Missing Answers Per Question

The total number of questions in the questionnaire was 64 with 55 questions directly relating directly to research constructs and the rest being demographically related. The total number of nonresponses for the 55 questions in the survey was 59. With a sample of 93 respondents the total number of survey questions to be answered was 5115. Thus, the percentage of nonresponses was 1.15% for the 93 respondents.

5.1.3 Imputation

In order to deal with nonresponses to questions or missing values, imputation was used. Imputation is the process of filling in missing data points through the use of a statistical method that provides a suitable substitute for this missing value (Zikmund et al., 2010). This was done by making use of the mean substitution imputation method which is suitable in cases of an overall low level of nonresponses (Hair et al., 2010). To implement this the mean score per question was calculated according to industry and the missing values imputed based on the industry category of the respondent. This process was followed since it would provide a more representative substitute value than the mean of the entire sample.

5.2 Sample Description

As mentioned previously 93 usable responses to the questionnaire were received. As described in Section 4.4 the primary sampling method used was that of purposive with snowballing thereafter. Appendix C outlines the various means employed to ensure that data was gathered from a sample of adequate size, as alluded to in Section 4.6.

5.2.1 Education

The highest level of education of 63.4% of the respondents was of a postgraduate degree, 21.5% was an undergraduate degree and 14% having a diploma or certificate. This shows a highly educated sample of respondents and is graphically presented in Figure 8.

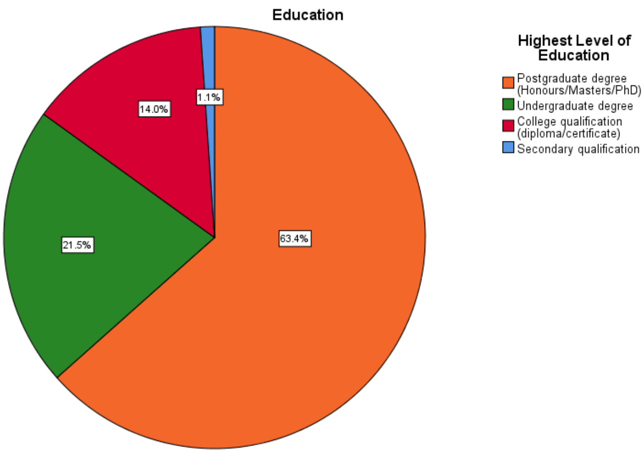


Figure 8: Pie Chart Depicting the Level of Education of Respondents

5.2.2 Age

The 26 -33 years age group had the largest group of respondents with 44.1% followed by the 34 – 41 years with 24%. Thus, respondents between the ages of 26 and 41 made up 68.1% of the sample. These were followed by the 42-49 years, the 50 years or older group and the 18-25 years group with 12.9%, 9.7% and 7.5% respectively. This is presented as a pie chart in Figure 9.

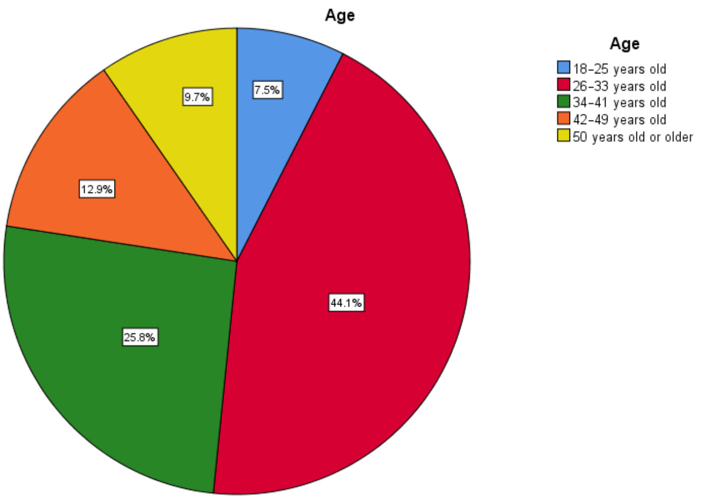


Figure 9: Pie Chart Depicting the Age of Respondents

5.2.3 Gender

Majority of respondents were male with 81.7% and females with 17.2% and 1.1% did offer a response to this question. This is graphically represented in Figure 10.

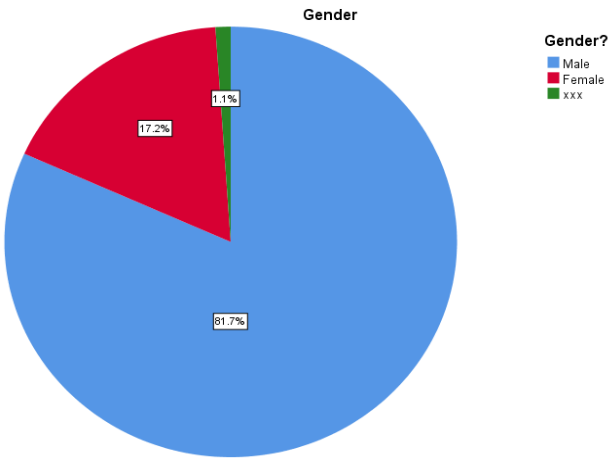


Figure 10: Pie Chart Depicting the Gender Split of Respondents

5.2.4 Industry

The top four industries from which responses were obtained were from information and communication (29%), financial and insurance (20.4%), transportation and storage (16.1%) and professional, scientific and technical activities (14%). These accounted for slightly below 80% of the total responses. Table 3 provides an outline of all the industries from which responses were obtained.

Table 3: Responses Per Industry

Industry	Frequency	Percent
Information and communication	27	29,0
Manufacturing	5	5,4
Mining and quarrying	3	3,2
Professional, scientific and technical activities	13	14,0
Transportation and storage	15	16,1
Other service activities	4	4,3
Arts, entertainment and recreation	1	1,1
Construction	1	1,1
Education	5	5,4
Financial and insurance activities	19	20,4
Total	93	100

5.2.5 Occupation Level

Looking at the level at which respondents operate within their organisations as shown in Table 4, 12.9% responded as being entry level employees. Of all the categories, the largest proportion of respondents (32.3%) identified themselves as middle management. Senior management responses accounted for 25.8% of responses and junior management made up 22.6%. the smallest proportion of respondents were Owners and C-level executives with 6.5% of the total.

Table 4: Level of Occupation

Occupation Level	Frequency	Percent
Entry Level	12	12.9
Junior Management	21	22.6
Middle management	30	32.3
Senior management	24	25.8
Owner/C-Level executive	6	6.5
Total	93	100.0

5.2.6 Big Data Analytics Knowledge

Of the total respondents, more than half of respondents (53.8%) declared an intermediate level of knowledge of Big Data Analytics and 13.2% having advanced knowledge. Furthermore, 27.5% of respondents identified themselves as having basic knowledge of Big Data Analytics and 5.4% declared that they had no knowledge of Big Data Analytics. These values are shown in Table 5. Further to this 39.6% of respondents identified themselves as analytics professionals.

Table 5: Level of Big Data Analytics Knowledge

	Frequency	Percent
None	5	5.4
Basic	25	26.9
Intermediate	50	53.8
Advanced	13	14.0
Total	93	100.0

5.3 Validity

Validity was tested through the use of Pearson correlation between each question within a construct and the construct total score (Hair et al., 2010; Kinnear & Gray, 2012). The results of these analyses is provided in Appendix D.4. All correlations were significant with the correlation coefficients ranging from 0.61 to 0.92 with the exception of the fourth modularity question (MOD4) with $r(91)=.23$, $p = .028$, which is slightly below the recommendation of 0.3. This shows that in general convergent validity is present within the constructs, however, the fourth modularity construct question (which forms part of the Big Data Analytics Infrastructure Flexibility second order construct) will be noted for consideration to be removed in the reliability analysis. As will be recalled from Section 4.6 a pre-test was also conducted to test content validity through feedback from industry experts.

5.4 Reliability

The reliability of constructs was checked through the use of the Cronbach's Alpha measure. This was done for each of the constructs as summarised in Table 6. Table 6 shows that nine of the twelve constructs had Cronbach's Alpha values of greater than 0.7 which indicates good to very good reliability (Zikmund et al., 2010). The compatibility, modularity and relational knowledge constructs have Cronbach's Alpha values between 0.65 and 0.7 which are considered as a representation of fair reliability (Zikmund et al., 2010)

It should be noted that one item each was removed from the Modularity and Relational Knowledge constructs in order to improve Cronbach's Alpha scores to acceptable level (Hair et al., 2010). These items were MOD4 (which was already noted from the validity analysis for potential removal) and RK4 respectively. Question RK4 dealt with working with customers and maintaining customer relationships.

Table 6: Cronbach's Alpha Values

Constructs	Cronbach's Alpha
Connectivity	0.71
Compatibility	0.67
Modularity	0.67
Planning	0.92
Investment Decisions	0.85
Coordination	0.74
Control	0.91
Technical Knowledge	0.82
Technical Management Knowledge	0.76
Business Knowledge	0.71
Relational Knowledge	0.65
DDDM	0.93

5.5 Factor Analysis

Factor analysis is a technique that is used to delineate the structure of the variables being used in the analysis (Hair et al., 2010). This section describes the two factor analysis techniques undertaken in this study which are Confirmatory Factor Analysis (CFA) and Exploratory Factor Analysis (EFA). Due to the sensitivity of CFA to sample size and the sample size in this study being somewhat low and lower than recommended fit metrics being attained for some constructs, both CFA and EFA were conducted (Beavers et al., 2013; Hair et al., 2010). Furthermore, CFA and EFA were conducted since it allowed for gaining a useful understanding and insight into the model.

5.5.1 Confirmatory Factor Analysis

Confirmatory Factor Analysis is generally used to confirm or test assumptions or concepts relating to a, possibly already validated, structure of a set of variables (Pallant, 2010). Since the constructs used in this study are from previously tested and validated research by Wamba et al. (2017), Cao et al. (2015) and (Brynjolfsson et al., 2011), Confirmatory Factor Analysis was conducted to test the validity and model fit with the collected data. However, the sample size available for this study being 93, could pose a challenge for CFA as large sample size has been cited as a requirement for CFA (Hair et al., 2010; Pallant, 2010). The quantification of a large sample is a topic which has not been agreed upon and researchers have offered varying approaches such as setting a value of 300 as a comfort factor for number of cases when undertaking factor analysis (Pallant, 2010). Furthermore, CFA has also been noted to being susceptible and sensitive to very large sample sizes (Hair et al., 2010).

CFA was conducted for all first order and second order constructs of Big Data Analytics Capabilities and Data-Driven Decision-Making, using the IBM SPSS AMOS package. These AMOS models are provided in Appendix D.2. Table 7 provides a summary of the key metrics from the second order constructs. Big Data Analytics Infrastructure Flexibility has a good Comparative Fit Index (CFI) of 0.99 and Data-Driven Decision-Making slightly below the recommended 0.92 with a CFI of 0.905. However, Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise had lower than recommended CFI values with 0.835 and 0.789 respectively. The Root Mean Square Error of Approximation (RMSEA) for Big Data Analytics Infrastructure Flexibility meets the recommendation of less than 0.08 with 0.027 while Big Data Analytics Management Capabilities, Big Data Analytics Personnel Expertise and Data-Driven Decision-Making have larger values of 0.128, 0.121 and 0.135. For the Standard Root Mean Square Residual (SRMR) measure Big Data Analytics Management Capabilities and Data-Driven Decision-Making meet the recommendation of less than 0.08 with values of 0.0718 and 0.0535 respectively while Big Data Analytics Infrastructure Flexibility and Big Data Analytics Personnel Expertise have SRMR values that are above the recommendation for good fit. Although the recommended values have been questioned and are subject to a number of factors, it is noted that the model fit challenge could be due to the low sample size in the data set (Hair et al., 2010; Pallant, 2010). Furthermore, with the exception of a few, the factor loadings (shown in Appendix D.2) were acceptable across the second order constructs.

Table 7: CFA Metrics for Second Order Constructs

	CFI	RMSEA	SRMR	Chi-squared
BDAIF	0.99	0.03	0.63	34.11
BDAMC	0.84	0.13	0.07	365.43
BDAPE	0.79	0.12	0.14	231.04
DDDM	0.91	0.14	0.05	93.69

As noted in the above discussion, there were shortcomings in some of the results obtained for the loadings, CFI, SRMR and RMSEA. Therefore, to further explore and gain more confidence in the model, it was decided to conduct Exploratory Factor Analysis (EFA) which will be discussed in Section 5.5.2 (Beavers et al., 2013). Ideally, after conducting CFA it would not be necessary to conduct EFA.

5.5.2 Exploratory Factor Analysis

Exploratory Factor Analysis (EFA) is often done in research to explore, possibly immature, variable structures and their interrelationships (Hair et al., 2010). In this study, since the model used is based on tested and validated constructs, EFA will serve as a secondary method to assessing the validity of the model constructs after conducting CFA. However, before undertaking EFA the constructs were assessed or tested for the appropriateness of factor analysis.

Two statistical tests as well as visual inspection of correlations were used for testing whether factor analysis is appropriate and can be applied to the model. These were the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy (MSA) as well as Bartlett's Test of Sphericity which were conducted for each first-order construct (Hair et al., 2010). Bartlett's Test of Sphericity statistically tests whether inter-item correlations are manifested (Hair et al., 2010). Principal Component Analysis (PCA) was applied as the dimension reduction technique after confirmation of the appropriateness of factor analysis.

Visual inspection of the correlation matrices was conducted and confirmed that there were correlations between each variable and at least one other variable with a correlation coefficient of greater than 0.3 (Hair et al., 2010; Pallant, 2010). The correlation matrices are provided in Appendix D.3 for reference.

The KMO Measure of Sampling Adequacy results, as summarised in Table 8, indicated that none of the constructs were below the 0.5 threshold of being unacceptable (Hair et al., 2010). However, it is noted that the connectivity construct has the lowest score for KMO at 0.596 which falls marginally below the level of mediocre. Compatibility, modularity, coordination, business knowledge and relational knowledge had measure between 0.6 and 0.7 which translates to mediocre. The investment decision making, technical knowledge and technological management constructs had "middling" measures of between 0.7 and 0.8 while planning, control and meritorious had Data-Driven Decision-Making had meritorious measures higher than 0.8 (Field, 2013). Thus, the KMO test results provides an indication of factor analysis being appropriate for this study.

The Bartlett's Test of Sphericity for all constructs are significant which indicates that the correlations between variables are sufficient and thus further confirms the applicability of factor analysis. The results of Bartlett's Test of Sphericity are also included in Table 8.

Table 8: KMO and Bartlett's Test of Sphericity

Construct	Kaiser-Meyer-Olkin Measure of Sampling Adequacy.	Meaning	Bartlett's Test of Sphericity		
			Approx. Chi-Square	df	Sig.
Connectivity	0.60	Miserable	61.08	3	0.000
Compatibility	0.65	Mediocre	41.92	3	0.000
Modularity	0.61	Mediocre	53.89	6	0.000

Planning	0.82	Meritorious	277.57	6	0.000
Investment Decision Making	0.79	Middling	151.01	6	0.000
Coordination	0.66	Mediocre	62.51	3	0.000
Control	0.87	Meritorious	552.09	28	0.000
Technical Knowledge	0.76	Middling	138.95	6	0.000
Technological Management	0.71	Middling	101.29	6	0.000
Business Knowledge	0.62	Mediocre	122.69	6	0.000
Relational Knowledge	0.69	Mediocre	67.95	6	0.000
DDDM	0.89	Meritorious	631.38	45	0.000

Principal Component Analysis (PCA) is a dimension reduction technique which is used to capture the maximum amount of variance or information while making use of the minimum amount of factors (Hair et al., 2010; Pallant, 2010). This was done using the Principal Component extraction method with Kaiser's criterion (eigenvalue greater than 1 rule) with a setting of 25 for the maximum number of iterations for convergence in SPSS (Pallant, 2010). The Varimax rotation method was used which is an orthogonal rotation method that minimises the number of variables with high factor loadings (Hair et al., 2010; Pallant, 2010). Factor rotation is used due to it generally reducing ambiguities and hence improving the interpretation of the factors as opposed to unrotated solutions (Hair et al., 2010; Pallant, 2010).

Table 9: PCA Results Summary

Construct	Number of Items	Number of Components Extracted	Cumulative % of variance
Connectivity	3	1	63.70
Compatibility	3	1	60.50
Modularity	3	1	47.77
Planning	4	1	80.93
Investment Decision Making	4	1	68.56
Coordination	3	1	65.83
Control	8	2	78.86
Technical Knowledge	4	1	65.94
Technological Management	4	1	58.97
Business Knowledge	4	2	81.43
Relational Knowledge	3	1	52.69
DDDM	10	1	61.87

Table 9 provides a summary of the results from the PCA which was conducted for all of the first order constructs. All constructs loaded onto one component besides control and business knowledge which loaded onto two component each.

The items loaded onto each of the two components from PCA for the control construct are summarised in Table 10. The first component (Control1 in Table 10) explained 62.76% of the variance and component 2 (Control2 in Table 10) explained 16.09% of the variance.

Table 10: Control Construct Components

Component	Included Items				% Variance
Control1	COL1	COL2	COL3	COL4	62.76
Control2	COL5	COL6	COL7	COL8	16.09

Table 11 summarises the components and item loadings for the Business Knowledge construct. The total variance explained for the two components (with eigenvalues greater than one) is 81.43% with component 1 contributing 56.12% and component 2 contributing 25.31%.

Table 11: Business Knowledge Construct Components

Component	Included Items		% Variance
BusKnowl1	BK3	BK4	56.12
BusKnowl2	BK1	BK2	25.31

5.6 Normality

The assumption that the distribution of data is normal is made in the application of parametric statistical analysis and thus affects the tests that can be validly applied (Hair et al., 2010). Therefore, the data collected in this research was tested through the use of the Shapiro-Wilk and Kolmogorov-Smirnov tests for normality. These tests for normality test the significance of the difference between the construct and a normal distribution (Hair et al., 2010).

Table 12 shows the results of the tests for normality for each of the first order constructs which shows that four of the constructs (Connectivity, Modularity, Business Knowledge, DDDM) are normally distributed while the rest are not. Furthermore, there is a discrepancy between the two tests for the Control and Business Knowledge constructs. Thus as recommended by Hair et al. (2010) graphical plots were also examined to assess the normality of the constructs. The histograms and Q-Q plots are provided in Appendix D.5.

Table 12: Tests of Normality Results - First Order Constructs

1st Order Constructs	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Connectivity	0,073	93	.200*	0,980	93	0,168
Compatibility	0,106	93	**0,012**	0,972	93	**0,040**
Modularity	0,075	93	.200*	0,987	93	0,480
Planning	0,100	93	**0,024**	0,962	93	**0,008**
Investment Decision Making	0,118	93	**0,003**	0,963	93	**0,010**
Coordination	0,108	93	**0,009**	0,970	93	**0,032**
Control	0,080	93	0,188	0,968	93	**0,021**
Technical Knowledge	0,121	93	**0,002**	0,941	93	**0,000**
Technological Management	0,139	93	**0,000**	0,943	93	**0,001**
Business Knowledge	0,105	93	**0,013**	0,978	93	0,121
Relational Knowledge	0,155	93	**0,000**	0,934	93	**0,000**
DDDM	0,064	93	.200*	0,980	93	0,159

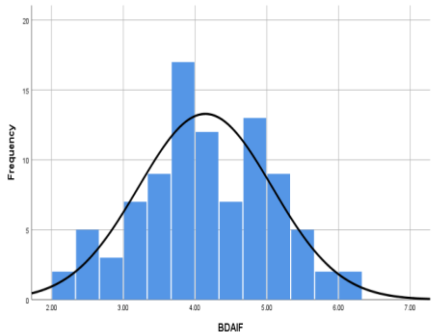
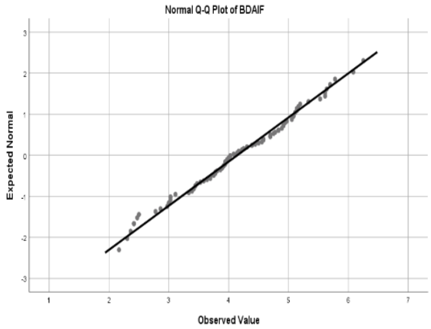
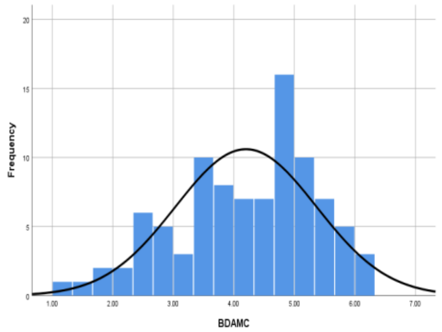
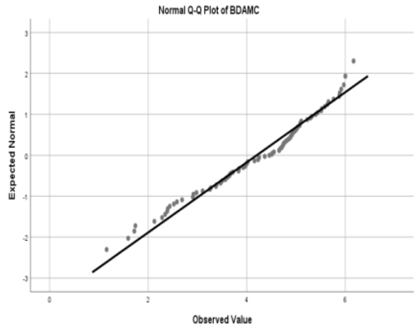
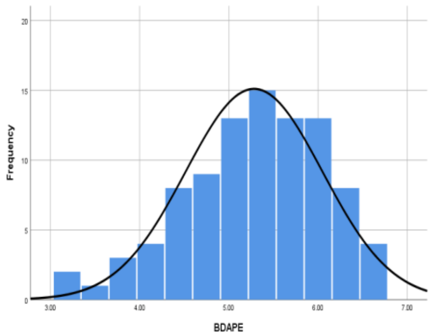
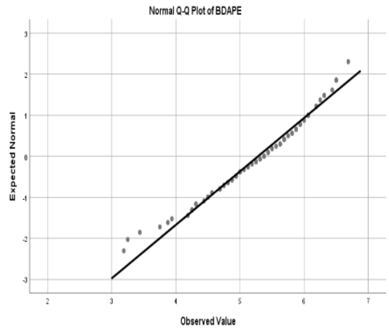
The second order constructs Big Data Analytics Infrastructure Flexibility (BDAIF), Big Data Analytics Management Capabilities (BDAMC) and Big Data Analytics Personnel Expertise (BDAPE) were also tested for normality. Table 13 shows the results of the statistical tests for normality (Shapiro-Wilk and Kolmogorov-Smirnov) which reveals that Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise are not normally distributed while Big Data Analytics Infrastructure Flexibility is.

Table 13 Tests of Normality Results - Second Order Constructs

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
BDAIF	0,056	93	.200	0,986	93	0,425
BDAMC	0,114	93	**0,004**	0,966	93	**0,017**
BDAPE	0,070	93	.200	0,973	93	**0,051**

The histograms and Q-Q plots for the second order Big Data Analytics Capabilities constructs are shown in Table 14 which allows for visual evaluation of the normality.

Table 14: Histograms and Q-Q Plots – Second Order Constructs

<table>
<tr><th>Histogram</th><th>Q-Q Plot</th></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>
</td><td>
</td></tr>
<tr><td>
</td><td>
</td></tr>
</table>

5.7 Linearity

A linear relationship between the independent and dependent variables is an assumption made for the statistical testing of relationships such as correlation analysis (Pallant, 2010). The relationships between the second order constructs and the dependent variable (Data-Driven Decision-Making) were evaluated through scatter plots. Figures 11 Figure 11to Figure 13 shows the scatter plots relevant to research hypotheses 1 to 3 respectively.

A visual inspection ofFigure 11 provides an indication of linearity between Big Data Analytics Personnel Expertise to Data-Driven Decision-Making.

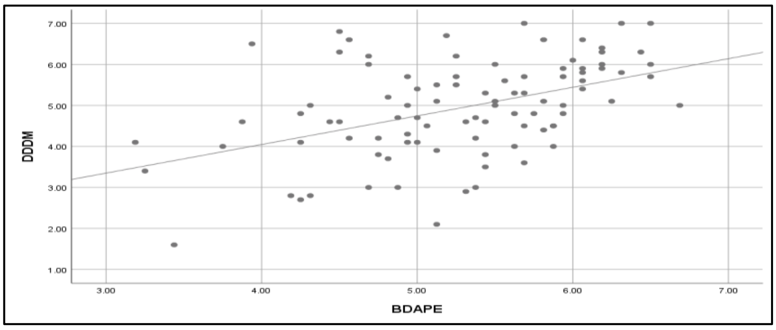


Figure 11: Scatter Plot of BDAPE to DDDM

Figure 12 also suggests linear relationship between Big Data Analytics Infrastructure Flexibility (BDAIF) and Data-Driven Decision-Making (DDDM).

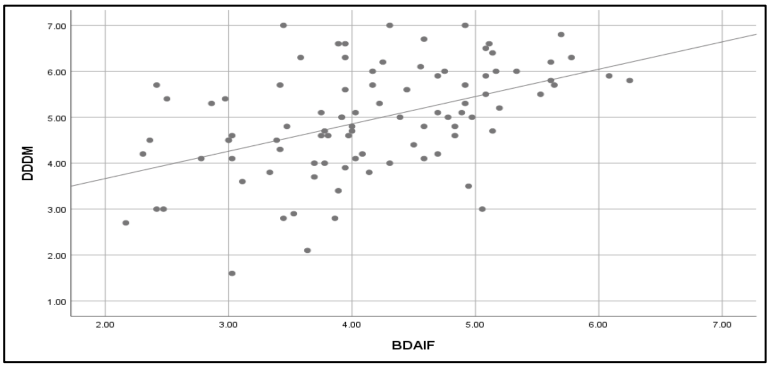


Figure 12: Scatter Plot of BDAIF to DDDM

The scatter plot depicted in Figure 13 suggests a linear relationship between Big Data Analytics Management Capabilities (BDAMC) and Data-Driven Decision-Making (DDDM) respectively.

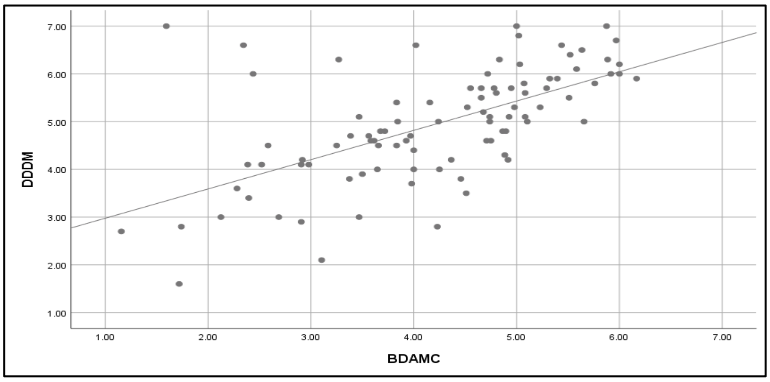


Figure 13: Scatter Plot of BDAMC to DDDM

5.8 Homoscedasticity

Homoscedasticity or homogeneity of variance is an assumption for correlation and regression analysis which are tests to be utilised in this study (Field, 2013; Pallant, 2010). This is evaluated through inspecting the scatter plots of standardised residuals and predicted dependent values for the equal variance of the error terms. The scatter plots are provided in Appendix D.6 and provide little indication of violation of homoscedasticity.

5.9 Multicollinearity

Multicollinearity is a further assumption of regression testing that needs to be validated to ensure that independent variables are not highly correlated (Hair et al., 2010; Pallant, 2010). A correlation between variables of higher than 0.7 is an indication of possible multicollinearity (Pallant, 2010). Appendix D.7 provides the correlation coefficient values which range between 0.23 and 0.67 which are within the recommended range (Pallant, 2010).

The recommendation is that, at a minimum, the tolerance value be greater than 0.1 and Variance Inflation Factor (VIF) be less 10. Therefore, the tolerance and VIF values provided in Table 15 indicate that there is little concern of multicollinearity.

Table 15: Collinearity Statistics

	Collinearity Statistics	
	Tolerance	VIF
BDAIF	0.54	1.85
BDAMC	0.54	1.87
BDAPE	0.87	1.15

5.10 Research Hypothesis 1

Hypothesis one posited that there is a positive relationship between Big Data Analytics Infrastructure Flexibility and Data-Driven Decision-Making. A summary of the descriptive statistics for Big Data Analytics Infrastructure Flexibility and Data-Driven Decision-Making is provided in Table 16. The mean value of 4.14 for Big Data Analytics Infrastructure Flexibility shows that the average response for Big Data Analytics Infrastructure Flexibility is slightly above "Neutral".

Table 16: Descriptive Statistics for BDAIF and DDDM

	N	Mean	Median	Mode	Std. Deviation
BDAIF	93	4.14	4.03	3.94	0.93
DDDM	93	4.94	5.00	4.60	1.17

In order to test the relationship between Big Data Analytics Infrastructure Flexibility and Data-Driven Decision-Making, a Pearson correlation was conducted, after confirming that all assumptions were met. Table 17 shows the results of the Pearson correlation which illustrates a significant, medium, positive correlation between Big Data Analytics Infrastructure Flexibility and Data-Driven Decision-Making with $r(91)=.47$, $p < .001$.

Table 17: Pearson Correlation between BDAIF and DDDM

		BDAIF	DDDM
BDAIF	Pearson Correlation	1	**.473****
	Sig. (2-tailed)		**0.000**
	N	93	93
DDDM	Pearson Correlation	**.473****	1
	Sig. (2-tailed)	**0.000**	
	N	93	93

**. Correlation is significant at the 0.01 level (2-tailed).

5.11 Research Hypothesis 2

The second hypothesis stated that there is a positive relationship between Big Data Analytics Management Capabilities and Data-Driven Decision-Making. This section provides the results related to this hypothesis. Table 18 provides and overview of the descriptive statistics for Big Data Analytics Management Capabilities and Data-Driven Decision-Making. From this data, it is noticed that Big Data Analytics Management Capabilities has a mean of 4.2 which is slightly above the "Neutral" response.

Table 18: Descriptive Statistics for BDAMC and DDDM

	N	Mean	Median	Mode	Std. Deviation
BDAMC	93	4.20	4.46	2.91	1.17
DDDM	93	4.94	5.00	4.60	1.17

Once again, due to the Shapiro-Wilk test showing that the Big Data Analytics Management Capabilities data is non-normal, both the Pearson and Spearman's correlation tests were conducted. The results of these are shown in Table 19 and Table 20 respectively. The results for the Pearson and Spearman's correlation tests show similar results of a significant, large, positive correlation between Big Data Analytics Management Capabilities and Data-Driven Decision-Making.

Table 19: Pearson Correlation between BDAMC and DDDM

		BDAMC	DDDM
BDAMC	Pearson Correlation	1	**.612****
	Sig. (2-tailed)		**0.000**
	N	93	93
DDDM	Pearson Correlation	**.612****	1
	Sig. (2-tailed)	**0.000**	
	N	93	93

**. Correlation is significant at the 0.01 level (2-tailed).

The correlation results are given as r(91)=.61, p < .001 and rs(91)=.64, p < .001 for the relationship between Big Data Analytics Management Capabilities and Data-Driven Decision-Making.

Table 20: Spearman Correlation between BDAMC and DDDM

			BDAMC	DDDM
Spearman's rho	BDAMC	Correlation Coefficient	1.000	**.639****
		Sig. (2-tailed)		**0.000**
		N	93	93
	DDDM	Correlation Coefficient	**.639****	1.000
		Sig. (2-tailed)	**0.000**	
		N	93	93

**. Correlation is significant at the 0.01 level (2-tailed).

5.12 Research Hypothesis 3

The third hypothesis stated that Big Data Analytics Personnel Expertise and Data-Driven Decision-Making have a positive relationship. Table 21 provides a summary of the descriptive statistics for Big Data Analytics Personnel Expertise and Data-Driven Decision-Making which shows that Big Data Analytics Personnel Expertise has a mean of 5.28 while Data-Driven Decision-Making has a mean of 4.94. This provides an indication that there is a positive view of both of these constructs amongst respondents.

Table 21: Descriptive Statistics for BDAPE and DDDM

	N	Mean	Median	Mode	Std. Deviation
BDAPE	93	5.28	5.38	5.69	0.77
DDDM	93	4.94	5.00	4.60	1.17

The Shapiro-Wilk test for normality (Table 13) showed that Big Data Analytics Personnel Expertise did not have a normal distribution and thus Pearson and Spearman's correlation tests were conducted. The results of these tests are shown in TablesTable 22 and Table 23. The output from the Pearson correlation shows a significant positive relationship between Big Data Analytics Personnel Expertise and Data-Driven Decision-Making with $r(91) = .46$, $p < .001$

Table 22: Pearson Correlation between BDAPE and DDDM

		BDAPE	DDDM
BDAPE	Pearson Correlation	1	**.457****
	Sig. (2-tailed)		**0.000**
	N	93	93
DDDM	Pearson Correlation	**.457****	1
	Sig. (2-tailed)	**0.000**	
	N	93	93
**. Correlation is significant at the 0.01 level (2-tailed).			

The Spearman's correlation results show that there is a significant, positive relationship between Big Data Analytics Personnel Expertise and Data-Driven Decision-Making with rs(91)=.45, $p < .001$. These results are similar to those obtained from the Pearson's correlation.

Table 23: Spearman Correlation between BDAPE and DDDM

			BDAPE	DDDM
Spearman's rho	BDAPE	Correlation Coefficient	1.000	**.448**
		Sig. (2-tailed)		**0.000**
		N	93	93
	DDDM	Correlation Coefficient	.448	1.000
		Sig. (2-tailed)	0.000	
		N	93	93

5.13 Proposition

The proposition described in Section 260 states that Big Data Analytics Capabilities predict Data-Driven Decision-Making. This was tested through applying multiple regression analysis between the Big Data Analytics Capabilities second order constructs (Big Data Analytics Infrastructure Flexibility, Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise) and Data-Driven Decision-Making.

The results of the multiple regression are summarised in Table 24 and Table 25 with R^2= .429, $F(3,89)=24.08$, $p < .001$. The adjusted R-Square value of 0.429 provides insight that 42.9% of the variance in Data-Driven Decision-Making is explained by the model. The output of the ANOVA provided in Appendix D.8

Table 24: Regression Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.669[a]	0.448	0.429	0.88417

The regression model for Data-Driven Decision-Making (DDDM), based on the results summary provided in Table 25, is:

DDDM = 0.365 + 0.068(BDAIF) + 0.473(BDAMC) + 0.277(BDAPE).

However, it is noted that the constant term and Big Data Analytics Infrastructure Flexibility are not statistically significant, unique contributors to the model and thus are removed. The resulting model is:

DDDM = 0.473(BDAMC) + 0.277(BDAPE).

This indicates that Big Data Analytics Management Capabilities is the strongest contributor to the model ($\beta=.47$, $p < .001$) followed by Big Data Analytics Personnel Expertise ($\beta=.28$, $p < .001$).

Table 25: Regression Coefficients

	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	Sig.
(Constant)	0.365	0.678		0.538	0.592
BDAIF	0.085	0.135	0.068	0.633	0.528
BDAMC	0.475	0.108	**0.473**	4.394	**0.000**
BDAPE	0.422	0.129	**0.277**	3.278	**0.001**

Thus, the proposition is upheld with Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise being the significant contributors to the model and thus predictors of Data-Driven Decision-Making.

This chapter presented the results relating to the research hypotheses and the proposition defined in Chapter 3. This was done through an initial description of the data preparation process and the sample which was then followed by validity, reliability and assumption checks and finally the results of the tests conducted in order to address the hypotheses and proposition. The next chapter will provide a more detailed discussion of the results as they relate to the literature.

6 Discussion of Results

The objective of this research was to investigate the relationship of Big Data Analytics Capabilities (comprising Big Data Analytics Infrastructure Flexibility, Big Data Analytics Personnel Expertise and Big Data Analytics Management Capability) and Data-Driven Decision-Making as depicted in Figure 6. Understanding these relationships will allow for business leaders to make informed decisions regarding how they deal with Big Data Analytics Capabilities and Data-Driven Decision-Making in their organisations as well as provide a building block upon which further research could be based. This was undertaken through the testing of the hypotheses and proposition, described in Chapter 3, through the use of the methodology presented in Chapter 4. Chapter 5 outlined the results of the statistical tests that were conducted in order to test the hypotheses and proposition.

This chapter presents a discussion of the results from the various hypotheses reported in Chapter 5. The results of each of the three hypotheses will be discussed separately followed by a discussion of the proposition's results.

Research has revealed that supporting decision-making through the use of Big Data Analytics is a means of generating value for organisations (Günther et al., 2017; Sheng et al., 2017; Weill & Woerner, 2015). It has also been noted in literature that generating value from Big Data is contingent on Big Data Analytics Capabilities which cover Big Data Analytics Infrastructure Capability, Big Data Analytics Personnel Expertise and Big Data Analytics Management Capability as described in Section 2.5 (Wamba et al., 2017). Thus, the contribution of this study is providing a description of the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making. This study therefore heeds the call for further research (empirical studies in particular) in the fields of Big Data, Big Data Analytics relating to decision-making (Janssen et al., 2017; Ji-fan Ren et al., 2017). In order to address the objective of this study, the research hypotheses were structured to test relationships between the second order capabilities (Big Data Analytics Infrastructure Capability, Big Data Analytics Personnel Expertise and Big Data Analytics Management Capability) and Data-Driven Decision-Making. Further to this the research proposition tested the relationship of Big Data Analytics Capabilities as predictors of Data-Driven Decision-Making through the application of multiple regression analysis.

6.1 Discussion of Hypothesis 1

The first hypothesis in Section 3.2 was defined as:

H1: Big Data Analytics Infrastructure Flexibility is positively related to Data-Driven Decision-Making.

Recalling that a seven-point Likert scale was used in the questionnaire, the responses to the Big Data Analytics Infrastructure Flexibility construct questions were slightly above “Neutral” ($M = 4.14, SD = 0.93$). The mean can be used as an acceptable measure of central tendency since, from Section 5.6, Big Data Analytics Infrastructure Flexibility was shown to have a normal distribution from both the Shapiro-Wilk and Kolmogorov-Smirnov tests for normality. The mean indicates that, on average, respondents to the questionnaire were neither positive nor negative regarding their Big Data Analytics Infrastructure Flexibility. Taking a further look into the first order constructs comprising Big Data Analytics Infrastructure Flexibility reveals that connectivity, compatibility and modularity have means of 3.88, 4.13 and 4.29 respectively. These values provide an indication of the respondents’ impression of these constructs in their work environment and further insights could be gleaned. Although the mean values for the first order constructs of Big Data Analytics Infrastructure Flexibility are in the region of four or Neutral, insight can be gained about each of the constructs by analysing the individual constructs. The results obtained for Big Data Analytics Infrastructure Flexibility compatibility and modularity, in comparison with the results obtained by Akter et al. (2016), are not considerably different with the means for Big Data Analytics Infrastructure Flexibility compatibility and Big Data Analytics Infrastructure Flexibility modularity obtained by Akter et al. (2016) being 4.54 and 4.47. However, the mean of 3.88 for Big Data Analytics Infrastructure Flexibility connectivity falls within a different range than the 4.53 obtained by Akter et al. (2016) which warrants further analysis.

On examination, it is noticed that connectivity has the lowest mean of the Big Data Analytics Infrastructure Flexibility construct which, at 3.88, falls below Neutral. This shows, overall, respondents are slightly toward the negative region of the Likert scale for this first order construct. An important consideration that the Big Data Analytics Infrastructure Flexibility connectivity construct deals with is having access to leading analytics systems and the ability to share analytics insights. However, with the mean score of Big Data Analytics Infrastructure Flexibility connectivity being below Neutral this provides some indication of respondents having a slightly negative experience in this area. On further analysis it is noticed that, at the question level, of the three questions relating to Big Data Analytics Infrastructure Flexibility connectivity question CN3 had a mean of 3.39 which is closer to “Moderately disagree” while questions CN1 and CN2 had means of 4.31 and 3.95 respectively which indicates that CN3 contributed to reducing the mean of Big Data Analytics Infrastructure Flexibility.

Question CN3 states “There are no identifiable communications bottlenecks within my work environment for sharing analytics insights.” (Wamba et al., 2017, p. 360). It is noted that the mean for question CN3 is the lowest of all questions in the questionnaire which increases the interest into what this question entails. Since the mean value for this question is 3.39 which tends toward “Moderately disagree”, it suggests that, in general, respondents feel that there are some communication bottlenecks relating to the sharing of analytics insights. Kowalczyk and Buxmann (2015) have highlighted this challenge of communication impediments in organisations by way of tensions between analytics professionals and decision makers. They further provide some guidance in possible measures to improve the communication challenge by means of, what they term as, organisational tactics and communication tactics. In addition to this Bumblauskas et al. (2017) have noted that translating data and analytics into action requires having valid information in a timely manner which is affected if there are bottlenecks in sharing information and analytics insights.

The Big Data Analytics Infrastructure Flexibility compatibility construct with a mean of 4.13 being slightly above Neutral provides little insight. However, by not being more negative it could be inferred that, in general, respondents did not feel that the Big Data Analytics Infrastructure had compatibility issues. As mentioned by Ramanathan et al. (2017) infrastructure incompatibility is a hindrance to analytics.

In the Big Data Analytics Infrastructure Flexibility modularity construct, questions dealing with reusability of software components or modules in building new systems and applications (MOD1 and MOD3) had mean values of 4.42 and 4.58 respectively. This puts them in the range between Neutral and Moderately agree which provides an impression that modularity in the sense of reusable software components in BDA infrastructure is leaning towards being somewhat positive.

The Data-Driven Decision-Making responses, however, had a higher mean at 4.94 (*M = 4.94, SD = 1.17*) providing an indication that respondents had a more positive inclination toward Data-Driven Decision-Making in their environment. The mean of 4.94 puts the responses to Data-Driven Decision-Making at marginally below “Somewhat agree”. Question item DDDM6 which states “In my environment, we consider data a tangible asset." had the highest mean of 5.32. The limitation of self-reporting and due to the sample comprising managers and analytics professionals could have influenced the responses as becoming more data-driven is, at a minimum, an aspiration (Vidgen et al., 2017).

The results of the correlation between Big Data Analytics Infrastructure Flexibility and Data-Driven Decision making showed a significant, medium, positive correlation between the two variables with $r(91) = .47$, $p < .001$. This indicates that there is a positive relationship between Big Data Analytics Infrastructure Flexibility and Data-Driven Decision-Making meaning that an increase in Big Data Analytics Infrastructure Flexibility coincides with an increase in Data-Driven Decision-Making. In other words, in situations where Big Data Analytics Infrastructure Flexibility exists respondents are likely to undertake Data-Driven Decisions. It is noted that the correlation does not give any indication of causality or prediction but provides an indication that as an initial step the two variables are positively, moderately related i.e. they vary or change in the same direction moderately.

The findings in this hypothesis of a significant, positive correlation between Big Data Analytics Infrastructure Flexibility and Data-Driven Decision-Making are consistent with literature where Kache and Seuring (2017) mention that through providing the ability of processing data and making information available to decision makers, Big Data Analytics Infrastructure has an influence on Data-Driven Decision-Making. This was alluded to in Section 3.2. An additional congruence between the findings of this hypothesis and literature is where Rejikumar et al. (2018) found that managers' intention to adopt Data-Driven Decision-Making was influenced by technology and infrastructure as mentioned in Sections 2.5 and 2.6.

6.2 Discussion of Hypothesis 2

The second hypothesis, as stated in Section 3.2, is:

H2: Big Data Analytics Management Capabilities and Data-Driven Decision-Making are positively related.

The descriptive statistics for hypothesis 2 provided in Table 18 shows that with Big Data Analytics Management Capabilities having a mean of 4.20 (*M = 4.20, SD = 1.17*), the average response was slightly more positive than Neutral. As described in Section 2.5 Big Data Analytics Management Capabilities has first order constructs that deal with planning, coordination, control and investment decision making capabilities relating to Big Data Analytics (Akter et al., 2016; Wamba et al., 2017). The means of Big Data Analytics Management Capabilities planning, Big Data Analytics Management Capabilities coordination, Big Data Analytics Management Capabilities control and Big Data Analytics Management Capabilities investment decision making are 4.49, 4.16, 4.13 and 4.02 respectively. Akter et al. (2016) reported results in a similar range although somewhat more positive with means of 4.89, 4.60 and 4,58 for Big Data Analytics Management Capabilities planning, coordination and control, respectively. However, Big Data Analytics Management Capabilities investment decision making result Akter et al. (2016) differed by comparatively more with a mean of 4.85. This indicates that the respondents in this study were more neutral toward the investment decision making relating to big data analytics in their environments.

Big Data Analytics Management Capabilities planning with a mean of 4.49 showed that respondents tended to be more positive relating to the Big Data Analytics planning or strategic activities in their environments than the other first order constructs of BDAC. Upon further examination of the results it the question PLAN1 which states "We continuously examine innovative opportunities for the strategic use of business analytics." had a mean of 4.98 (Wamba et al., 2017, p. 360). This is the highest mean value from all the questions in Big Data Analytics Management Capabilities planning and, to extend further, from all other Big Data Analytics Management Capabilities first order constructs as well. This provides an insight that, on average, respondents felt that they search for strategic application of business analytics in an ongoing manner. The importance of a strategic definition of how and where the application of analytics will create value is highlighted by Vidgen et al. (2017). Additionally, Ransbotham et al. (2015b) assert that analytics strategies need to continuously progress and develop, even in analytically mature environments or organisations, which reinforces the importance of planning and the concept of continuous improvement and exploration. This is vital in the continuously evolving and developing technologically world

The Big Data Analytics Management Capabilities coordination first order construct explored the dimension of cross functional and cross departmental coordination and collaboration. The mean value of 4.13 indicates that respondents were for the most part Neutral about coordination and collaboration in their environments. Wamba et al. (2015), in the context of emergency services, highlighted that availability and access to information is not sufficient in extreme events and requires the collaboration and coordination of various stakeholders for effective management of the situation. This view is echoed by Kowalczyk and Buxmann (2015) who further elucidate that the different stakeholders have different contributions to make such as the analytics skills from the analysts and domain knowledge from the decision maker. This coordination contributes to effective data-driven decision-making.

The Big Data Analytics Management Capabilities investment decision making construct explores the extent to which the consequences of an investment are considered when making investment decisions relating to business analytics. The second Big Data Analytics Management Capabilities investment decision making question (IDM2), which states “When we make business analytics investment decisions, we project how much these options will help end users make quicker decisions.” had the highest mean value of 4.55 (Wamba et al., 2017, p. 360). This result highlights that, on average, respondents are somewhat inclining to “Moderately Agree”. This is of interest and ties in with Sheng et al. (2017) where they state that investment in Information Technology (IT) and data analytics skills would support timely decisions by executives through the (data analytics) techniques to analyse the Big Data. This is also corroborated by Sharma et al. (2014) who posit that the value realised from investment in analytics is realised through enhanced decision-making which is enabled by analytics.

The Big Data Analytics Management Capabilities control construct dealt with aspects of Big Data Analytics such as performance management of the analytics function as well as operationalising analytics in business processes. Côrte-Real et al. (2017) describe that business processes are one of the stages in which Big Data Analytics can create value which provides some insight into the substance of Big Data Analytics Management Capabilities control as a construct. The mean value of 4.02 for Big Data Analytics Management Capabilities control indicates that respondents, in general, were Neutral regarding the operationalising and performance management of the analytics function in their environment.

To test the hypothesis that Big Data Analytics Management Capabilities and Data-Driven Decision-Making are positively related, correlation analysis was conducted. As mentioned in Section 5.11. The nature of the data showed that it was not normally distributed which lead to both a Pearson and Spearman’s correlation being conducted. The results from these revealed a significant, large and positive relationship between Big Data Analytics Management Capabilities and Data-Driven Decision-Making with $r(91) = .61$, $p < .001$ and $r_s(91) = .64$, $p < .001$. In other words, this gives an indication that, for example, when a high level of Big Data Analytics Management Capabilities is present then respondents are likely to undertake Data-Driven Decisions. Compared to the results of the correlation analysis from hypothesis 1 it is noted that there is a larger correlation between Big Data Analytics Management Capabilities and Data-Driven Decision-Making than Big Data Analytics Infrastructure Flexibility and Data-Driven Decision-Making which provides an indication that Big Data Analytics Management Capabilities and Data-Driven Decision-Making have a stronger association than Big Data Analytics Infrastructure Flexibility and Data-Driven Decision-Making.

Therefore, as highlighted by Kache and Seuring (2017) the important opportunities provided by the first order constructs of Big Data Analytics Management Capabilities which include coordination and control, referred to as integration and collaboration and governance and compliance by Kache and Seuring (2017), are significantly, positively related to Data-Driven Decision-Making. This provides an indication of some of the factors that need to be considered when developing management capabilities related to Big Data Analytics.

6.3 Discussion of Hypothesis 3

The third hypothesis to be tested, as outlined in Section 3.2, is as follows:

H3: Big Data Analytics Personnel Expertise and Data-Driven Decision-Making are positively related.

Table 21 provides the descriptive statistics for Big Data Analytics Personnel Expertise and Data-Driven Decision-Making. In contrast to Big Data Analytics Infrastructure Flexibility and Big Data Analytics Management Capabilities, Big Data Analytics Personnel Expertise has a mean of 5.28 ($M = 5.28, SD = 0.77$) which is more than one Likert Scale point above and translates to being slightly more positive than “Moderately Agree”. The insight that this provides is that respondents, in general, tend to have positive inclination of their expertise relating to Big Data Analytics.

The Big Data Analytics Personnel Expertise second order construct comprises technical knowledge, technological management knowledge, business knowledge and relational knowledge as first order constructs which had means of 4.65, 5.57, 5.18 and 5.73 respectively. Therefore respondents, in general, had the most positive inclination of 5.73 for relational knowledge and the lowest for technical knowledge with a mean of 4.65. In comparison, means of 4.88, 4.85, 4.96 and 4.85 were obtained for technical knowledge, technological management knowledge, business knowledge and relational knowledge by Akter et al. (2016). Therefore, with the exception of technical knowledge the results obtained for the first order constructs in this study were higher than those obtained by Akter et al. (2016). The difference being particularly substantial in technological management knowledge and relational knowledge from which a general insight may be gleaned which is that surveying capabilities and expertise may be useful for organisations in gauging the expertise within the organisation. Grover et al. (2018) underscore that expertise, domain-specific business knowledge and analytics skills and knowledge, amongst others, are essential and need to be continuously reviewed or assessed. This is further substantiated by Ransbotham et al. (2015b) who highlight the importance of planning for analytics talent by hiring talent and developing expertise through effective use of human resource practises.

The Big Data Analytics Personnel Expertise relational knowledge construct deals with the capability of team work and relationships with people such as customers or colleagues. The result obtained of 5.73 for this construct provides an indication that the respondents tend to, in general, have a positive view of their ability to conduct teamwork, teach others and maintain healthy customer relationships (Akter et al., 2016). Investigating this construct further reveals that respondents, in general, related most positively with question items RK2 and RK3 which state “I am very capable in terms of executing work in a collective environment.” and “I am very capable in terms of teaching others.” respectively (Wamba et al., 2017, p. 360).

Technical skills and expertise relating to Big Data Analytics have been noted to be a challenge in organisations (Vidgen et al., 2017). The result of the Big Data Analytics Personnel Expertise technical knowledge construct at 4.65 being the lowest of all the personnel expertise provides an indication that, in general, respondents were more positive about other expertise (such as relational and business knowledge) than the technical knowledge. However, it should be noted, that at 4.65 the inclination of respondents relating to technical knowledge is tending towards Moderately Agree.

The Big Data Analytics Personnel Expertise business knowledge construct with a mean of 5.18 provides an indication that respondents are somewhat positive regarding the business knowledge relating to Big Data Analytics in their environment. Upon further analysis, it is noticed that the question BK2 which relates to the capability of interpreting business problems and the development of suitable solutions, had a somewhat lower mean than other questions at 4.72. This provides an interesting insight since literature alludes to the gap of employees being able to solve business problems or answer business questions through the application of analytics (Kowalczyk & Buxmann, 2015; Ransbotham et al., 2015a). This challenge could be solved through the recommendation by Kowalczyk and Buxmann (2015) of closer collaboration or through the closing of the gap in analytics skills present in management asserted by Ransbotham et al. (2015a).

The Big Data Analytics Personnel Expertise technological management knowledge construct explored the ability to understand technological trends, learn new technologies and the view of analytics being a means to an end (Akter et al., 2016; Wamba et al., 2017). The Big Data Analytics Personnel Expertise technological management knowledge construct had a mean of 5.57 which indicates that respondents in general tended have a moderately positive view of their technological management knowledge.

The direct testing of the hypothesis was done by means of correlation analysis. As alluded to in Section 5.12, the results from the correlations, Pearson and Spearman's rho, of $r(91) = .46$, $p < .001$ and $rs(91) = .45$, $p < .001$ respectively, indicates that Big Data Analytics Personnel Expertise and Data-Driven Decision-Making have a significant, medium, positive relationship.

Therefore from the results it is noted that the relationship between Big Data Analytics Personnel Expertise and Data-Driven Decision-Making provides some indication of supporting literature in that there is general need for skills and capabilities, in managers, that need to use analytics for decision-making (Ransbotham et al., 2015b, 2016). Further to this the responses and results are in line with the view where skills and data management capabilities are mentioned to be improvement opportunities in supporting decision-making through the use of Big Data Analytics as was the case with the Big Data Analytics Personnel Expertise business knowledge construct (Kache & Seuring, 2017; Kiron et al., 2014b).

In conclusion to the hypotheses discussion, it is noted that all three second order constructs of BDAC have significant, positive correlations with Data-Driven Decision-Making, all of which are congruent with literature as discussed. Big Data Analytics Personnel Expertise and Big Data Analytics Infrastructure Flexibility have medium correlations with Data-Driven Decision-Making while Big Data Analytics Management Capabilities has a large correlation with Data-Driven Decision-Making. The next section discusses the results of testing the proposition which was done through the means of multiple regression analysis.

6.4 Discussion of Proposition

The proposition to be tested in this study was between Big Data Analytics Capabilities (comprising second order constructs Big Data Analytics Infrastructure Flexibility, Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise) and Data-Driven Decision-Making. The results from testing this proposition through multiple regression analysis between the Big Data Analytics Infrastructure Flexibility, Big Data Analytics Management Capabilities, Big Data Analytics Personnel Expertise (independent variables) and Data-Driven Decision-Making (dependent variable) are key to this study. This is because they provide further insight into the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making which is the purpose of this research, as mentioned in Section 1.5. The multiple regression analysis resulted in an adjusted R-square value of 0.429. This result provides the insight that 42.9% of the variance in Data-Driven Decision-Making is explained by the model. As presented in Section 0, the multiple regression analysis produced a regression model of:

DDDM = 0.365 + 0.068(BDAIF) + 0.473(BDAMC) + 0.277(BDAPE).

This model includes the non-significant contributors to the model which are the constant term and Big Data Analytics Infrastructure Flexibility as shown in Table 25. The regression model provides insight as well as empirical support that Big Data Analytics Personnel Expertise and Big Data Analytics Management Capabilities are significant predictors of Data-Driven Decision-Making. As noted in Section 2.9, empirical support is lacking in the fields of Big Data, Big Data Analytics and Data-Driven Decision-Making as alluded to by Sheng et al. (2017), Sivarajah et al. (2017), Janssen et al. (2017) as well as Sharma et al. (2014).

The model coefficients of the unique significant contributors to the regression model conveys that Big Data Analytics Personnel Expertise has a standardised coefficient of 0.277 and Big Data Analytics Management Capabilities has a standardised coefficient of 0.473. This provides an indication that Big Data Analytics Management Capabilities is a stronger predictor of Data-Driven Decision-Making than Big Data Analytics Personnel Expertise. Furthermore, the model indicates that when looking at increasing Data-Driven Decision-Making in personnel, through Big Data Analytics Capabilities, organisations should look at improving Big Data Analytics Personnel Expertise and Big Data Analytics Management Capabilities. Ideally, both Big Data Analytics Personnel Expertise and Big Data Analytics Management Capabilities should be addressed, however, it is often the case that there are some level of constraints on resources. In this case Big Data Analytics Management Capabilities would be the first priority due to it being a stronger predictor of Data-Driven Decision-Making.

6.5 Conclusion

In conclusion, this chapter provided a discussion of the results obtained in this research, provided in Chapter 5, which aimed to investigate the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making. To review the results, it was noted that all three second order constructs of Big Data Analytics Capabilities i.e. Big Data Analytics Infrastructure Flexibility, Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise had significant, positive correlations (at 99% confidence level) with Data-Driven Decision-Making. Further to the hypotheses, the proposition provided further insight by testing the second order constructs of Big Data Analytics Capabilities as predictors of Data-Driven Decision-Making which revealed Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise as significant predictors of Data-Driven Decision-Making.

Previous studies have recommended conducting surveys and providing empirical support in the fields of Big Data, Big Data Analytics and Data-Driven Decision-Making to substantiate and provide depth to the field which often comprises anecdotal evidence (Dremel et al., 2017). Thus, this study contributes sought after empirical support in the investigation of the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making.

7 Conclusion

This chapter is the culmination of the study and serves to link previous chapters through a discussion of the principal findings, the implications for management, limitations of this study and recommendations for future research.

The principal findings section is related to the research purpose, literature, research hypotheses and proposition which tie back to Chapters 1, 2 and 3. Further to this the practical implications for management, from the research findings (Chapters 5 and 6), are discussed. This discussion on the implications for management relates to Section 1.3.1, which highlighted the business need for study. Additionally, the sections on limitations and future research in this chapter cover the theoretical and methodological limitations of this study which relate to sections discussed in Chapters 2 and 4 respectively and also potential future studies related to this research are provided.

7.1 Principal Findings

This study aimed to investigate the relationship between Big Data Analytics Capabilities (comprising Big Data Analytics Infrastructure Flexibility, Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise) and Data-Driven Decision Making. As noted in Section 2.9, the relationship between Big Data Analytics Capabilities and firm performance has been research by Wamba et al. (2017) and the relationship between Data-Driven Decision-Making and firm performance has also been tested (Brynjolfsson et al., 2011). However, the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making has, to the best of the author's knowledge, not been investigated. Thus, this study complements previous research and contributes by growing the empirical research of Big Data and Big Data Analytics relating to decision-making. This has been noted by Janssen et al. (2017) to be limited. The investigation of the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making was conducted through testing three research hypotheses (using correlation analysis) and a research proposition (using multiple regression analysis).

This study found, through the testing of the research hypotheses, that Big Data Analytics Infrastructure Flexibility, Big Data Analytics Management Capability and Big Data Analytics Personnel Expertise (second order constructs of Big Data Analytics

Capabilities) have medium to large, significant, positive correlations with Data-Driven Decision-Making. Big Data Analytics Infrastructure Flexibility and Big Data Analytics Personnel Expertise have medium correlations ($r(91)=.47$, $p < .001$ and $r(91)=.46$, $p < .001$ respectively) with Data-Driven Decision-Making and Big Data Analytics Management Capability has a large correlation with Data-Driven Decision-Making with $rs(91)=.64$, $p < .001$. These findings provide a contribution to literature which have highlighted the need for empirical findings in the fields of Big Data Analytics as well as Data-Driven Decision-Making by means of providing insight into the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making (Sivarajah et al., 2017; Wamba et al., 2015).

The findings of the hypotheses were further built upon through the testing of the research proposition. The testing of the research proposition through multiple regression analysis revealed that Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise are significant predictors of Data-Driven Decision-Making while Big Data Analytics Infrastructure Flexibility is not. Drawing on an insight from Kiron et al. (2014b), a possible reason for the finding of Big Data Analytics Infrastructure Flexibility not being a significant predictor, is that there could be an increase in access to data but the rate of data management and analytics skills are not improving or growing at the same rate. This suggests that although managers may have access to the data they could still be stifled by not having the necessary skills and expertise to make use of the data, particularly for decision-making.

Once again, the findings from the multiple regression analysis provide empirical support in relating Big Data Analytics Capabilities and Data-Driven Decision-Making which is a contribution sought after in literature (Janssen et al., 2017). Furthermore, the findings from the research proposition as well as the hypotheses contribute to the scholarship in management relating to Big Data which was highlighted by George et al. (2014) as being limited.

The findings from this study are summarised in Figure 14 and closes the loop going back to Figure 1 which formed part of the introduction to the research problem in this study.

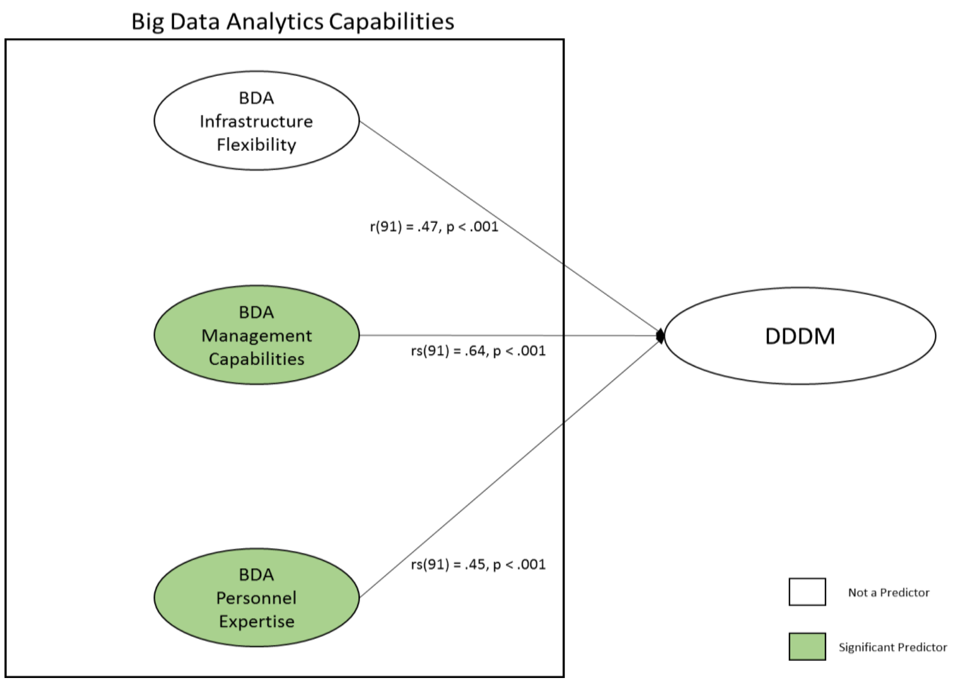


Figure 14: Summary of Findings

The findings from this study which reveal characteristics of the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making aid in addressing gap in research, noted by Sheng et al. (2017) as well as Janssen et al. (2017), which relates to using Big Data and Big Data Analytics to inform decisions. The insight provided by this study indicates that the Big Data Analytics Capabilities of Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise will likely be significant contributors in the promotion of using Big Data Analytics in informing decisions (though Data-Driven Decision-Making).

Ransbotham et al. (2015a) highlighted the need for a combination of analytics skills as well as business knowledge in order to generate value. The findings from this research confirm this, in that Big Data Analytics Personnel Expertise, which comprises analytics and technical skills, as well as Big Data Analytics Management Capabilities, which includes business and relational knowledge, were found to be significant predictors of Data-Driven Decision-Making.

Although this study provides a contribution to literature and academic aspects of Big Data Analytics, Big Data Analytics Capabilities and Data-Driven Decision-Making there are also practical implications from this research on management which could be employed in practice. These implications are discussed in the next section.

7.2 Implications for Management

This section discusses the practical implications of the findings in this study on management while building on the foundations from literature. In a business environment in which Big Data is ubiquitous, it is natural inclination for management to aim to exploit it in order to extract value for the business and possibly gain competitive advantage (Grover et al., 2018; Günther et al., 2017). Data-Driven Decision-Making has been shown to have a positive effect on firm performance by Brynjolfsson and McElheran (2016a) and has been noted to being a means to gaining competitive advantage (Ransbotham et al., 2016). This study aimed to gain an understanding of the relationship between Big-Data Analytics Capabilities and Data-Driven Decision-Making with the view of providing practical insights for management into the possible Big Data Analytics Capabilities levers to pull which are likely to foster Data-Driven Decision-Making.

Relating this to the current study, the findings from the hypotheses were that Big Data Analytics Infrastructure Flexibility, Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise have positive, significant correlations with Data-Driven Decision-Making are that each of these Big Data Analytics Capabilities are positively associated with Data-Driven Decision-Making. The practical insight and implications for management from these results are that an increase each of these constructs is likely to be associated with an increase in Data-Driven Decision-Making. This provides a good starting point for management as Weill and Woerner (2015), highlight that Data-Driven (they refer to evidence-based) Decision-making should be emphasised in organisations.

Furthermore, the finding of Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise being significant predictors of Data-Driven Decision-Making provides powerful insights that can be used by management practitioners. The rest of this section on the management implications of this study will focus on the practical steps that can be taken by management in light of the finding that Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise are significant predictors of Data-Driven Decision-Making. The goal for managers is to enable employees to enhance their intuition and experience based decision-making through the use of data and analytics while fostering a critical thinking mindset (Shah et al., 2012).

To recap, Big Data Analytics Management Capabilities comprises the dimensions of planning, coordination, control and investment decision making (Wamba et al., 2017). Ransbotham et al. (2016) noted the importance of having a Big Data Analytics plan and have found that analytically mature organisations are more likely to have formalised analytics plans and strategies. Thus, management need to be cognisant of whether a cogent analytics plan exists as well as ensuring that it is aligned with the current business processes and strategy. If an analytics plan does not exist then it should be developed as highlighted by Vidgen et al. (2017) as well as Kache and Seuring (2017). Furthermore, in developing and reviewing the analytics plan it is imperative that management ensure that congruence is maintained between the analytics and business strategies. Additionally, an important management implication is that the congruence and alignment should follow through into the implementation of these strategies through the analytics and business processes, which is key to deriving value from Big Data Analytics (Sheng et al., 2017).

Coordination is an important component in Big Data Analytics Management Capabilities in that, as Kowalczyk and Buxmann (2015) allude to, there is a need for analysts and business decision makers to collaborate since they bring differing skillsets together. This is also further supported by Wamba et al. (2015) who highlight the need for coordination between various stakeholders to gain effective management in high pressure emergency situations. This can be extended to business environments which often have high pressure situations and thus the management implication of this is that business leaders need to ensure that there are means for coordination and collaboration between analysts and other business stakeholders. This may require a careful look at the way departments (in the organisation) are structured as well as the defined business processes that may hinder collaboration in which case interventions need to take place. These interventions could take the form of modifying the organisational design and updating the business processes. Furthermore, in order to foster more coordination and collaboration, managers should explore training of analytics personnel (such as data scientists) to effectively work and collaborate with colleagues from business departments (Harris & Mehrotra, 2014).

Big Data Analytics Personnel Expertise, which was found as the second significant predictor of Data-Driven Decision-Making consists of the dimensions of technical knowledge, management of technology, relational knowledge and business knowledge. The first implication for managers from the analysis of Big Data Analytics Personnel Expertise is that management needs to understand the extent to which they have or lack the expertise in their departments or organisations (Kache & Seuring, 2017). Technical expertise is often a challenge in the application of Big Data Analytics and thus the implication for managers is to identify if this is a challenge (in their environment) and grow and develop these skills (Sivarajah et al., 2017). This could be achieved through conscientious recruitment and also through training programs. To effectively employ Big Data Analytics for decision-making there is also a need for a combination of technical expertise as well as business knowledge (Ransbotham et al., 2015a). An important note for managers here is that, as noted by Ransbotham et al. (2016) and Ransbotham et al. (2015a), there is a need for managers to grow their skill sets to include becoming comfortable with analytics. This could mean that as a talent management practise, managers should be empowered with the skills to be able to use Big Data Analytics and hence use it as a basis for decision-making. This would also be effective in fostering an analytics and data-driven culture (Kiron et al., 2014b).

In conclusion, the findings from this study have several practical implications for management to move their organisations towards improving Big Data Analytics Capabilities as well as encouraging Data-Driven Decision-Making.

7.3 Limitations

This section describes the limitations of the study which include aspects relating to the research design choices, practical implementation as well as the theoretical boundary to which this study extends.

The first limitation that is noted is that of the sampling technique used in the study. The judgemental and snowballing sampling technique employed in this study is a limitation in that it was noticed that two industries accounted for 50% of all responses received. This may be indicative of the nature of application of Big Data Analytics amongst various industries or could be attributed to the non-random sampling methods employed. Therefore the limitation of the snowball sampling technique is that biases may be introduced (Zikmund et al., 2010).

The second limitation noted is that of a self-administered questionnaire as the measurement instrument utilised in this study. Since the responses are provided by the respondents themselves, this could have introduced some biases into the responses to the questions. An example of this being the use of the seven-point scale, which could have introduced the acquiescence bias (Akter et al., 2016).

The third limitation noted in this study is that of the sample size obtained. Although several means were adopted with the aim of attaining a large sample the total number of usable responses in the sample was 93. Sample size is a factor in the type of statistical analysis that is recommended and also has an effect on the results from statistical analysis (Zikmund et al., 2010). An example of this is that sample size impacts the results and applicability of CFA as was noticed in this study. Furthermore, it is recommended that a minimum sample size of 200 be obtained to undertake Structured Equation Modelling (SEM) which although was not part of the analysis methods adopted, precluded its consideration for analysis post data collection (Hair et al., 2010).

This study was limited in scope to Big Data Analytics Capabilities comprising the three second order constructs (Big Data Analytics Infrastructure Flexibility, Big Data Analytics Management Capabilities and Big Data Analytics Personnel Expertise) and Data-Driven Decision-Making. This limitation has the aspect of not taking into account other factors which could include Big Data Analytics Information Quality or Big Data Analytics System Quality (Ji-fan Ren et al., 2017). Related to this limitation is that this study was based on a the Big Data Analytics Capabilities construct from literature, explored by Wamba et al. (2017) and Akter et al. (2016). The limitation in this is that Big Data Analytics Capabilities could incorporate different or additional constructs and also has been noted to potentially have contextual dependencies (Wamba et al., 2017).

Finally, it is noted that the researcher conducting this study, although having conducted technical research previously, has had limited previous experience in business research which could have introduced some inaccuracies where judgement was required during the course of the study.

7.4 Future Research

This study focussed on investigating the relationship between Big Data Analytics Capabilities and Data-Driven Decision-Making and provides a suitable foundation of empirical support upon which further research could be explored. This section provides a description of some possibilities for future research which could be considered.

Future studies could include additional constructs related to Big Data Analytics which include but are not limited to Big Data Analytics System Quality, Big Data Analytics Information Quality, organisation culture and leadership style (Ji-fan Ren et al., 2017). In addition to this, future research could include testing of the mediating or moderating effect of Data-Driven Decision-Making on the relationship between Big Data Analytics

Capabilities and firm performance. Furthermore, organisational culture could be a factor that could be included in the study (Kiron et al., 2014b).

Big Data and predictive analytics factors such as assimilation and routinisation have been tested in relation to organisational performance (Gunasekaran et al., 2017). These factors could be considered in future research relating to Data-Driven Decision-Making as well as firm or organisational performance.

Further to this, future research could look to building upon the sample size limitation in this study and undertake a Structured Equation Modelling (SEM) approach in a similar study to provide further sought after empirical support in this field. Additionally, studies which look into the contextual factors of organisations and their potential effect on Big Data Analytics, Big Data Analytics Capabilities as well as Data-Driven Decision-Making could be considered.

This study was conducted as a cross sectional study and thus conducted at a point in time. A longitudinal study would be able to provide insight into the possible changes of the relationships between the variables over time.

8 References

Adair, J. E. (2013). The art of effective decision making. In *Decision Making and Problem Solving* (2nd ed.). London: Kogan Page Ltd.

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment? *International Journal of Production Economics*, *182*, 113–131. https://doi.org/10.1016/j.ijpe.2016.08.018

Arunachalam, D., & Kumar, N. (2018). Benefit-based consumer segmentation and performance evaluation of clustering approaches: An evidence of data-driven decision-making. *Expert Systems with Applications*, *0*, 1–24. https://doi.org/10.1016/j.eswa.2018.03.007

Ashby, N. J. S. (2017). Numeracy predicts preference consistency: Deliberative search heuristics increase choice consistency for choices from description and experience. *Judgment and Decision Making*, *12*(2), 128–139. Retrieved from http://www.sjdm.org/journal/16/16912/jdm16912.pdf

Ayat, N., Akbarinia, R., Afsarmanesh, H., & Valduriez, P. (2014). Entity resolution for probabilistic data. *Information Sciences*, *277*, 492–511. https://doi.org/10.1016/j.ins.2014.02.135

Azar, O. H. (2014). The default heuristic in strategic decision making: When is it optimal to choose the default without investing in information search? *Journal of Business Research*, *67*(8), 1744–1748. https://doi.org/10.1016/j.jbusres.2014.02.021

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., Skolits, G. J., & Esquivel, S. L. (2013). Practical Considerations for Using Exploratory Factor Analysis in Educational Research. *Practical Assessment, Research & Evaluation*, *18*(6). Retrieved from http://pareonline.net/getvn.asp?v=18&n=6

Bernroider, E. W. N., & Schmöllerl, P. (2013). A technological, organisational, and environmental analysis of decision making methodologies and satisfaction in the context of IT induced business transformations. *European Journal of Operational Research*, *224*(1), 141–153. https://doi.org/10.1016/j.ejor.2012.07.025

Borgonovo, E., Cappelli, V., Maccheroni, F., & Marinacci, M. (2018). Risk analysis and decision theory: A bridge. *European Journal of Operational Research*, *264*(1), 280–293. https://doi.org/10.1016/j.ejor.2017.06.059

Borgonovo, E., & Marinacci, M. (2015). Decision analysis under ambiguity. *European Journal of Operational Research*, *244*(3), 823–836. https://doi.org/10.1016/j.ejor.2015.02.001

Borrero, S., & Henao, F. (2017). Can managers be really objective? Bias in multicriteria decision analysis. *Academy of Strategic Management Journal*, *16*(1), 244–260.

Braun, M. T., Kuljanin, G., & DeShon, R. P. (2018). Special considerations for the acquisition and wrangling of big data. *Organizational Research Methods*, *21*(3), 633–659. https://doi.org/10.1177/1094428117690235

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strength in numbers: How does data-driven decision making affect firm performance? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1819486

Brynjolfsson, E., & McElheran, K. (2016a). Data in action: Data-Driven decision making in U.S. manufacturing. *CES 16-06*, (January), 1–55.

Brynjolfsson, E., & McElheran, K. (2016b). The rapid adoption of data-driven decision-making. *American Economic Review*, *106*(5), 133–39. Retrieved from http://dx.doi.org/10.1257/aer.p20161016

Bumblauskas, D., Nold, H., Bumblauskas, P., & Igou, A. (2017). Big data analytics: transforming data to action. *Business Process Management Journal*, *23*(3), 703–720. https://doi.org/10.1108/BPMJ-03-2016-0056

Cao, G., Duan, Y., & Li, G. (2015). Linking business analytics to decision making effectiveness: A path model analysis. *IEEE Transactions on Engineering Management*, *62*(3), 384–395. https://doi.org/10.1109/TEM.2015.2441875

Chavez, R., Yu, W., Jacobs, M. A., & Feng, M. (2017). Data-driven supply chains, manufacturing capability and customer satisfaction. *Production Planning and Control*, *28*(11–12), 906–918. https://doi.org/10.1080/09537287.2017.1336788

Chen, D. Q., Preston, D. S., & Swink, M. (2015). How the use of big data analytics affects value creation in supply chain management. *Journal of Management Information Systems*, *32*(4), 4–39. https://doi.org/10.1080/07421222.2015.1138364

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of big data analytics in European firms. *Journal of Business Research*, *70*, 379–390. https://doi.org/10.1016/j.jbusres.2016.08.011

Donovan, S. J., Güss, C. D., & Naslund, D. (2015). Improving dynamic decision making through training and self-reflection. *Judgment and Decision Making*, *10*(4), 284–295.

Dremel, C., Herterich, M. M., Wulf, J., Waizmann, J.-C., & Brenner, W. (2017). How Audi AG established big data analytics in its digital transformation. *MIS Quarterly Executive*, *16*(2), 81–100.

Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big data consumer analytics and the transformation of marketing. *Journal of Business Research*, *69*(2), 897–904. https://doi.org/10.1016/j.jbusres.2015.07.001

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage Publications.

Frisk, J. E., & Bannister, F. (2017). Improving the use of analytics and big data by changing the decision-making culture: A design approach. *Management Decision*. https://doi.org/10.1108/MD-07-2016-0460

Frisk, J. E., Lindgren, R., & Mathiassen, L. (2014). Design matters for decision makers: Discovering IT investment alternatives. *European Journal of Information Systems*, *23*(4), 442–461. https://doi.org/10.1057/ejis.2013.13

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, *35*(2), 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

George, G., Haas, M. R., & Pentland, A. (2014). Big data and management. *Academy of Management Journal*, *57*(2), 321–326. https://doi.org/10.1111/risa.12257

George, G., Osinga, E. C., Lavie, D., & Scott, B. A. (2016). Big data and data science methods for management research. *Academy of Management Journal*, *59*(5), 1493–1507. https://doi.org/10.5465/amj.2016.4005

Goes, P. B. (2014). Big Data and IS research. *MIS Quarterly*, *38*(3).

Grover, V., Chiang, R. H. L., Liang, T.-P., & Zhang, D. (2018). Creating strategic business value from big data analytics: A research framework. *Journal of Management Information Systems*, *35*(2), 388–423. https://doi.org/10.1080/07421222.2018.1451951

Gunasekaran, A., Papadopoulos, T., Dubey, R., Wamba, S. F., Childe, S. J., Hazen, B., & Akter, S. (2017). Big data and predictive analytics for supply chain and organizational performance. *Journal of Business Research*, *70*, 308–317. https://doi.org/10.1016/j.jbusres.2016.08.004

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information Systems*, *26*(3), 191–209. https://doi.org/10.1016/j.jsis.2017.07.003

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. *Information and Management*, *53*(8), 1049–1064. https://doi.org/10.1016/j.im.2016.07.004

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.

Harris, J. G., & Mehrotra, V. (2014). Getting value from your data scientists. *MIT Sloan Management Review*, *56*(1), 15–18. Retrieved from http://sloanreview.mit.edu/article/getting-value-from-your-data-scientists/

Horita, F. E. A., de Albuquerque, J. P., Marchezini, V., & Mendiondo, E. M. (2017). Bridging the gap between decision-making and emerging big data sources: An application of a model-based framework to disaster management in Brazil. *Decision Support Systems*, *97*, 12–22. https://doi.org/10.1016/j.dss.2017.03.001

Intezari, A., & Pauleen, D. J. (2017). Conceptualizing wise management decision-making: A grounded theory approach. *Decision Sciences*, *49*(2), 335–400. https://doi.org/10.1111/deci.12267

Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, *70*, 338–345. https://doi.org/10.1016/j.jbusres.2016.08.007

Ji-fan Ren, S., Fosso Wamba, S., Akter, S., Dubey, R., & Childe, S. J. (2017). Modelling quality dynamics, business value and firm performance in a big data analytics environment. *International Journal of Production Research*, *55*(17), 5011–5026. https://doi.org/10.1080/00207543.2016.1154209

Kache, F., & Seuring, S. (2017). Challenges and opportunities of digital information at the intersection of big data analytics and supply chain management. *International Journal of Operations & Production Management*, *37*(1), 10–36.

https://doi.org/10.1108/IJOPM-02-2015-0078

Kinnear, B. P. R., & Gray, C. D. (2012). *IBM SPSS statistics 19 made simple*. Hove and New York: Psychology Press. https://doi.org/10.1111/j.1751-5823.2012.00187_13.x

Kiron, D., Prentice, P. K., & Ferguson, R. B. (2014a). Raising the bar with analytics. *MIT Sloan Management Review*, *55*(2), 28–34.

Kiron, D., Prentice, P. K., & Ferguson, R. B. (2014b). The analytics mandate. *MIT Sloan Management Review*, *55*(4), 1.

Kock, A., & Gemünden, H. G. (2016). Antecedents to decision-making quality and agility in innovation portfolio management. *Journal of Product Innovation Management*, *33*(6), 670–686. https://doi.org/10.1111/jpim.12336

Kowalczyk, M., & Buxmann, P. (2015). An ambidextrous perspective on business intelligence and analytics support in decision processes: Insights from a multiple case study. *Decision Support Systems*, *80*, 1–13. https://doi.org/10.1016/j.dss.2015.08.010

Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management*, *34*(3), 387–394. https://doi.org/10.1016/j.ijinfomgt.2014.02.002

Lantz, B. (2013). Equidistance of Likert-type scales and validation of inferential methods using experiments and simulations. *Electronic Journal of Business Research Methods*, *11*(1), 16–28. https://doi.org/10.1111/j.1365-2929.2004.02012.x

Liberatore, M. J., Pollack-Johnson, B., & Clain, S. H. (2017). Analytics capabilities and the decision to invest in analytics. *Journal of Computer Information Systems*, *57*(4), 364–373. https://doi.org/10.1080/08874417.2016.1232995

Marcus, B., Weigelt, O., Hergert, J., Gurt, J., & Gelléri, P. (2017). The use of snowball sampling for multi source organizational research: Some cause for concern. *Personnel Psychology*, *70*(3), 635–673. https://doi.org/10.1111/peps.12169

Martin, R. L., & Golsby-Smith, T. (2017). Management is much more than science: The limits of data-driven decision making. *Harvard Business Review*, (Sept-Oct), 129–135.

Matthias, O., Fouweather, I., Gregory, I., & Vernon, A. (2017). Making sense of big data:

Can it transform operations management? *International Journal of Operations & Production Management*, *37*(1), 37–55. https://doi.org/10.1108/IJOPM-02-2015-0084

Pallant, J. (2010). *SPSS survival manual*. Maidenhead: McGraw-Hill Education.

Persson, A., & Ryals, L. (2014). Making customer relationship decisions: Analytics v rules of thumb. *Journal of Business Research*, *67*(8), 1725–1732. https://doi.org/10.1016/j.jbusres.2014.02.019

Pigni, F., Piccoli, G., & Watson, R. (2016). Digital data streams: Creating value from the real-time flow of big data. *California Management Review*, *58*(3), 5–25. https://doi.org/10.1525/cmr.2016.58.3.5

Popovič, A., Hackney, R., Coelho, P. S., & Jaklič, J. (2012). Towards business intelligence systems success: Effects of maturity and culture on analytical decision making. *Decision Support Systems*, *54*(1), 729–739. https://doi.org/10.1016/j.dss.2012.08.017

Raguseo, E. (2018). Big data technologies: An empirical investigation on their adoption, benefits and risks for companies. *International Journal of Information Management*, *38*(1), 187–195. https://doi.org/10.1016/j.ijinfomgt.2017.07.008

Ramanathan, R., Philpott, E., Duan, Y., & Cao, G. (2017). Adoption of business analytics and impact on performance: a qualitative study in retail. *Production Planning and Control*, *28*(11–12), 985–998. https://doi.org/10.1080/09537287.2017.1336800

Ransbotham, S., Kiron, D., & Prentice, P. K. (2015a). Minding the analytics gap. *MIT Sloan Management Review*, *56*(Spring), 63–68.

Ransbotham, S., Kiron, D., & Prentice, P. K. (2015b). The talent dividend. *MIT Sloan Management Review*, *56*(4), 1–12. Retrieved from http://sloanreview.mit.edu/projects/analytics-talent-dividend/

Ransbotham, S., Kiron, D., & Prentice, P. K. (2016). Beyond the hype: The hard work behind analytics success. *MIT Sloan Management Review*, *57*(3), n/a. https://doi.org/10.1016/j.aqpro.2013.07.003

Rejikumar, G., Aswathy Asokan, A., & Sreedharan, V. R. (2018). Impact of data-driven decision-making in Lean Six Sigma: an empirical analysis. *Total Quality Management and Business Excellence*, 1–18.

https://doi.org/10.1080/14783363.2018.1426452

Rindfleisch, A., Malter, A. J., Ganesan, S., & Moorman, C. (2008). Cross-sectional versus longitudinal survey research: Concepts, findings, and guidelines. *Journal of Marketing Research*, *45*(3), 261–279. https://doi.org/10.1509/jmkr.45.3.261

Saunders, M., & Lewis, P. (2012). *Doing research in business & management: An essential guide to planning your project* (1st ed.). Harlow: Pearson.

Saunders, M., Lewis, P., & Thornhill, A. (2008). *Research methods for business students*. https://doi.org/10.1007/s13398-014-0173-7.2

Shah, S., Horne, A., & Capella, J. (2012). Good data won't guarantee good decisions. *Harvard Business Review*, *90*(4), 23–25. https://doi.org/Article

Sharma, R., Mithas, S., & Kankanhalli, A. (2014). Transforming decision-making processes: A research agenda for understanding the impact of business analytics on organisations. *European Journal of Information Systems*, *23*(4), 433–441. https://doi.org/10.1057/ejis.2014.17

Sheng, J., Amankwah-Amoah, J., & Wang, X. (2017). A multidisciplinary perspective of big data in management research. *International Journal of Production Economics*, *191*, 97–112. https://doi.org/10.1016/j.ijpe.2017.06.006

Shivakumar, R. (2014). How to tell which decisions are strategic. *California Management Review*, *56*(3), 78–97. https://doi.org/10.1525/cmr.2014.56.3.78

Simon, H. A. (1979). Rational decision making in business organizations. *American Economic Association Rational*, *69*(4), 493–513.

Simon, H. A. (1982). *Models of bounded rationality: Volume 2*. Cambridge, Mass.: MIT Press.

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of big data challenges and analytical methods. *Journal of Business Research*, *70*, 263–286. https://doi.org/10.1016/j.jbusres.2016.08.001

U.S. Census Bureau. (2015). 2015 Management and organizational practices survey. Retrieved from https://econhelp.census.gov/mops

Vidgen, R., Shaw, S., & Grant, D. B. (2017). Management challenges in creating value from business analytics. *European Journal of Operational Research*, *261*(2), 626–

639. https://doi.org/10.1016/j.ejor.2017.02.023

Viera, C. A., & Freer, K. (2015). Barriers and enablers to data-driven decision making by high school counselors and advisors. *Performance Improvement*, *54*(10), 30–40. https://doi.org/10.1002/pfi.21533

Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. (2015). How "big data" can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, *165*, 234–246. https://doi.org/10.1016/j.ijpe.2014.12.031

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, *70*, 356–365. https://doi.org/10.1016/j.jbusres.2016.08.009

Weill, P., & Woerner, S. L. (2015). Thriving in an increasingly digital ecosystem. *MIT Sloan Management Review*, *56*(4), 27–34. https://doi.org/10.1287/isre.1100.0318

Winkler, J., Kuklinski, C. P. J. W., & Moser, R. (2015). Decision making in emerging markets: The Delphi approach's contribution to coping with uncertainty and equivocality. *Journal of Business Research*, *68*(5), 1118–1126. https://doi.org/10.1016/j.jbusres.2014.11.001

World Economic Forum. (2016). *Digital transformation of industries*. Retrieved from http://reports.weforum.org/digital-transformation/wp-content/blogs.dir/94/mp/files/pages/files/digital-enterprise-narrative-final-january-2016.pdf

Zikmund, W. G., Babin, B. J., Carr, J. C., & Griffin, M. (2010). *Business research methods* (8th ed.). Mason, OH: South-Western Cengage Learning.

9 Appendices

A Questionnaire

Sources: (U.S. Census Bureau, 2015; Cao et al., 2015; Gupta & George, 2016; Wamba et al., 2017)

Demographics (Ji-fan Ren et al., 2017)		**Select Most appropriate**
Education	Primary qualification	
	Secondary qualification	
	College qualification (diploma/certificate)	
	Undergraduate degree	
	Postgraduate degree (Master/PhD)	
Age	18–25 years old	
	26–33 years old	
	34–41 years old	
	42–49 years old	
	50 years old or older	
Gender	Male	
	Female	
Industry	Accommodation and food service activities	
	Administrative and support service activities	
	Agriculture, forestry and fishing	
	Arts, entertainment and recreation	
	Construction	
	Education	
	Electricity, gas, steam and air conditioning supply	
	Financial and insurance activities	
	Human health and social work activities	
	Information and communication	
	Manufacturing	
	Mining and quarrying	
	Professional, scientific and technical activities	
	Public administration and defence; compulsory social security	
	Real estate activities	
	Transportation and storage	

	Water supply; sewerage, waste management	
	Wholesale and retail trade; repair of motor vehicles	
	Other service activities	
Work Environment	Do you work in a technical environment such as engineering or software development?	Yes/No
Role	Are you in a managerial or decision-making role?	Yes/No
	Are you an analytics professional?	Yes/No
BDA infrastructure flexibility (Wamba et al., 2017)	**Connectivity**	
	Compared similar working environments, my work environment has the foremost available analytics systems.	Seven-point Likert Scale (1-7)
	All other (e.g., remote, branch, and mobile) offices are connected to the central office for sharing analytics insights.	Seven-point Likert Scale (1-7)
	My work environment utilizes open systems network mechanisms to boost analytics connectivity.	Seven-point Likert Scale (1-7)
	There are no identifiable communications bottlenecks within my work environment for sharing analytics insights.	Seven-point Likert Scale (1-7)
	Compatibility	
	Software applications can be easily used across multiple analytics platforms.	Seven-point Likert Scale (1-7)
	Our user interfaces provide transparent access to all platforms.	Seven-point Likert Scale (1-7)
	Information is shared seamlessly across our organization, regardless of the location.	Seven-point Likert Scale (1-7)
	Modularity	
	Reusable software modules are widely used in new system development.	Seven-point Likert Scale (1-7)
	End users utilize object-oriented tools to create their own applications	Seven-point Likert Scale (1-7)
	Analytics personnel utilize object-oriented technologies to minimize the development time for new applications.	Seven-point Likert Scale (1-7)
	The legacy system within our organization restricts the development of new applications.	Seven-point Likert Scale (1-7)

BDA management capabilities (Wamba et al., 2017)	**Planning**	
	We continuously examine innovative opportunities for the strategic use of business analytics.	Seven-point Likert Scale (1-7)
	We enforce adequate plans for the utilization of business analytics.	Seven-point Likert Scale (1-7)
	We perform business analytics planning processes in systematic ways.	Seven-point Likert Scale (1-7)
	We frequently adjust business analytics plans to better adapt to changing conditions.	Seven-point Likert Scale (1-7)
	Investment Decision-making	
	When we make business analytics investment decisions, we estimate the effect they will have on the productivity of the employees' work.	Seven-point Likert Scale (1-7)
	When we make business analytics investment decisions, we project how much these options will help end users make quicker decisions.	Seven-point Likert Scale (1-7)
	When we make business analytics investment decisions, we estimate whether they will consolidate or eliminate jobs.	Seven-point Likert Scale (1-7)
	When we make business analytics investment decisions, we estimate the cost of training that end users will need.	Seven-point Likert Scale (1-7)
	When we make business analytics investment decisions, we estimate the time managers will need to spend overseeing the change.	Seven-point Likert Scale (1-7)
	Coordination	
	In my work environment, business analysts and line people meet regularly to discuss important issues.	Seven-point Likert Scale (1-7)
	In my work environment, business analysts and line people from various departments regularly attend cross-functional meetings.	Seven-point Likert Scale (1-7)
	In my work environment, business analysts and line people coordinate their efforts harmoniously.	Seven-point Likert Scale (1-7)
	In my work environment, information is widely shared between business analysts and line people	Seven-point Likert Scale (1-7)

	so that those who make decisions or perform jobs have access to all available know-how.	
	Control	
	In my work environment, the responsibility for analytics development is clear.	Seven-point Likert Scale (1-7)
	We are confident that analytics project proposals are properly appraised.	Seven-point Likert Scale (1-7)
	We constantly monitor the performance of the analytics function.	Seven-point Likert Scale (1-7)
	Our analytics department is clear about its performance criteria.	Seven-point Likert Scale (1-7)
	Our company is better than competitors in connecting (e.g., communication and information sharing) parties within a business process.	Seven-point Likert Scale (1-7)
	Our company is better than competitors in reducing cost within a business process.	Seven-point Likert Scale (1-7)
	My work environment is better than others in bringing complex analytical methods to bear on a business process.	Seven-point Likert Scale (1-7)
	My work environment is better than competitors in bringing detailed information into a business process.	Seven-point Likert Scale (1-7)
BDA personnel expertise (Wamba et al., 2017)	**Technical knowledge**	
	I am very capable in terms of programming skills (e.g., structured programming, web-based application, CASE tools, etc.).	Seven-point Likert Scale (1-7)
	I am very capable in terms of managing project life cycles.	Seven-point Likert Scale (1-7)
	I am very capable in the areas of data management and maintenance.	Seven-point Likert Scale (1-7)
	I am very capable in the areas of distributed computing.	Seven-point Likert Scale (1-7)
	I am very capable in decision support systems (e.g., expert systems, artificial intelligence, data warehousing, mining, marts, etc.).	Seven-point Likert Scale (1-7)
	Technological management knowledge	

	I show superior understanding of technological trends.	Seven-point Likert Scale (1-7)
	I show superior ability to learn new technologies.	Seven-point Likert Scale (1-7)
	I am very knowledgeable about the critical factors for the success of our organization.	Seven-point Likert Scale (1-7)
	I am very knowledgeable about the role of business analytics as a means, not an end.	Seven-point Likert Scale (1-7)
	Business knowledge	
	I understand our organization's policies and plans at a very high level.	Seven-point Likert Scale (1-7)
	Our analytics personnel are very capable in interpreting business problems and developing appropriate solutions.	Seven-point Likert Scale (1-7)
	I am very knowledgeable about business functions.	Seven-point Likert Scale (1-7)
	I am very knowledgeable about the business environment.	Seven-point Likert Scale (1-7)
	Relational knowledge	
	I am very capable in terms of managing projects.	Seven-point Likert Scale (1-7)
	I am very capable in terms of executing work in a collective environment.	Seven-point Likert Scale (1-7)
	I am very capable in terms of teaching others.	Seven-point Likert Scale (1-7)
	I work closely with customers and maintain productive user/client relationships.	Seven-point Likert Scale (1-7)
Data Driven Decision Making (Guangming Cao, Yanqing Duan, and Gendao Li) (U.S. Census Bureau, 2015)	We use data-based insight for the creation of new service/product	Seven-point Likert Scale (1-7)
	We depend on data-based insights for decision making	Seven-point Likert Scale (1-7)
	We are open to new ideas that challenge current practice based on data-driven insight	Seven-point Likert Scale (1-7)
	We have the data to make decisions	Seven-point Likert Scale (1-7)

	To what extent do you use data to support decision making in this environment?	Seven-point Likert Scale (1-7)
	We consider data a tangible asset	Seven-point Likert Scale (1-7)
	We base our decisions on data rather than on instinct	Seven-point Likert Scale (1-7)
	We are willing to override our own intuition when data contradict our viewpoints	Seven-point Likert Scale (1-7)
	We continuously assess and improve the business rules in response to insights extracted from data	Seven-point Likert Scale (1-7)

B Pilot Study Feedback Template

1. 1. How long did it take you to complete the questionnaire?
2. 2. What is your opinion of the length of the questionnaire? ◯ About right ◯ Too short ◯ Too Long
3. 3. What is your opinion of the clarity of the questions? ◯ Poor ◯ Satisfactory ◯ Good ◯ Other If "OTHER" provide details:
4. 4. What is your opinion of the structure and format of the questionnaire? ◯ Poor ◯ Satisfactory ◯ Good ◯ Other If "OTHER" provide details:
5. 5. Did you have any difficulties completing the questionnaire? ◯ Yes ◯ No ◯ Other If "OTHER" provide details:
6. 6. Does the questionnaire omit any issues you consider to be important to Big Data Analytics Capabilities and Data-Driven Decision Making? ◯ Yes ◯ No ◯ Other If "OTHER" provide details:

C Data Collection

Table 26: Data Collection Process

		Number of Members
LinkedIn Groups	Data Science and AI Professionals	17 187
	Big Data | Analytics | Strategy | Finance | Innovation	237 982
	Business Intelligence Professionals (BI, Big Data, Analytics, IoT)	220 310
	Business Analytics, Big Data and Artificial Intelligence	197 358
	Data Mining, Big Data, Data Visualisation and Data Science	171 524
Educational Institutions Groups	MSc in Big Data classes (1st and 2nd Year)	100
	Post graduate Diploma in Data Science	60
Internal Groups	Software Design and Analytics group	30
	Internal Big Data Workgroup	55
Direct Messages (email and instant messaging)	Researchers Network	60
		844 666,00

D Data Analysis

D.1 Code Book

Variable Values

Value		Label
Education	1	Primary qualification
	2	Secondary qualification
	3	College qualification (diploma/certificate)
	4	Undergraduate degree
	5	Postgraduate degree (Honours/Masters/PhD)
	999	xxx
Age	1	18–25 years old
	2	26–33 years old
	3	34–41 years old
	4	42–49 years old
	5	50 years old or older
	999	xxx
Gender	1	Male
	2	Female
	999	xxx
Industry	1	Accommodation and food service activities
	10	Information and communication
	11	Manufacturing
	12	Mining and quarrying
	13	Professional, scientific and technical activities
	14	Public administration and defence; compulsory social security
	15	Real estate activities
	16	Transportation and storage
	17	Water supply; sewerage, waste management

	18	Wholesale and retail trade; repair of motor vehicles
	19	Other service activities
	2	Administrative and support service activities
	3	Agriculture, forestry and fishing
	4	Arts, entertainment and recreation
	5	Construction
	6	Education
	7	Electricity, gas, steam and air conditioning supply
	8	Financial and insurance activities
	9	Human health and social work activities
	999	xxx
TechEnv	1	Yes
	2	No
DMRole	1	Yes
	2	No
AnProf	1	Yes
	2	No
	999	xxx
AnKn	1	None
	2	Basic
	3	Intermediate
	4	Advanced
JobLev	1	Entry Level
	2	Junior Management
	3	Middle management
	4	Senior management
	5	Owner/C-Level executive
	999	xxx

Construct Questions	1	Strongly disagree
	2	Disagree
	3	Moderately disagree
	4	Neutral
	5	Moderately agree
	6	Agree
	7	Strongly agree
	999[a]	xxx

D.2 CFA: AMOS Models

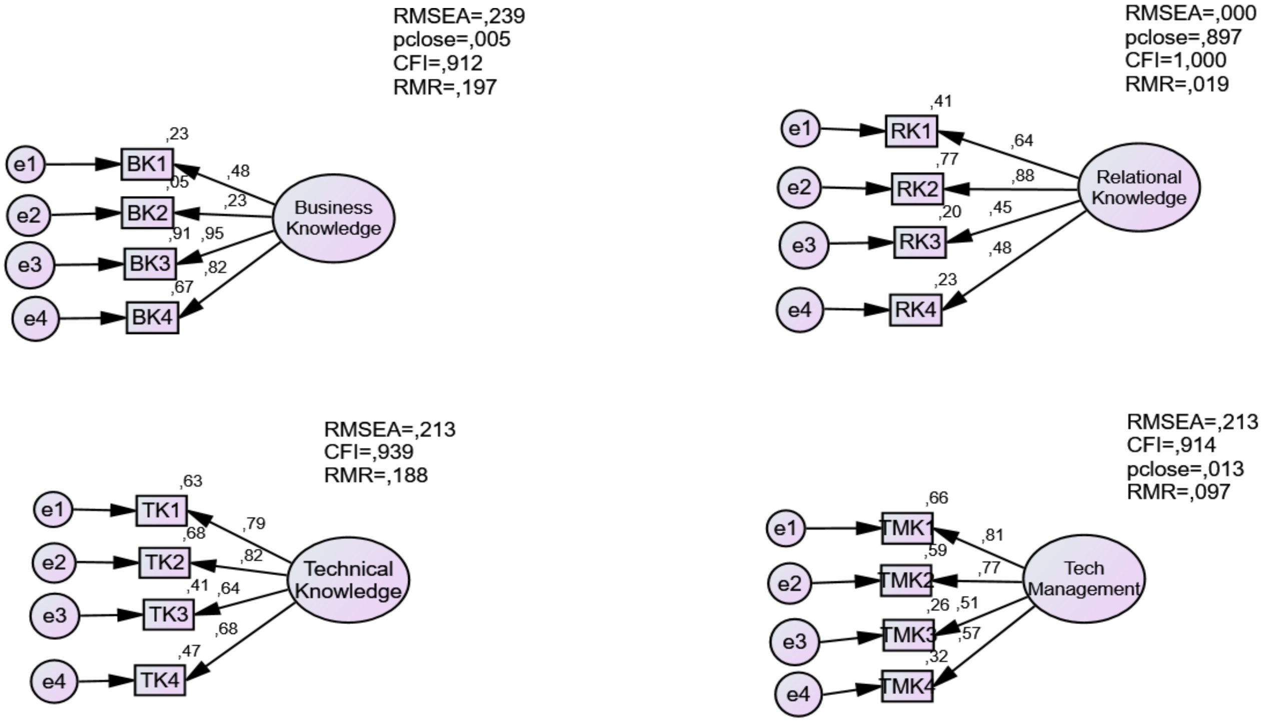


Figure 15: CFA First Order Constructs BDAPE

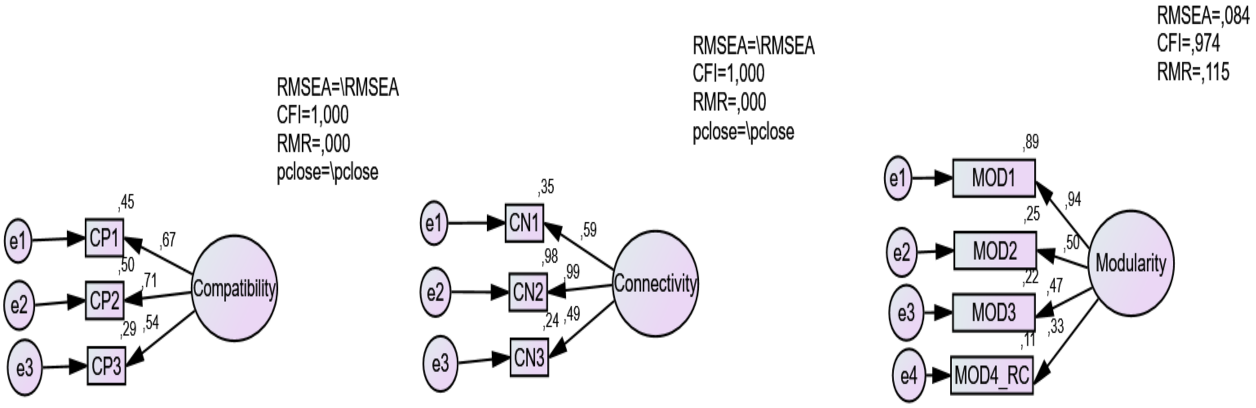


Figure 16: CFA First Order Constructs BDAIF

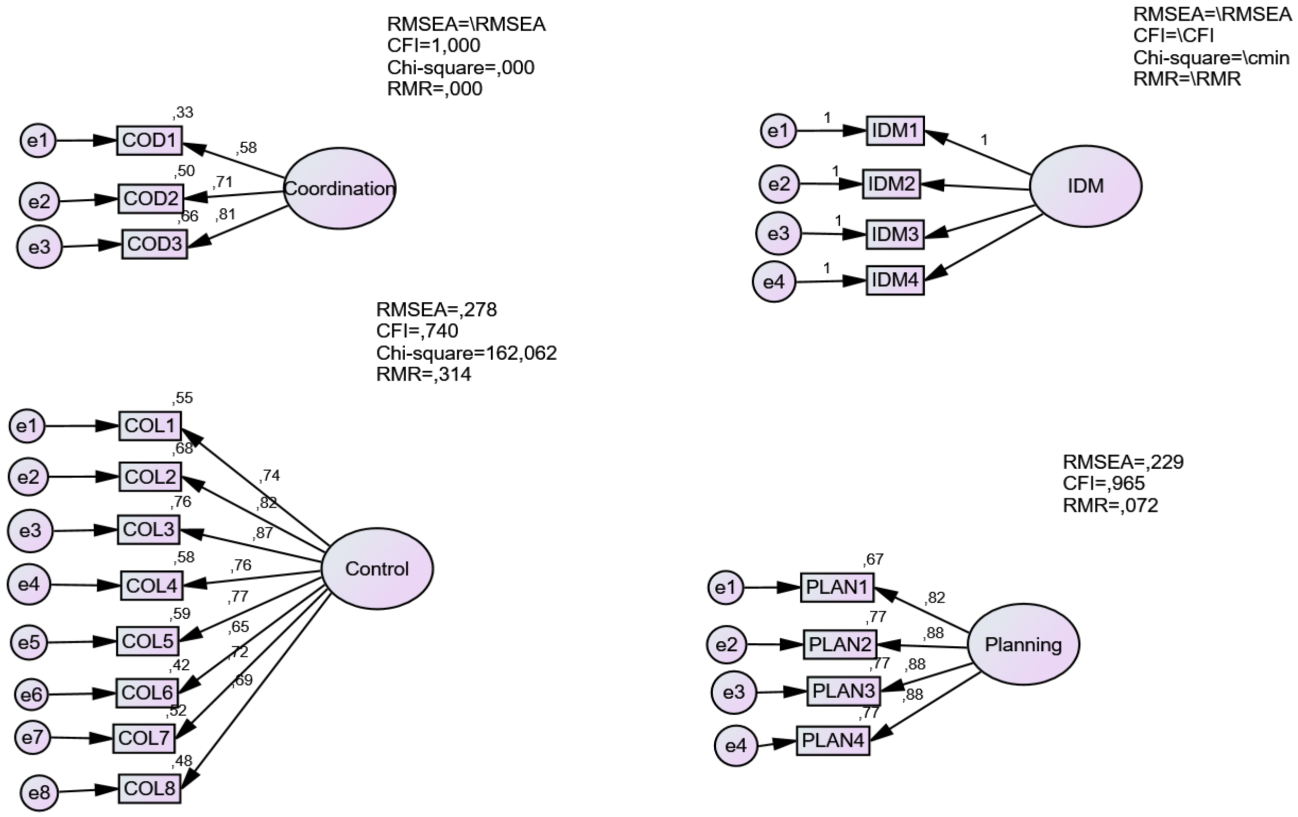


Figure 17: CFA First Order Constructs BDAMC

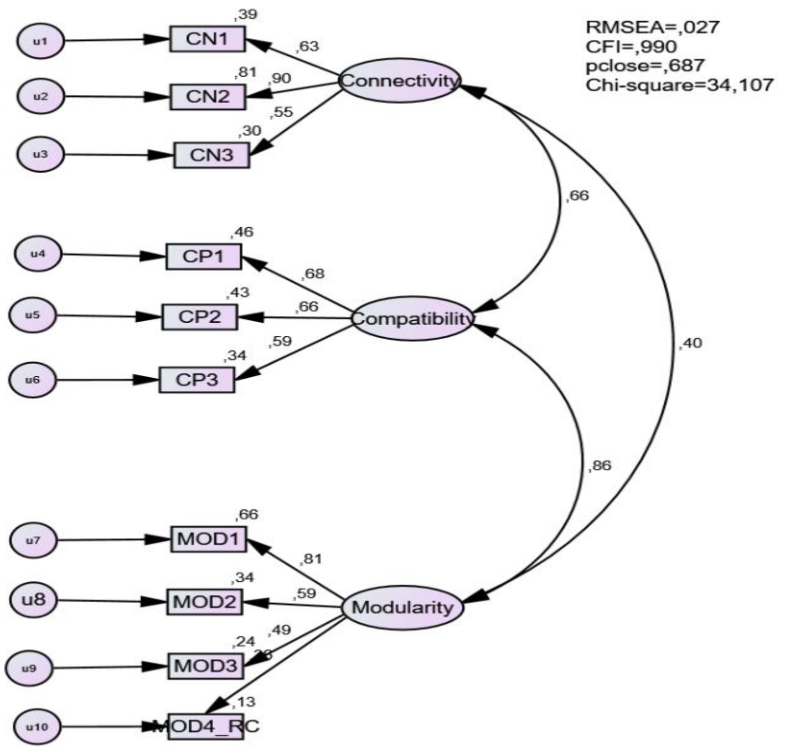


Figure 18: CFA BDAIF

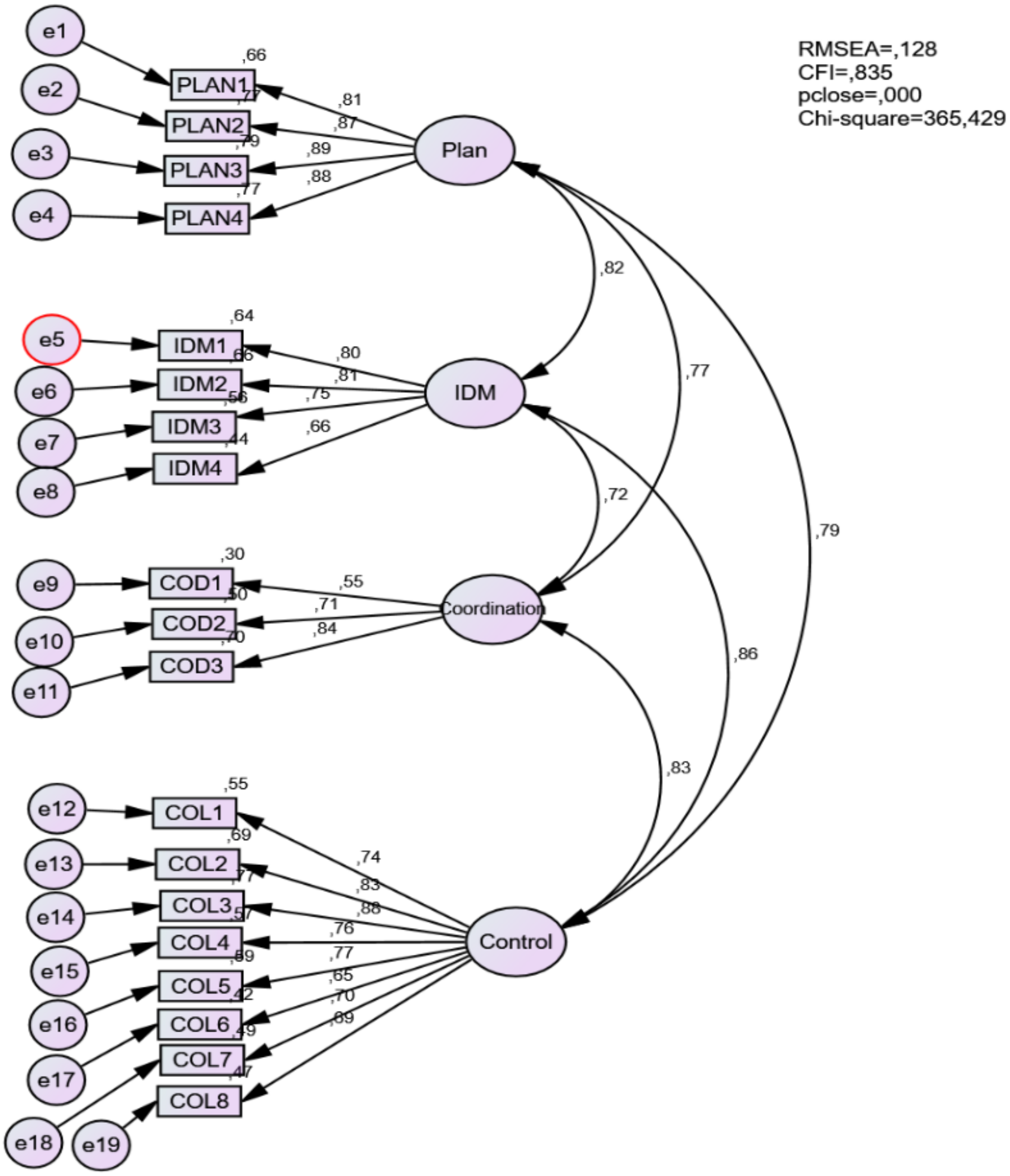


Figure 19: CFA BDAMC

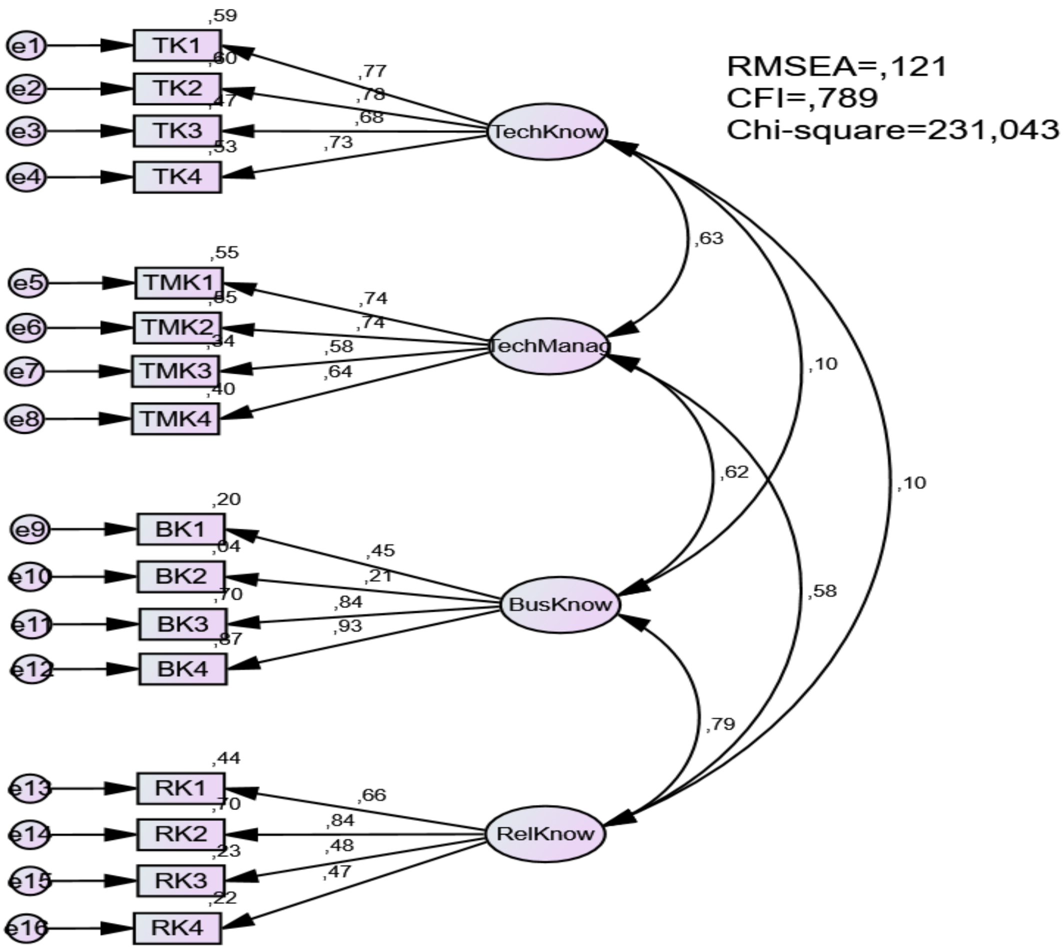


Figure 20: CFA BDAPE

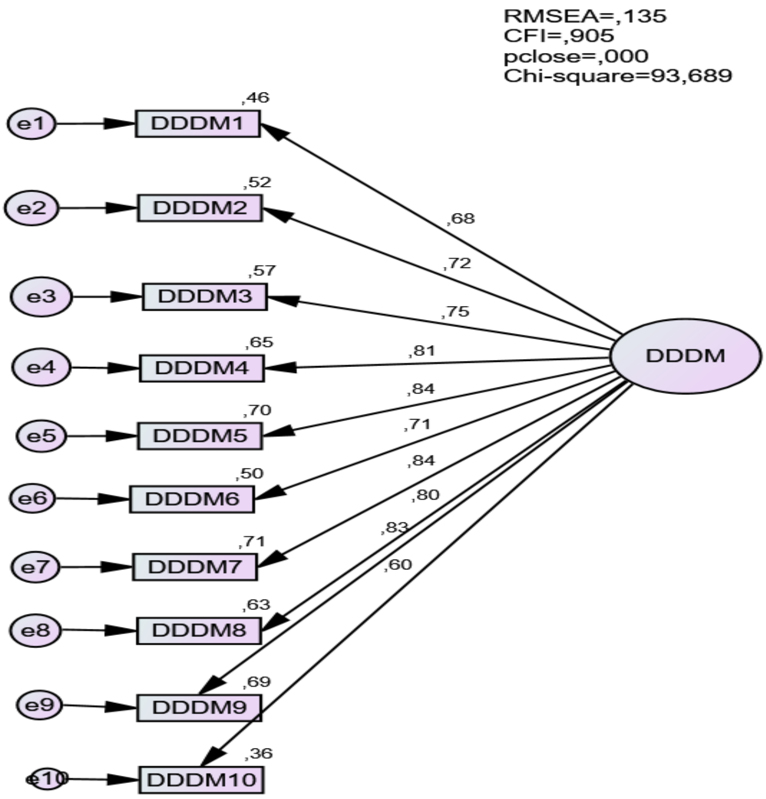


Figure 21: CFA DDDM

D.3 EFA: Correlation Matrices

Table 27: EFA Correlation Matrices – First Order Construct Questions

Correlation Matrix

		CN1	CN2	CN3
Correlation	CN1	1,000	0,582	0,287
	CN2	**0,582**	1,000	0,482
	CN3	0,287	**0,482**	1,000

		CP1	CP2	CP3
Correlation	CP1	1,000	0,477	0,362
	CP2	**0,477**	1,000	0,380
	CP3	**0,362**	**0,380**	1,000

		MOD1	MOD2	MOD3	MOD4_RC
Correlation	MOD1	1,000	0,463	0,438	0,320
	MOD2	**0,463**	1,000	0,311	0,151
	MOD3	**0,438**	**0,311**	1,000	0,025
	MOD4_RC	**0,320**	0,151	0,025	1,000

		PLAN1	PLAN2	PLAN3	PLAN4
Correlation	PLAN1	1,000	0,782	0,674	0,700
	PLAN2	**0,782**	1,000	0,763	0,743
	PLAN3	**0,674**	**0,763**	1,000	0,811
	PLAN4	**0,700**	**0,743**	**0,811**	1,000

		IDM1	IDM2	IDM3	IDM4
Correlation	IDM1	1,000	0,647	0,567	0,542
	IDM2	**0,647**	1,000	0,596	0,495
	IDM3	**0,567**	**0,596**	1,000	0,637
	IDM4	**0,542**	**0,495**	**0,637**	1,000

		COD1	COD2	COD3
Correlation	COD1	1,000	0,409	0,470
	COD2	**0,409**	1,000	0,578
	COD3	**0,470**	**0,578**	1,000

		COL1	COL2	COL3	COL4	COL5	COL6	COL7	COL8
Correlation	COL1	1,000	0,798	0,669	0,588	0,430	0,338	0,461	0,450
	COL2	**0,798**	1,000	0,807	0,700	0,518	0,366	0,493	0,408
	COL3	**0,669**	**0,807**	1,000	0,716	0,620	0,497	0,563	0,532
	COL4	**0,588**	**0,700**	**0,716**	1,000	0,573	0,385	0,473	0,426
	COL5	**0,430**	**0,518**	**0,620**	0,573	1,000	0,753	0,699	0,699
	COL6	**0,338**	**0,366**	**0,497**	**0,385**	**0,753**	1,000	0,615	0,740
	COL7	**0,461**	**0,493**	**0,563**	**0,473**	**0,699**	**0,615**	1,000	0,720
	COL8	**0,450**	**0,408**	**0,532**	**0,426**	**0,699**	**0,740**	**0,720**	1,000

		TK1	TK2	TK3	TK4
Correlation	TK1	1,000	0,693	0,463	0,502
	TK2	**0,693**	1,000	0,488	0,534
	TK3	**0,463**	**0,488**	1,000	0,591
	TK4	**0,502**	**0,534**	**0,591**	1,000

		TMK1	TMK2	TMK3	TMK4
Correlation	TMK1	1,000	0,652	0,380	0,431
	TMK2	**0,652**	1,000	0,351	0,395
	TMK3	**0,380**	**0,351**	1,000	0,500
	TMK4	**0,431**	**0,395**	**0,500**	1,000

		BK1	BK2	BK3	BK4
Correlation	BK1	1,000	0,410	0,452	0,372
	BK2	**0,410**	1,000	0,209	0,164
	BK3	**0,452**	0,209	1,000	0,783
	BK4	**0,372**	0,164	**0,783**	1,000

		RK1	RK2	RK3	RK4
Correlation	RK1	1,000	0,559	0,270	0,330
	RK2	**0,559**	1,000	0,398	0,413
	RK3	0,270	**0,398**	1,000	0,191
	RK4	**0,330**	**0,413**	0,191	1,000

		DDDM 1	DDDM 2	DDDM 3	DDDM 4	DDDM 5	DDDM 6	DDDM 7	DDDM 8	DDDM 9	DDDM 10
Correlation	DDDM1	1,000	0,734	0,579	0,491	0,558	0,516	0,473	0,490	0,543	0,425
	DDDM2	**0,734**	1,000	0,577	0,585	0,593	0,486	0,572	0,485	0,579	0,437
	DDDM3	**0,579**	**0,577**	1,000	0,562	0,714	0,500	0,581	0,570	0,609	0,474
	DDDM4	**0,491**	**0,585**	**0,562**	1,000	0,761	0,575	0,682	0,606	0,653	0,472
	DDDM5	**0,558**	**0,593**	**0,714**	**0,761**	1,000	0,616	0,635	0,655	0,645	0,499
	DDDM6	**0,516**	**0,486**	**0,500**	**0,575**	**0,616**	1,000	0,674	0,521	0,529	0,384
	DDDM7	**0,473**	**0,572**	**0,581**	**0,682**	**0,635**	**0,674**	1,000	0,762	0,742	0,531
	DDDM8	**0,490**	**0,485**	**0,570**	**0,606**	**0,655**	**0,521**	**0,762**	1,000	0,735	0,431
	DDDM9	**0,543**	**0,579**	**0,609**	**0,653**	**0,645**	**0,529**	**0,742**	**0,735**	1,000	0,519
	DDDM10	**0,425**	**0,437**	**0,474**	**0,472**	**0,499**	**0,384**	**0,531**	**0,431**	**0,519**	1,000

D.4 Validity Testing: Correlations

Table 28: Pearson Correlations: First-Order Constructs and Construct-Total

	Sig.	Pearson Correlation
	ConnectivityTotal	
CN1	0,000	.777**
CN2	0,000	.875**
CN3	0,000	.736**
	CompatibilityTotal	
CP1	0,000	.781**
CP2	0,000	.778**
CP3	0,000	.772**
	ModularityTotal	
MOD1	0,000	.683**
MOD2	0,000	.723**
MOD3	0,000	.692**
MOD4	0,028	.228*
	PlanningTotal	
PLAN1	0,000	.873**
PLAN2	0,000	.915**
PLAN3	0,000	.905**
PLAN4	0,000	.906**
	IDMTotal	
IDM1	0,000	.841**
IDM2	0,000	.828**
IDM3	0,000	.839**
IDM4	0,000	.804**
	CODTotal	
COD1	0,000	.760**
COD2	0,000	.822**
COD3	0,000	.850**
	ControlTotal	
COL1	0,000	.747**
COL2	0,000	.799**
COL3	0,000	.850**
COL4	0,000	.764**
COL5	0,000	.839**
COL6	0,000	.746**
COL7	0,000	.796**
COL8	0,000	.789**

		TechKnowlTotal
TK1	0,000	.825**
TK2	0,000	.821**
TK3	0,000	.789**
TK4	0,000	.811**

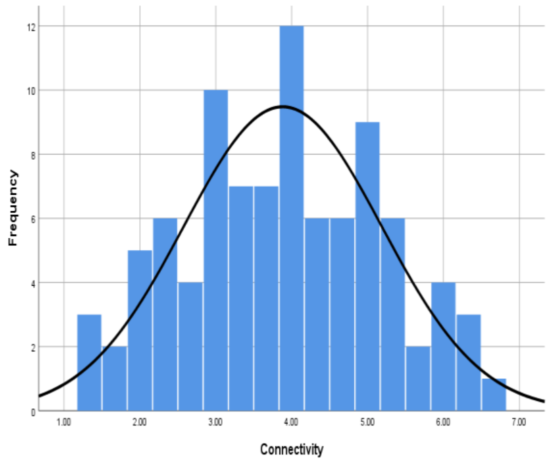
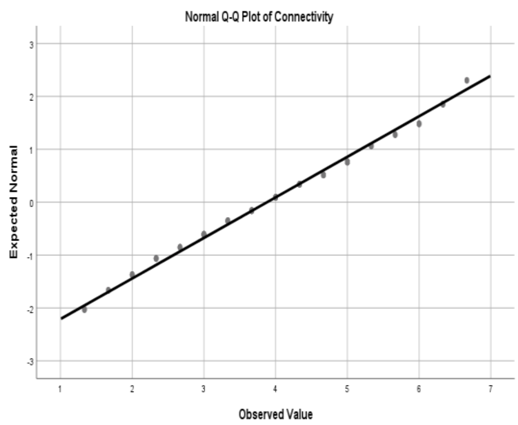
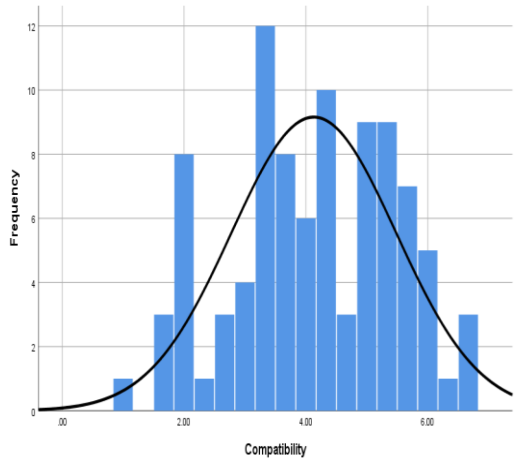
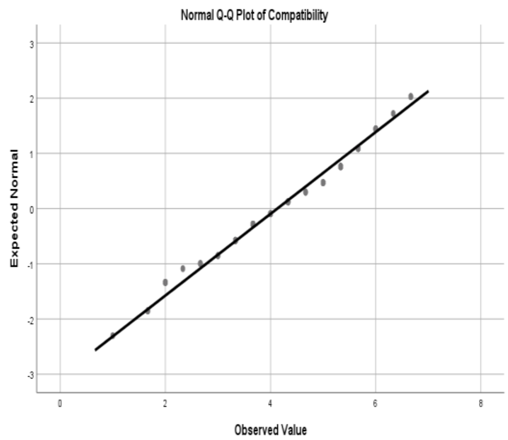
		TechManTotal
TMK1	0,000	.797**
TMK2	0,000	.787**
TMK3	0,000	.707**
TMK4	0,000	.775**

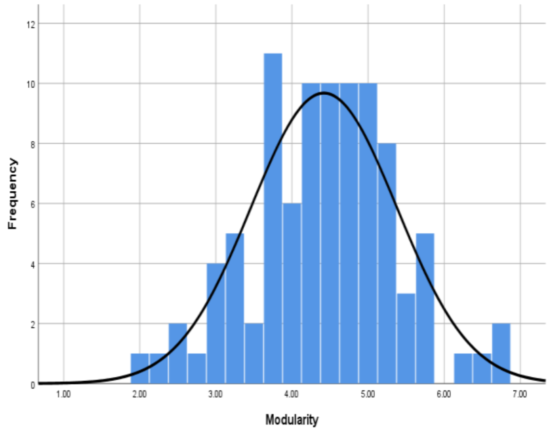
		BusKnowTotal
BK1	0,000	.788**
BK2	0,000	.638**
BK3	0,000	.792**
BK4	0,000	.738**

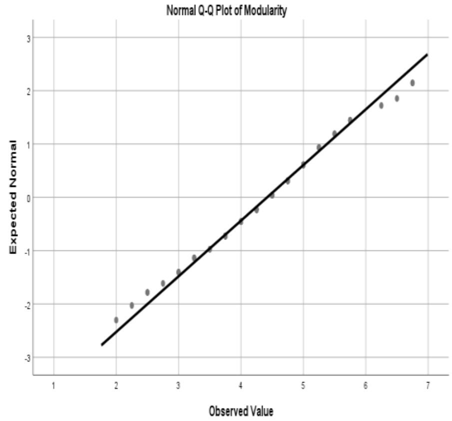
		RelKnowTotal
RK1	0,000	.714**
RK2	0,000	.763**
RK3	0,000	.606**
RK4	0,000	.770**

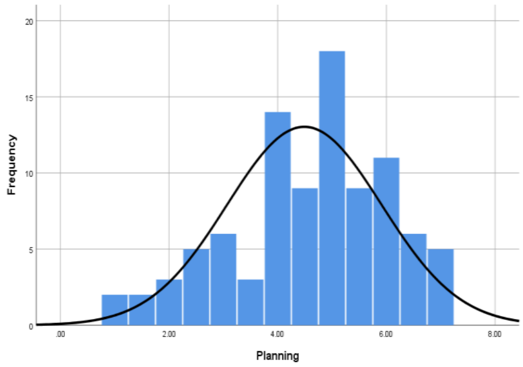
		DDDMTotal
DDDM1	0,000	.741**
DDDM2	0,000	.773**
DDDM3	0,000	.785**
DDDM4	0,000	.812**
DDDM5	0,000	.848**
DDDM6	0,000	.738**
DDDM7	0,000	.850**
DDDM8	0,000	.799**
DDDM9	0,000	.837**
DDDM1	0,000	.659**

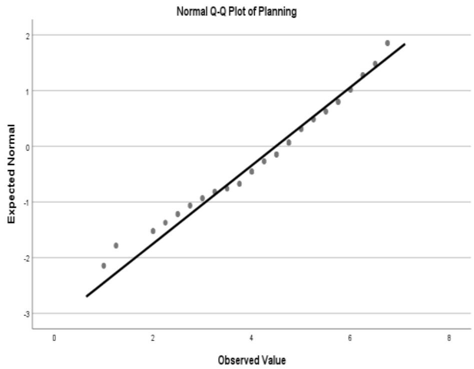
D.5 Normality Testing: Histograms and Q-Q Plot

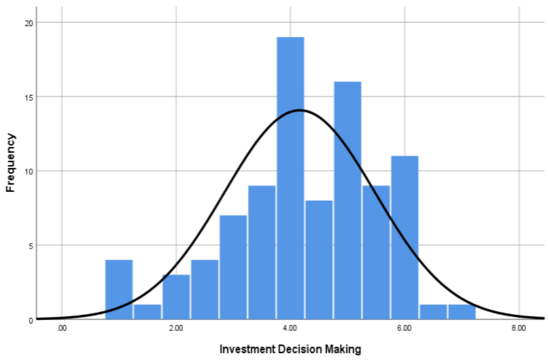
Histogram	Q-Q Plot
	
	

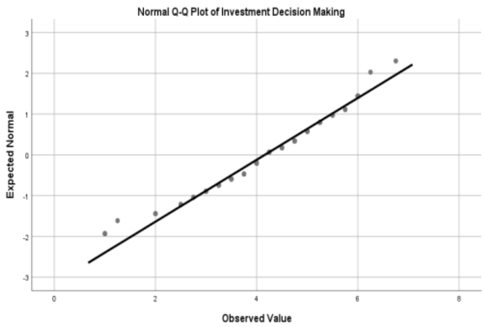

Frequency
Modularity

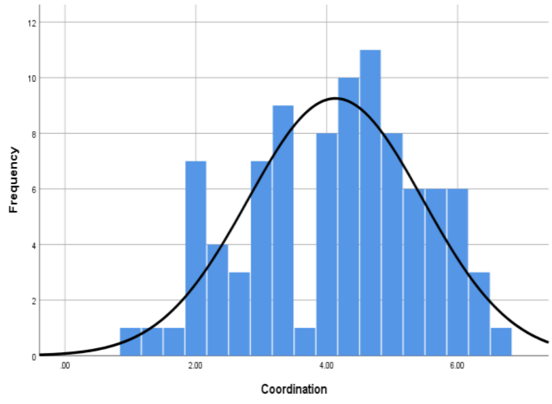

Normal Q-Q Plot of Modularity
Expected Normal
Observed Value

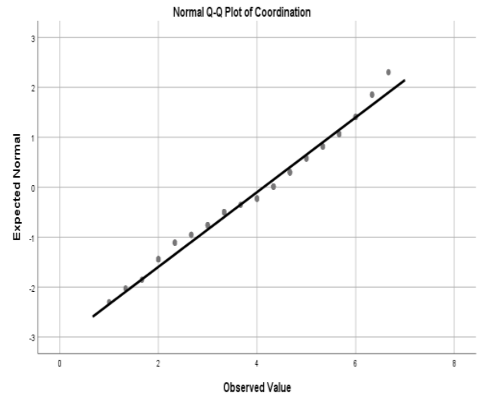

Frequency
Planning

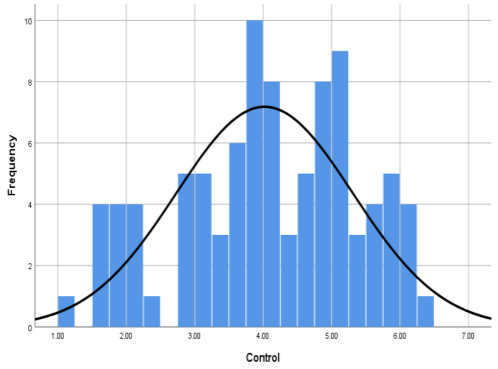

Normal Q-Q Plot of Planning
Expected Normal
Observed Value

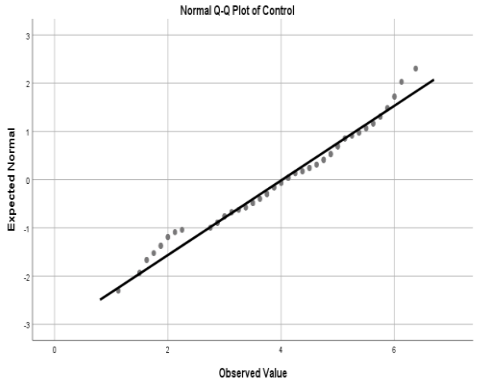

Frequency
Investment Decision Making

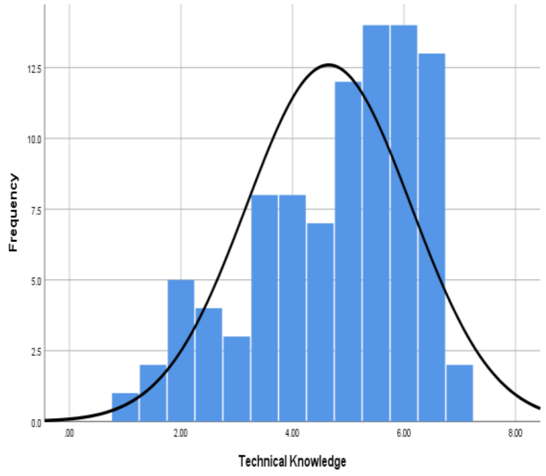

Normal Q-Q Plot of Investment Decision Making
Expected Normal
Observed Value

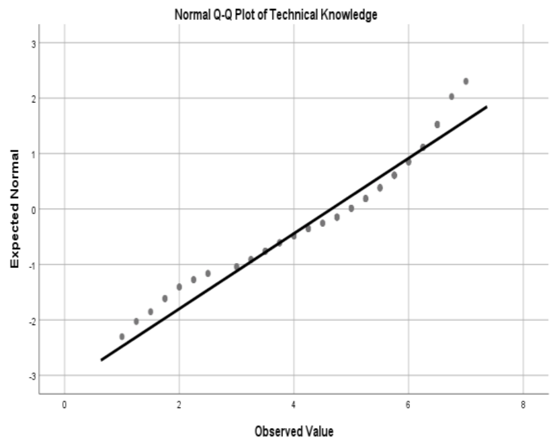

Frequency
Coordination

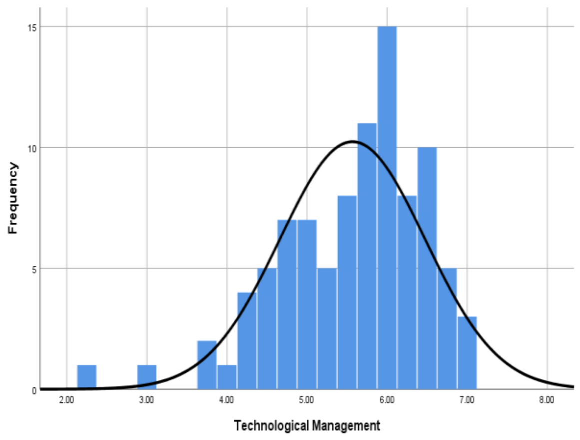

Normal Q-Q Plot of Coordination
Expected Normal
Observed Value

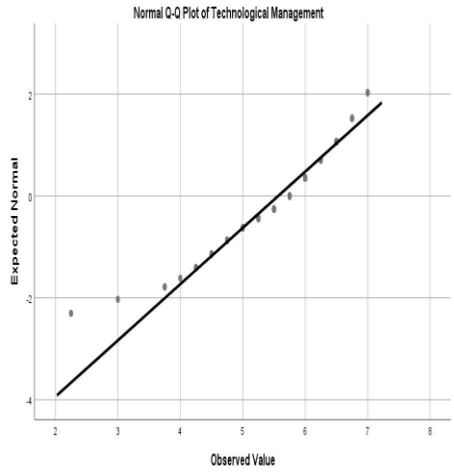

Frequency
Control

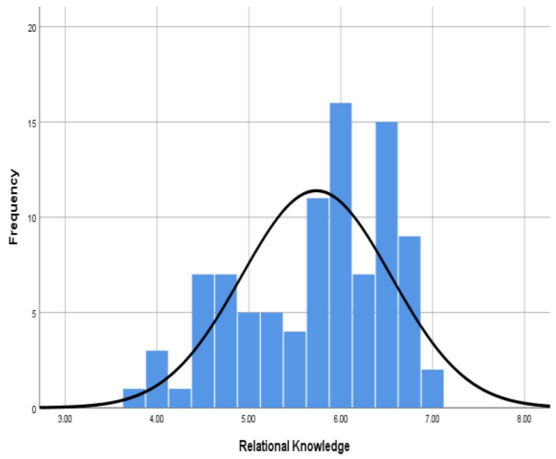

Normal Q-Q Plot of Control
Expected Normal
Observed Value

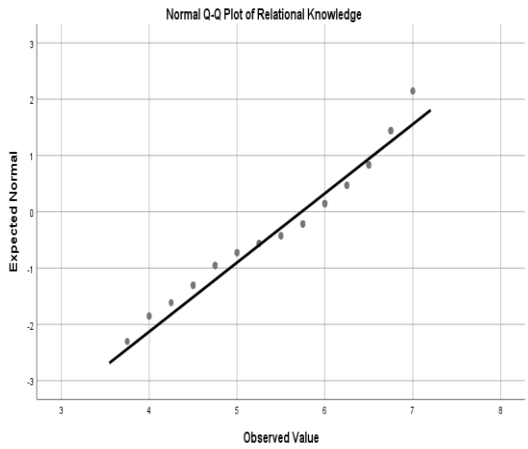

Frequency
Technical Knowledge


Normal Q-Q Plot of Technical Knowledge
Expected Normal
Observed Value


Frequency
Technological Management


Normal Q-Q Plot of Technological Management
Expected Normal
Observed Value


Frequency
Relational Knowledge


Normal Q-Q Plot of Relational Knowledge
Expected Normal
Observed Value

D.6 Homoscedasticity Testing Plots

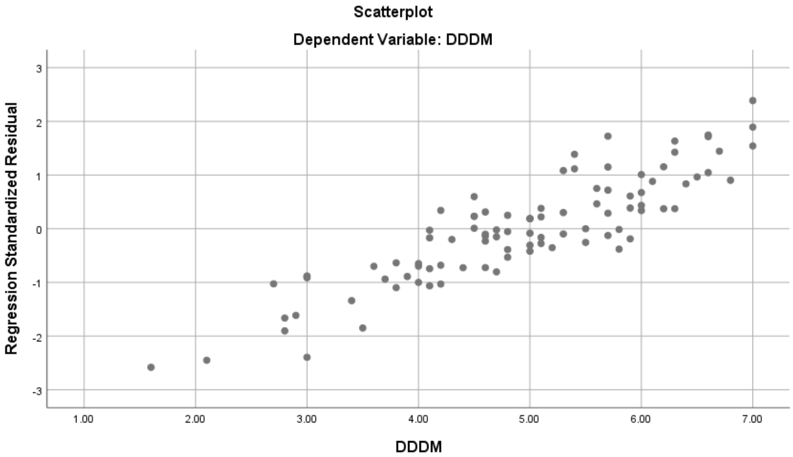


Figure 22: Scatter plot of Residuals: BDAIF and DDDM

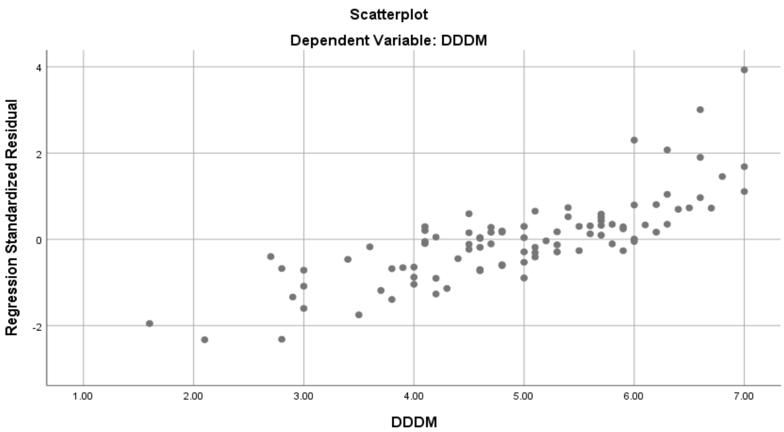


Figure 23: Scatter plot of Residuals: BDAMC and DDDM

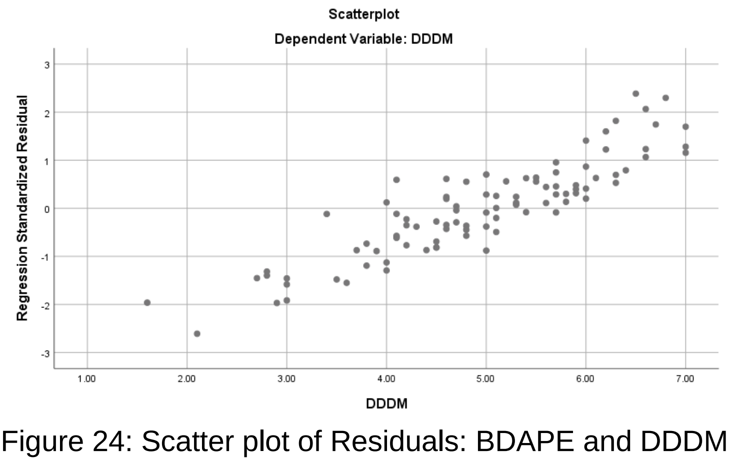


Figure 24: Scatter plot of Residuals: BDAPE and DDDM

D.7 Multicollinearity Testing Output

Table 29: Pearson Correlations Between Second Order Construct Variables

		DDDM	BDAIF	BDAMC	BDAPE
Pearson Correlation	DDDM	1.000	.473	.612	.457
	BDAIF	**.473**	1.000	.670	.319
	BDAMC	**.612**	**.670**	1.000	.336
	BDAPE	**.457**	**.319**	**.336**	1.000
Sig. (1-tailed)	DDDM	.	.000	.000	.000
	BDAIF	**.000**	**.**	.000	.001
	BDAMC	**.000**	**.000**	**.**	.000
	BDAPE	**.000**	**.001**	**.000**	.
N	DDDM	93	93	93	93
	BDAIF	93	93	93	93
	BDAMC	93	93	93	93
	BDAPE	93	93	93	93

D.8 Regression Output

Table 30: Regression ANOVA

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	56.468	3	18.823	24.077	.000[b]
	Residual	69.577	89	.782		
	Total	126.045	92			

a. Dependent Variable: DDDM

b. Predictors: (Constant), BDAPE, BDAIF, BDAMC

Table 31: Regression Coefficients

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Correlations			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	0,365	0,678		0,538	0,592					
	BDAIF	0,085	0,135	0,068	0,633	0,528	0,473	0,067	0,050	0,541	1,848
	BDAMC	0,475	0,108	0,473	4,394	0,000	0,612	0,422	0,346	0,535	1,870
	BDAPE	0,422	0,129	0,277	3,278	0,001	0,457	0,328	0,258	0,871	1,148

a. Dependent Variable: DDDM

D.9 Statistics Per Item

Table 32: Descriptive Statistics Per Question

Code	Name	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
		Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
CN1	Compared to similar working environments, my work environment has the foremost/leading available analytics systems.	93	1	7	4,31	1,567	-0,188	0,250	-0,670	0,495
CN2	All other offices (e.g., remote, branch, and mobile) are connected to the central office for sharing analytics insights.	93	1	7	3,95	1,747	-0,141	0,250	-1,070	0,495
CN3	There are no identifiable communications bottlenecks within my work environment for sharing analytics insights.	93	1	7	3,39	1,588	0,402	0,250	-0,820	0,495
CP1	Software applications can be easily used across multiple analytics platforms.	93	1	7	4,38	1,706	-0,327	0,250	-1,036	0,495
CP2	Our user interfaces provide transparent access to all platforms.	93	1	7	4,05	1,618	-0,262	0,250	-0,894	0,495
CP3	Information is shared seamlessly across our organization, regardless of the location.	93	1	7	3,96	1,893	-0,173	0,250	-1,176	0,495
MOD1	Reusable software modules are widely used in new system development.	93	1	7	4,42	1,734	-0,443	0,250	-0,807	0,495
MOD2	End-users can use reusable software tools/packages to create their own analytics applications.	93	1	7	3,87	1,801	-0,031	0,250	-1,224	0,495
MOD3	Analytics personnel utilize object-oriented technologies to minimize the development time for new applications.	93	1	7	4,58	1,378	-0,758	0,250	0,237	0,495
MOD4	The legacy system(s) within our organization restricts the development of new applications.	93	1	7	4,81	1,734	-0,616	0,250	-0,550	0,495

Code	Name	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
		Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
PLAN1	We continuously examine innovative opportunities for the strategic use of business analytics.	93	1	7	4,98	1,496	-0,938	0,250	0,514	0,495
PLAN2	We enforce adequate plans for the utilization of business analytics.	93	1	7	4,47	1,626	-0,401	0,250	-0,613	0,495
PLAN3	We perform business analytics planning processes in systematic ways.	93	1	7	4,32	1,603	-0,528	0,250	-0,480	0,495
PLAN4	We frequently adjust business analytics plans to better adapt to changing conditions.	93	1	7	4,18	1,601	-0,322	0,250	-0,809	0,495
IDM1	When we make business analytics investment decisions, we estimate the effect they will have on the productivity of the employees' work.	93	1	7	4,13	1,702	-0,219	0,250	-0,821	0,495
IDM2	When we make business analytics investment decisions, we project how much these options will help end users make quicker decisions.	93	1	7	4,55	1,585	-0,459	0,250	-0,403	0,495
IDM3	When we make business analytics investment decisions, we estimate whether they will consolidate or eliminate jobs.	93	1	7	3,97	1,514	-0,175	0,250	-0,693	0,495
IDM4	When we make business analytics investment decisions, we estimate the time managers will need to spend overseeing the change.	93	1	7	4,00	1,567	-0,329	0,250	-0,705	0,495
COD1	In my work environment, business analysts and line people meet regularly to discuss important issues.	93	1	7	4,19	1,562	-0,434	0,250	-0,707	0,495
COD2	In my work environment, business analysts and line people from various departments regularly attend cross-functional meetings.	93	1	7	4,17	1,666	-0,337	0,250	-1,079	0,495

Code	Name	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
		Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
COD3	In my work environment, information is widely shared between business analysts and line people so that those who make decisions or perform jobs have access to all available know-how.	93	1	7	4,03	1,710	-0,251	0,250	-0,843	0,495
COL1	In my work environment, the responsibility for analytics development is clear.	93	1	7	4,16	1,702	-0,339	0,250	-0,861	0,495
COL2	We are confident that analytics project proposals are properly appraised.	93	1	7	4,06	1,545	-0,435	0,250	-0,828	0,495
COL3	We constantly monitor the performance of the analytics function.	93	1	7	4,17	1,579	-0,528	0,250	-0,575	0,495
COL4	Our analytics department is clear about its performance criteria.	93	1	7	4,09	1,592	-0,226	0,250	-0,806	0,495
COL5	Our company is better than competitors in connecting (e.g., communication and information sharing) parties within a business process.	93	1	7	3,83	1,685	-0,043	0,250	-0,682	0,495
COL6	Our company is better than competitors in reducing cost within a business process.	93	1	7	3,83	1,672	-0,092	0,250	-0,934	0,495
COL7	My work environment is better than others in bringing complex analytical methods to bear on a business process.	93	1	7	4,04	1,668	0,017	0,250	-0,886	0,495
COL8	My work environment is better than competitors in bringing detailed information into a business process.	93	1	7	3,98	1,622	-0,027	0,250	-0,644	0,495
TK1	I am very capable in terms of programming skills (e.g., structured programming, web-based application, CASE tools, etc.).	93	1	7	4,63	1,977	-0,493	0,250	-1,031	0,495

Code	Name	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
		Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
TK2	I am very capable in the areas of data management and maintenance.	93	1	7	4,74	1,574	-0,485	0,250	-0,438	0,495
TK3	I am very capable in the areas of distributed computing.	93	1	7	4,66	1,879	-0,450	0,250	-1,094	0,495
TK4	I am very capable in decision support systems (e.g., expert systems, artificial intelligence, data warehousing, mining, marts, etc.).	93	1	7	4,57	1,832	-0,575	0,250	-0,916	0,495
TMK1	I show superior understanding of technological trends.	93	1	7	5,37	1,121	-0,957	0,250	1,980	0,495
TMK2	I show superior ability to learn new technologies.	93	1	7	5,68	1,235	-1,198	0,250	1,764	0,495
TMK3	I am very knowledgeable about the critical factors for the success of our organization.	93	3	7	5,56	1,047	-0,595	0,250	-0,119	0,495
TMK4	I am very knowledgeable about the role of business analytics as a means, not an end.	93	2	7	5,67	1,313	-1,181	0,250	0,944	0,495
BK1	I understand our organization's policies and plans at a very high level.	93	1	7	5,23	1,490	-1,284	0,250	1,277	0,495
BK2	Our analytics personnel are very capable in interpreting business problems and developing appropriate solutions.	93	1	7	4,72	1,402	-0,478	0,250	-0,051	0,495
BK3	I am very knowledgeable about business functions.	93	2	7	5,33	1,210	-0,672	0,250	0,305	0,495
BK4	I am very knowledgeable about the business environment.	93	2	7	5,43	1,097	-0,524	0,250	0,077	0,495
RK1	I am very capable in terms of managing projects.	93	2	7	5,69	1,053	-0,998	0,250	1,698	0,495
RK2	I am very capable in terms of executing work in a collective environment.	93	4	7	6,09	0,803	-0,673	0,250	0,142	0,495

Code	Name	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
		Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
RK3	I am very capable in terms of teaching others.	93	1	7	5,99	1,058	-1,723	0,250	5,001	0,495
RK4	I work closely with customers and maintain productive user/client relationships.	93	1	7	5,17	1,626	-1,044	0,250	0,585	0,495
DDDM1	In my environment, we use data-based insight for the creation of new service/product.	93	1	7	4,91	1,494	-0,630	0,250	0,030	0,495
DDDM2	In my environment, we depend on data-based insights for decision-making.	93	1	7	4,80	1,550	-0,384	0,250	-0,730	0,495
DDDM3	In my environment, we are open to new ideas that challenge current practice based on data-driven insight.	93	1	7	5,09	1,479	-0,872	0,250	0,342	0,495
DDDM4	In my environment, we have the data to support decision-making.	93	1	7	5,00	1,437	-0,719	0,250	0,485	0,495
DDDM5	In my environment, we use data to support decision making.	93	1	7	5,12	1,421	-0,794	0,250	0,569	0,495
DDDM6	In my environment, we consider data a tangible asset.	93	1	7	5,32	1,446	-1,092	0,250	1,159	0,495
DDDM7	In my environment, we base our decisions on data rather than on instinct.	93	1	7	4,92	1,576	-0,657	0,250	-0,197	0,495
DDDM8	In my environment, we are willing to override our own intuition when data contradicts our viewpoints.	93	1	7	4,98	1,511	-0,581	0,250	-0,423	0,495
DDDM9	In my environment, we continuously assess and improve the business rules in response to insights extracted from data.	93	1	7	4,65	1,530	-0,384	0,250	-0,593	0,495
DDDM10	In my environment, we use data from external sources (suppliers, customers, outside data providers) in decision-making.	93	1	7	4,62	1,474	-0,466	0,250	-0,398	0,495

E Ethical Clearance Letter

Gordon Institute of Business Science
University of Pretoria

02 August 2018

Dhoodhat Zaheer

Dear Zaheer

Please be advised that your application for Ethical Clearance has been approved.

You are therefore allowed to continue collecting your data.

Please note that approval is granted based on the methodology and research instruments provided in the application. If there is any deviation change or addition to the research method or tools, a supplementary application for approval must be obtained

We wish you everything of the best for the rest of the project.

Kind Regards

GIBS MBA Research Ethical Clearance Committee

Gordon Institute of Business Science
Reg. No. 99/19816/08
26 Melville Road, Illovo, Johannesburg
PO Box 787602, Sandton, 2146, South Africa
telephone (+27) 11 771 4000
fax (+27) 11 771 4177
website gibs.co.za
University of Pretoria