Submitted in part fulfilment of the requirements for the degree of

Master of Science in Business Analytics

Determinants of big data analytics capability and data-driven decision-making

By

Hanis Qistina Binte Mohammad Rizal

URN: 6762483

FACULTY OF ARTS AND SOCIAL SCIENCES

UNIVERSITY OF SURREY

JANUARY 2024

Word count: 10552 words

Executive Summary

In this applied dissertation, the modern organisational landscape is presented with not only challenges but many opportunities as it grapples with a surge in data volume. The continuous growth in data volume highlights the significance of big data being an essential tool for organisations in various industries. Big data which is characterised by its substantial volume, variety, and velocity, requires advanced analytics tools, giving rise to big data analytics which plays an important role, especially in decision-making. While existing literature focuses largely on organisational implications, there is insufficient research on the individual's big data analytic capabilities and their role in decision-making. This dissertation aims to contribute to this growing body of research by exploring the determinants of an individual’s big data analytics capabilities and their impact on decision-making effectiveness, decision-making efficiency, and subsequently the quality of big data-driven decision-making. By exploring these aspects further, this research aims to provide a more comprehensive understanding of the relationship between an individual’s capabilities in the context of big data analytics and their data-driven decision-making processes. The results from these explorations offer organisations insightful information for effective resource management and are essential for organisations who are looking to improve and optimise their decision-making processes in the current era of big data.

Extensive literature was reviewed to understand the large field of big data, big data analytics, big data analytics and data-driven decision-making. Existing literature revealed that many organisations are utilising big data in their own way to create new products and services, enhance them and create new business models. These advantages are achievable by adopting big data analytics into their organisation processes. The utilisation of big data analytics is a valuable methodology for the storage and analysis of data, facilitating improved extraction of information, forecasting, and decision-making. (Chatterjee et al., 2023). It refers to the process of analysing large volumes of data to extract value for businesses and employees (Batistic and Van Der Laken, 2019). An organisation's or an individual’s proficiency in gathering and analysing extensive amounts of data to capture valuable insights is referred to as big data analytics capabilities. It leverages on an organisation's ability to merge data assets with physical IT assets and human resources to gain a competitive advantage. Big data analytics also plays a pivotal role in enabling organisations to derive meaningful insights for informed decision-making (Dubey et al., 2019). Leveraging data and information is the most effective way to

make and execute better and faster decisions (Redman, 2008). When employees of organisations are making more accurate decisions, the costs of business operations are reduced significantly and thus improve the overall performance of the organisation (Chatterjee et al., 2023).

Primary quantitative research was done to investigate and measure an individual's big data analytics capabilities and data-driven decision-making skills. An online survey was posted through online platforms to gather responses from any volunteers. The collection period for the questionnaire lasted for a month before. The data was compiled, pre-processed, and cleaned before it was used for any analysis. Regression analysis, in particular Ordinary Least Square Regression, was performed to analyse the data gathered. Regression analysis was utilised in this dissertation to develop predictive models to help answer the proposed research questions. Regression analysis is a statistical technique to help determine the relationship between one dependent variable with one or more independent variables. The dependent variables for this study were decision-making effectiveness, decision-making efficiency, and big data decision-making quality. The independent variables used to predict what affected decision-making effectiveness and decision-making efficiency were the different big data analytics capabilities. To predict what affected big data decision-making quality, decision-making efficiency and decision-making effectiveness were used. The findings from the analysis showed there was indeed a positive and significant association between big data analytics capabilities and decision-making effectiveness or decision-making efficiency. The big data analytics capabilities which were the most significant contributor to decision-making effectiveness in particular were big data knowledge exchange and flexible infrastructure for big data. The big data analytics capability that was the most significant predictor of decision-making efficiency was process integration. The findings from the analysis also showed there was a positive association between decision-making effectiveness and decision-making efficiency with big data decision-making quality.

The theoretical and practical implications resulting from this study highlight shows that there is a need for an organisational strategy that is comprehensive and integrates investments in process optimisation, technology infrastructure, and knowledge exchange. Organisations should aim to adopt a dynamic and integrated approach to decision-making in the era of big data. By doing so, their employees will be in a better position to handle the intricacies of big data and use it to make smarter decisions when incorporating these components into their strategy planning. However, this entails not just offering the required tools and technology but also fostering an attitude that prioritises constant learning, and flexibility among employees so they would not get left behind. Enforcing this culture would guarantee that the outcomes of decision-making of their employees would remain high-quality, efficient, and effective. By improving the big data analytics capabilities of their employees, it would improve their decision-making effectiveness and efficiency, and in turn, improve the quality of big data decision-making. Not only does this approach enhance an individual's decision results, but it also contributes to the overall performance and competitiveness of organisations as a whole in the current big data era.

In conclusion, this study offers insights into the interconnectedness between the big data analytics capability, decision-making effectiveness, decision-making efficiency, and the quality of big data decision-making among individuals of various organisations. It highlights the positive influence of an individual's big data analytics capability on their effective and efficient data-driven decision-making and in turn, the quality of big data decision-making. From the findings, organisations should intend to harness the full potential of big data for improved decision outcomes of their employees. This in turn would improve the overall performance of their organisation and have a competitive advantage against rivals.

Declaration of Originality

I hereby declare that this thesis has been composed by myself and has not been presented or accepted in any previous application for a degree. The work, of which this is a record, has been carried out by myself unless otherwise stated and where the work is mine, it reflects personal views and values. All quotations have been distinguished by quotation marks and all sources of information have been acknowledged by means of references including those of the Internet. I agree that the University has the right to submit my work to the plagiarism detection sources for originality checks.

Full Name: Hanis Qistina Binte Mohammad Rizal

Date: 07/01/2024

Signature:

Acknowledgements

I would like to put on record my heartfelt thanks to those who have walked this journey with me. There are many who truly deserve to be recognised for their contribution to the completion of my dissertation.

First and foremost, I would like to express my gratitude to my dissertation supervisor, Dr Saqib Shamim, for his guidance throughout the development of my dissertation. His advice and patience when assisting me were very much appreciated.

I would also like to thank Dr Athina Ioannou who motivated me throughout this period and provided valuable insights and concepts when developing my regression models.

Finally, I am deeply grateful to my parents, brothers and aunt for their endless support, reassurance, and prayers during the entirety of this Master Programme. Thank you for always being there for me since day one.

Table of Contents

List of Tables

List of Abbreviations

BD Big Data

BDA Big Data Analytics

BDAC Big Data Analytics Capabilities

BDC Big Data Collaboration

BDDMQ Big Data Decision Making Quality

BDKE Big Data Knowledge Exchange

DDDM Data-Driven Decision-Making

DMEF Decsion-Making Effectiveness

DMEFF Decsion-Making Efficiency

FIBD Flexible Infrastructure for Big Data

LFBD Leadership Focus on Big Data

PI Process Integration

R Routinising

Introduction

1.1 Introduction to Big Data Analytics and Decision Making

The growing volume and complexity of data offer both opportunities and challenges in the ever-changing world of modern organisations. According to the research done by Dell Technologies (2020), it was revealed that organisations manage an estimated 13.53 petabytes (PB) of data. That is a 40% increase from 2018 and a much higher increase of 831% in 2016. This shows that more data is being generated each day and will most likely continue to grow in volume at an increasing rate in the upcoming years. Data is said to be often generated by other organisations, by social media users or provided by devices of the Internet of Things (IoT) (Janssen et al., 2017). With data being generated at an unprecedented rate in the modern digital age, Big Data (BD) has become a vital tool for business organisations across various industries. BD is essential in today's age and will play a pivotal role in the global economy. Redman (2008) also describes the mutually beneficial relationship between data and a business organisation, i.e., where there is data smoke, there is business fire.

BD is defined as datasets that are both big and high in velocity and variety which makes them challenging to handle using traditional tools and techniques (Elgendy and Elragal, 2014). According to Dietrich, Heller, and Yang (2021), BD is characterised by three main features defined by the three V's – volume, variety, and velocity. Volume refers to the size and how big the data is. Velocity refers to the rate at which data is changing or created. Variety of data refers to the various types and formats of data and the different kinds of uses and ways to analyse them. Another definition of BD given by Manyika et al. (2011) is data whose scale, distribution, diversity, and timeliness require the use of new technical applications and analytics that, in turn, provide insights that open up new business opportunities for organisations.

One of the primary applications of BD in organisations lies in analytics. BD is closely associated to Big Data Analytics (BDA), which is required to create value for the data (Janssen et al., 2017). BDA refers to the data sets and analytical techniques in applications that are too large and intricate, hence requiring advanced and unique storage, management, analysis, and visualisation technologies (Chen et al, 2012). These tools offer insights to organisations, which could not be done with previous outdated technological tools (Wamba et. al, 2017). The ability to make use of all the information and knowledge available is an essential element of an organisation's success (Olszak, 2016). This ability to extract actionable intelligence from complex data sets makes BDA a key factor that empowers organisations and influences decision-making, innovation, and competitiveness within organisations (Roy, 2021).

However, conducting data analysis solely is not enough. The results of data analysis must be efficiently circulated among employees of the organisations to influence decisions and allow data to be part of their decision-making process rather than intuition. (Björkman, Franco and Caesarius, 2017). An individual's capacity to make wiser decisions is often where the value of BD lies. (Economist Intelligence Unit, 2012). The quality of decisions made by individuals are not only dependent on the data but also on the procedures followed during data collection and processing (Janssen et al., 2017). Based on research done by Janssen et al. (2017), BD can offer some unusual inflexion points for fresh perspectives and understandings and has the ability to enhance an individual's decision-making skills. Good decisions can lead to satisfied employees and improved competitive positions for businesses (Redman, 2008). Hence, BDA and data-driven decision-making have cemented their position as a crucial asset for organisations and employees striving to prosper in today's data-driven economy.

1.2 Research Problem and Research Gaps

Despite the increasing adoption of BDA in organisations and organisations investing heavily in acquiring and implementing BD technologies, there exists a significant gap in understanding the nuanced relationship between the big data analytics capabilities of individuals, who are the key decision makers within these organisations, and their effective and efficient data-driven decision-making processes and the quality of their decision made in the context of BD. A significant research gap exists in the studies focusing on an individual's contribution to their decision-making in the context of BD. Existing literature often focuses on organisational-level implications of big data analytics and their decision-making, overlooking an individual's role. There is a gap in understanding how an individual's ability in big data analytics directly influences the quality and efficiency of their decision-making. There is still much to explore about the intricate relationships that exist between a person's capability for BDA, decision-making effectiveness, and decision-making efficiency. For organisations looking to optimise decision processes in the current big data era, it is vital to unravel these dynamics. The specific measurements and indicators that define decision-making effectiveness and decision-making efficiency in the context of BD are not clearly defined. Effectiveness and Efficiency in decision-making within the world of BDA remains a relatively unexplored area. The factors that lead to effective and efficient decision-making processes must be recognised and quantified in order to provide valuable information for maximising the use of time and resources. While effective decision-making is often highlighted in studies, how it relates to decision quality in the context of BD is unclear.

1.3 Research Objectives and Research Questions

This research aims to focus on studying the determinants that influence the development of the big data analytics capabilities (BDAC) of employees within organisations and the subsequent impact of these capabilities on their data-driven decision-making. This research also aims to contribute to a deeper understanding of how organisations use big data to drive success by investigating the factors that shape the effectiveness of big data analytics and its influence on their decision-making process. It also aims to bridge the gap by exploring the relationship between the BDAC of individuals, decision-making effectiveness, decision-making efficiency, and the quality of decisions made with BD. There is also a shortage of literature on how the practical implications of big data analytics capabilities impact decision-making quality. This research seeks to address these theoretical and practical gaps by delving into the intricate dynamics of individual capabilities in BDA and their direct impact on decision-making effectiveness, efficiency, and quality. As organisations increasingly rely on data-driven insights, understanding and enhancing individual contributions to decision processes emerge as critical imperatives.

The research questions for this study are:

- How does an individual's big data analytics capability contribute to the effectiveness of decision-making of an individual?
- How does an individual's big data analytics capability contribute to the efficiency of decision-making of an individual?
- How does an individual's decision-making effectiveness affect their big data decision-making quality?
- How does an individual's decision-making efficiency affect their big data decision-making quality?

1.4 Chapter Structures

The next section reviews the literature on BDA and data-driven decision-making (DDDM). This section will provide more understanding of BDA and DDDM based on current literature by different scholars. Next, the methodology used to analyse the data is introduced. In this section, how data is being collected and analysed is discussed. Following that, the results and findings from analyses are thoroughly described. In this section, the outcomes of the model using the Stata software are discussed. The next section analyses and discusses the study's contributions. This section will summarise the paper's findings and compare them with the studies based on current literature. This paper closes with a discussion on limitations and future research directions.

Literature Review

2.1 Big Data

Based on existing literature, Big Data (BD) has been defined in multiple means by different scholars. Elgendy and Elragal (2014) consider BD as datasets that are both big and high in variety and velocity which makes them challenging to handle using traditional tools and techniques. Waller and Fawcett (2013) support this view and define BD as datasets that are too large for any regular systems which aids in processing data, hence requiring the use of new technologies to process them. BD is typically identified by three Vs: Volume, Velocity, and Variety, as per McAfee & Brynjolfsson (2012). However, Gandomi and Haider (2015) believe that three more Vs – Value, Variability, and Veracity should be added to the list that defines BD. Fan et al. (2014) refer to BD as the ocean of information, while Boyd and Crawford (2012) define it as a cultural, technological, and scholarly phenomenon. Nonetheless, according to research done by Jeble et al. (2018), what businesses do with the information they have matters more than the volume of data they have. Numerous organisations are utilising BD in a variety of ways. In digital marketing, Organisations use BD to obtain insights into the unmet demands and expectations of their consumers, which eventually contributes to the long-term viability of their organisation. (Cavlak and Cop, 2021). In the realm of e-commerce, BD is used to manage transactions, products, and services, which can result in a 30% increase in the business management's efficiency (Xie, 2021). BD is also proving to be a promising field in healthcare. The uses of BD in the healthcare industry include disease surveillance, epidemic control, clinical decision support, and population health management (Sabharwal et al., 2016). There are many significant benefits of using BD in the healthcare industries, one being the detection of diseases at early stages (Wang and Alexander, 2019). According to Manyika et al.'s (2011) research, BD enables organisations to create new products and services, enhance existing ones, and create new business models. These advantages are achievable by adopting big data analytics in various areas such as customer intelligence, supply chain intelligence, performance, quality and risk management, and fraud detection (Manyika et al., 2011). Additionally, organisations should strive to extract maximum value from the multitude of data they collect and store (Yousuf et al., 2020). Big Data Analytics (BDA) presents an effective means to achieve this objective. BDA is intricately linked to BD and is essential in unlocking the true potential of data (Janssen et al., 2017). The utilisation of Big Data Analytics (BDA) is a valuable methodology for the storage and analysis of data, facilitating improved extraction of information, forecasting, and decision-making. (Chatterjee et al., 2023).

2.2 Big Data Analytics

Big Data Analytics (BDA) refers to the process of analysing large volumes of data to extract the value for businesses and employees (Batistič and der Laken, 2019). Sabharwal and Miah (2021) define BDA as the computational intelligence techniques that are used to change raw data into insights that can be used to support decision-making. According to Saggi and Jain's (2018) research, BDA has been expressed in terms of six main components – data generation, data acquisition, data storage, advanced data analytics, data visualisation, and decision-making for value-creation. BDA is the future of management and can offer endless business opportunities, which in turn revolutionise various industries. BDA plays a crucial role in helping businesses gain a comprehensive understanding of the current and future trends and determine the most effective actions to achieve optimal outcomes (Lavelle et al., 2010). With the implementation of BDA, organisations can efficiently collect and integrate vast amounts of data, analyse the data acquired in different formats and structures, and transform the results into valuable insights for decision-making that surpass traditional methods (Garmaki et al., 2016).

BDA has been successfully implemented in various industries to allow them to gain insights and make informed decisions from big data. For instance, BDA is used in manufacturing to improve the performance and consumer loyalty of manufacturing industries, in which BDA helps businesses identify main strategies, metrics and concepts that were derived from BDA and marketing analytics (Mishra et al., 2023). BDA is also used to analyse the data of more than 8000 key enterprises in China's environmental protection industry (Ikegwu et al., 2022). It helps in identifying the unmet demands and expectations of consumers and achieving sustainable organisation success.

Despite BDA having a significant potential to enhance any firm's performance, research indicates that leaders of some organisations are reluctant to make major investments in BDA. This is often due to first hand experience with disappointing results or observing other organisations failing in their BDA investments (Woener et al., 2015). The findings of Manyika et al.'s (2011) research indicate a noteworthy disparity between the desired utilisation of Big

Data Analytics (BDA) by managers and its current actualisation, despite the crucial role BDA plays in generating value for organisations.

Chatterjee et al. (2023) raises the question of whether organisations can effectively convert vast amounts of stored data into valuable knowledge and information that can offer a competitive edge and enable them to survive. Chatterjee et al. (2023) also highlight potential challenges associated with the use of BDA in organisations, particularly regarding the analysis of confidential and personal data. Such challenges could raise concerns about privacy and security. Therefore, in ensuring a seamless implementation of BDA in any organisation, it is necessary for organisations to identify the risks they may face and at the same time find solutions to mitigate these risks (Neirottie et al., 2021).

2.3 Big Data Analytics Capabilities

Big Data Analytics Capabilities (BDAC) refers to the proficiency of an organisation's or an individual's in gathering and analysing extensive amounts of data to capture valuable insights. BDAC leverages on the ability of an organisation to merge data assets with physical IT assets and human resources to be ahead among other competitors. Achieving BDAC involves implementing comprehensive organisation-wide processes, roles, and structures that prioritize the cultivation of data, technology, and talent (Mikalef et. al, 2020). Based on research done by Wang (2019), BDAC entails acquiring, storing, processing, and analysing vast quantities of data by utilising advanced analytics techniques such as machine learning, data mining and predictive modelling.

According to Wamba et al.'s (2017) research, there are three main components of BDAC – BDA infrastructure flexibility, BDA management capability, and BDA personal expertise capabilities. BDA infrastructure flexibility involves BDA connectivity, capability, modularity, agility, security and risk management and data management (Kim et al., 2012; Lee et al., 2007; Weill et al., 2002). This refers to an organisation's ability to efficiently manage large and complex data sets. It involves handling data storage, processing, and retrieval, as well as expanding the infrastructure to meet the increasing demands of big data. BDA management involves BDA planning, control, and coordination (Fink and Neumann, 2007; Weill et al., 2002). This component refers to the organisation's ability to manage and govern the data effectively. It includes the ability to define data policies, procedures, and standards, as well as manage data quality, security, and privacy. BDA personal expertise capabilities refer to BDA technical knowledge, technology knowledge, business knowledge, and relational knowledge (Kim et al., 2012; Aral and Weill, 2007). This component refers to the organisation's ability to hire, train, and retain individuals with the necessary skills and expertise to work with big data. It includes the ability to identify the required skills, provide training and development opportunities and cultivate a mindset that values data-driven decision-making. However, Shamin et.al (2019) suggests that BDA management capability could also include leadership, talent management and culture, especially with regards to decision-making skills of organisations. BDAC is becoming a more crucial role in the decision-making process of businesses (Awan et al., 2021). BDAC demonstrates an organisation's ability to effectively mobilise and deploy BDA resources, utilise them, and align BDA planning with the firm's strategy to gain a competitive advantage in the industry and enhance its performance (Garmaki et al., 2016).

2.4 Data-Driven Decision Making

Making a decision involves selecting an choice from a range of options in order to accomplish a desired goal. According to Malakooti's (2010) research, decision-making is a process that is multi-dimensional and intricate, and often occurs spontaneously and naturally, though at times, decision-making may be a process that has to be systematically planned with much thought and foresight. Organisations are constantly trying to determine how they can use data to help them in decision-making, especially in today's era of big data (Visinescu et al., 2016). Decision-making is a processthat is supported by information or data that can be processed in a meaningful manner at any level of an organisation (Ahmed et al., 2022). It is believed that lately, individuals are increasingly making decisions based on the results of big data rather than their own intuition (Davenport 2006; Lavalle et al. 2010). The quality of decisions made in the course of an organisation's present operations is influenced by how efficiently and effectively it uses the data sets that are accessible (Ahmed et al., 2022). Grušovnik et al. (2017) suggest that by either making high-quality decisions, making faster decisions, or implementing more effective decisions, allows successful organisations to outperform their rivals in the industry. Organisations are increasingly challenged by data-driven insights, which can lead to effective decision-making for organisations (LaValle et al., 2010). Effective decision-making ensures clarity and purpose for organisations and individuals. Introducing data-driven insights enhances decision-making and strengthens business activities. Due to big data analytics' ability to produce deep data insights, the idea of data-driven insights has recently gained much attention by many (Awan et al., 2021). The characteristics of data-driven insights are linked to three approaches: normative, descriptive, and prescriptive insights as suggested by Sheshadri (2015) and Awan et al. (2021). The normative approach examines how individuals derive accurate decisions made by them. The descriptive approach describes how it is possible to make accurate decisions by individuals. The perspective approach allows the understanding of the decision-making process that provides better support for individuals when making decisions. When employees of organisations are making more accurate decisions, the costs of business operations are reduced significantly and thus improve the overall performance of the organisation (Chatterjee et al., 2021; Pham and Lo, 2023). The quality of decision-making is defined by Raghunathan (1999) as the ability of a decision-maker to make the right decision. This is also referred to as the quality of the decision made by the decision-maker. To understand how the quality of data-driven decision-making can be measured, the paper evaluates decision-making based on the effectiveness and efficiency of an individual's decision-making as suggested by Clark et al. (2017). This view can be supported Shamin et al. (2019) who state that the achievement of better results from data-driven decision-making and its effectiveness, is the outcome of quality decision-making. The degree to which decision-makers are satisfied with the results they have achieved when making decisions can be used to gauge the effectiveness of an individual's decision-making, and the resources used — such as time and money — that go into the process of decision-making can be used to gauge its efficiency (Kaltoft et al., 2013). Decision-making quality ensures that the key decision-makers of any organisation are certain of what they are doing and what they are aiming to accomplish. By doing so, decision-makers within organisations can make better judgments to contribute to the success of the organisation and gain a competitive advantage.

2.5 Relationship Between Big Data and Decision-Making

Big data analytics (BDA) plays a pivotal role in enabling organisations to derive meaningful insights for informed decision-making (Dubey et al., 2019). BDA helps an individual by simplifying the process of drawing conclusions from the data (Ghasemaghaei and Calic, 2019). To facilitate informed decision-making within an organisation's growth trajectory, it is imperative to establish efficient processes that encompass ongoing diagnosis, strategic action planning, and comprehensive implementation and evaluation of big data analytics (Sabharwal and Miah, 2021). As per Redman's (2008) assertion, leveraging data and information is the most effective way to make and execute better and faster decisions. By having access to better data and information in advance, utilising them to make prompt and informed decisions, and aligning the entire organisation, firms can operate with greater efficiency and confidence as suggested by Shamin et al. (2019). Similarly, based on Yiu’s (2012) research, with the use of big data, organisations can enhance their data-driven decision-making, which also enhances the organisation’s effectiveness and efficiency. Research done by Tseng et al. (2022) states that BDA allows individuals to analyse data effectively which facilitates accurate decision-making to improve the performance of organisations by reducing the costs of their operations. Having a good level of big data analytics capability can help the effectiveness and efficiency of decision-making at an individual and organisational level (Chatterjee et al., 2023) Mikalef et al. (2020) further contend that analytical applications have gained significance in evidence-based decision-making, given the growing reliance on BD in organisational decision-making. Kościelniak and Puto (2015) conducted a study on the usability of big data in the decision-making process across multiple organisations. From the study made by Kościelniak and Puto (2015), four stages of the decision-making process – to determine the authorised source of data on the performance of the organisation, evaluate employees, establish explicit management of business rules, and improve the effects of an organisation’s activity – were formulated. also propose A framework named B-DAD (Big – Data Analytics and Decisions) that was proposed by Elgendy and Elgaral (2016) illustrates the relationship between BD, BDA and Decision Making. The framework consists of five different phases – the intelligence phase, the design phase, the choice phase and, the implementation phase. Big data analytics and its relationship with decision-making has garnered significant attention and research. The indispensable role of BDA in enabling informed decision-making by providing crucial insights and information has been widely acknowledged. Several studies have delved into this association and its impact on organisational performance and strategic management. Di Berardino and Vona’s (2023) study sheds light on the various factors that contribute to effective decision-making, as well as the key impacts of utilising BD and BDA. It also underscores the importance of exploring novel organisational factors, data chain dynamics, and inhibitors to overcome obstacles in the decision-making process. Thirathon (2016) examines the impact of BDA on organisational performance. The study delved into the significance of data analysis and data-driven decision-making capabilities in determining the success of an organisation and the findings of the study revealed the crucial role of BDA in shaping decision-making processes and its consequential effect on organisational performance. Numerous organisations have successfully integrated big data analytics into their decision-making processes, as evidenced by the achievements of Amazon, Netflix, and Walmart. Amazon has harnessed the power of big data analytics to tailor customer experiences, optimize supply chain management, and employ data-driven strategies for business decisions. Netflix employs big data analytics to suggest personalised content to its users, enhance streaming quality, and make informed decisions about content creation and acquisition. Walmart has integrated big data analytics to refine inventory management, enrich customer experiences, and make sound decisions about pricing and promotions. These instances showcase how companies across diverse industries have effectively integrated big data analytics into their decision-making processes to gain a competitive edge, enhance customer experiences, and drive business growth. In summary, the relationship between big data analytics and decision-making is a complex and multifaceted area of study. Big data analytics has a significant impact on decision-making, organisational performance, and strategic management. It is essential to explore this relationship further to address challenges and capitalise on opportunities in various fields.

Methodology

3.1 Sample and Data Collection

Quantitative methodology was employed to collect and analyse data in this study. In this study, a survey method is used to collect data through a structured questionnaire. Data were collected from employees in various organisations across different industries in the United Kingdom (UK). Microsoft Forms was used to distribute the questionnaire. Microsoft Forms creates surveys and collects responses online making it easier for data collection. It also allows users to see real-time responses and has built-in analytics to evaluate responses. After the required responses have been collected, users can export results to Excel, making it easier for the pre-processing of data. The online questionnaire link was disseminated to potential participants through online channels. The questionnaire was accessible to respondents for one month. The timeline spanned from October 2023 to November 2023. A total of 53 responses were received during the entire data collection period, all of which were deemed usable for analysis. As it was compulsory for participants to answer each question before moving on to the next one, this study did not have any missing data. Several steps were undertaken to limit common method bias. Firstly, the implemented questionnaire ensured the preservation of anonymity and confidentiality of the acquired data. Secondly, the systematic randomisation of items in the questionnaire effectively minimizes the respondents' ability to distinguish between independent and dependent variables. Respondents of this study were all volunteers. A statement of consent was also collected beforehand, in which they had to agree, for a respondent to proceed with the questionnaire.

3.2 Questionnaire and Measures

The questionnaire consists of 14 main questions – 9 of which were adapted from Shamin et al (2019) while the remaining 5 questions were demographic. The questions were modified to shift the perspective from a firm's standpoint to an individual's perspective for easier collection of data. Big Data Knowledge Exchange (BDKE), Big Data Collaboration (BDC), Process Integration (PI), Routinising (R), Flexible Infrastructure for Big Data (FIBD), and Leadership Focus on Big Data (LFBD) were the factors chosen to measure Big Data Analytics Capabilities. Big Data Decision Making Quality (BDDMQ), Decision-Making Effectiveness (DMEF) and Decision-Making Efficiency (DMEFF) were the factors chosen to measure Data-Driven

Decision-Making. Each factor had different items related to the factor. There was a total of 35 items for all the different factors. Responses to the survey items were measured using a Likert Scale. All items were measured using a five-point Likert scale. The five-point Likert scale used in this research contains 5 options, which are: Strongly Agree (5), Agree (4), Neither Agree or Disagree (3), Disagree (2), and Strongly Disagree (1). By including Neither Agree or Disagree in the Likert Scale allowed participants to take a neutral stand when answering the questions.

3.3 Data Cleaning and Preparation

Before running the regression models in Stata, the data set first had to be cleaned and adjusted accordingly before loading it into the software. The raw data set was first downloaded into Microsoft Excel. To get the value of a variable for one respondent, the average of the items was calculated to get the final value of the factor. For example, BDKE had 5 different items (See Appendix). The average of all items was calculated to give the value of one respondent's BDKE. This was repeated for all respondents and for all 9 different factors. This was first done in Microsoft Excel before the final data set used was imported into Stata. From the raw data set, only the type of industry, number of employees, the results of each item, and the calculated averages of all the variables – BDKE, BDC, PI, R, FIBD, LFBD, BDDMQ, DMEF, and DMEFF – were used in the final data set that was imported into Stata.

3.4 Data Analysis

To analyse the data, quantitative techniques were utilised such as correlation analysis and regression analysis. Correlation coefficient is a statistical calculation of any form of correlation which implies a statistical association between any two variables. The correlation used for this analysis is called Pairwise Correlation. It measures the degree of association of any two variables measured and looks for a linear relationship between them. Regression analysis is a mathematical method that aids in determining the relationship between an independent variable with one or more independent variables. The independent variables in the model are also known as predictors. Even though there are multiple forms of regression analysis present, they all primarily assess the impact of one or more independent variables on a dependent variable. The regression model used for this study was Ordinary Least Square (OLS) Regression. OLS Regression analysis was used to develop four different predictive models to help answer each of the four research questions. For both basic exploratory analysis, correlation analysis, and regression analysis, the Stata software package was utilised.

There were two relevant control variables from the survey that were included in the regression models. Organisation size, measured by the number of employees, was included as a control variable since bigger organisations are more likely to have access to more resources (Chen et al., 2014). The type of industry an individual was working in was also included as a control variable in the regression model. The importance of big data analytics usage within organisations may vary depending on the type of industry the individual is in (Ghasemaghaei et al., 2017). The number of employees and type of industry were coded as a dummy variables in Stata.

Results & Findings

4.1 Descriptive Analysis

Table 1 below shows the summary of the demographic characteristics of the respondents. These characteristics focus on age, education, type of industry and the number of employees in their organisation. More than 50 percent of the respondents are in the 18-25 age group (58.5 percent), 22.6 percent of respondents are between 26-35, 15.1 percent of respondents are between 36-50 and only 3.8% of the respondents are above the age of 50. The majority of the respondents have an undergraduate degree, accounting for 69.9 percent of the sample size. 17.0 percent of respondents have a Postgraduate Degree, 7.5 percent of respondents hold a diploma or certificate and 5.7 percent of the respondents have a secondary school qualification. The largest percentage (26.4 percent) of respondents are engaged in "Other Services activities" followed by Professional, scientific, and technical activities which account for 13.2 percent and Human Health and Social Work Activities accounting for 11.3 percent. 41.5 percent of the respondents surveyed are working in organisations that have over 2000 employees. 26.4 percent of respondents are working in organisations that have 201-500 employees, and 20.8 percent of the respondents are working in organisations which have less than 200 employees. 7.5 percent of respondents are working in organisations with 501-1000 employees and only 3.8 percent are working in organisations with 1001-2000 employees.

Table 1 Summary of demographics of the respondents

Demographics		Frequency	Percentage
Age	18-25	31	58.5%
	26-35	12	22.6%
	36-50	8	15.1%
	Above 50	2	3.8%
Education	Postgraduate Degree (Masters/PhD)	9	17.0%
	Undergraduate Degree	37	69.8%
	College (Diploma/Certificate)	4	7.5%
	Secondary School Qualification	3	5.7%
	Primary School Qualification	0	0.0%
Industry	Accommodation and Food Services activities	0	0.0%
	Administrative and support Service activities	1	1.9%
	Agriculture, forestry, and fishing	0	0.0%
	Arts, entertainment, and recreation	2	3.8%

	Construction	1	1.9%
	Education	5	9.4%
	Electricity gas, steam, and air conditioning supply	0	0.0%
	Financial and insurance activities	3	5.7%
	Human health and social work activities	6	11.3%
	Information and communication	5	9.4%
	Manufacturing	1	1.9%
	Mining and quarrying	0	0.0%
	Professional, scientific, and technical activities	7	13.2%
	Public Administration and Defence; compulsory social security	5	9.4%
	Real estate activities	0	0.0%
	Transportation and storage	1	1.9%
	Water Supply, Sewerage, waste management	0	0.0%
	Wholesale and Retail trade, repair of motor vehicles and motorcycles	2	3.8%
	Other Services activities	14	26.4%
Number of Employees	Less than 200	11	20.8%
	201-500	14	26.4%
	501-1000	4	7.5%
	1001-2000	2	3.8%
	Above 2000	22	41.5%

Table 2 below shows the summarised mean of each item related to the variables. From the table below, most of the items are averaging above a score of 3.5 which would mean that the majority of respondents agree or strongly agree with the questions, with the exception of some. It is important to note that the means of the items related to big data collaboration (BDC) is the lowest compared to the other items. The mean of the items related to big data collaboration is around 2.9 which is lower than the neutral score of 3. This means that the majority of respondents are disagreeing with the statements.

Table 2 Mean of Each Item

Item	Mean	Item	Mean
BDKE 1	3.64	LFBD 1	3.7
BDKE 2	3.64	LFBD 2	3.83
BDKE 3	3.71	LFBD 3	3.66

BDKE 4	3.43	LFBD 4	3.6
BDKE 5	3.83	LFBD 5	3.6
BDC 1	2.96	LFBD 6	3.45
BDC 2	2.92	BDDMQ 1	3.62
BDC 3	2.83	BDDMQ 2	3.62
PI 1	3.62	BDDMQ 3	3.34
PI 2	3.55	BDDMQ 4	3.6
PI 3	3.72	DMEF 1	3.87
R 1	3.15	DMEF2	3.85
R 2	3.81	DMEF 3	3.83
R 3	3.6	DMEF 4	3.85
FIBD 1	3.09	DMEFF 1	3.55
FIBD 2	3.75	DMEFF2	3.51
FIBD 3	3.81	DMEFF 3	3.68
FIBD 4	3.69		

Table 3 below shows the summarised mean of each variable that was measured. From the table below, the majority of the variables are averaging above the neutral score of 3 with the exception of big data collaboration. This means that overall, respondents generally possess most of the big data analytic capabilities. It also means that the respondents have good efficient and effective decision-making skills and in turn, produce good quality of big data decisions.

Table 3 Mean of each variable

Variable	Mean
Big Data Knowledge Exchance	3.65
Big Data Collaboration	2.91
Process Integration	3.63
Routinising	3.52
Flexible Infrastructure for Big Data	3.58
Leadership Focus for Big Data	3.63
Big Data Decision Making Quality	3.54
Decision-Making Effectiveness	3.85
Decision-Making Efficiency	3.58

4.2 Correlation Analysis

When checking for multicollinearity, a high correlation coefficient could help identify the two variables which are highly related to each other, which causes multicollinearity. Multicollinearity can lead to inflated standard errors of the regression coefficients, making them imprecise and less reliable. First, a correlation analysis was performed in Stata to identify the correlation between any two variables. The correlation analysis used to measure the correlation coefficients was Pairwise Correlation. In Stata, the Pairwise correlation a pair of variables is treated separately and observations that only have valid values for each pair in the data set are included. The linear relationship between the variables is calculated. The correlation matrix also explains the strength of the relationship between the two variables. From Table 2 below, the numbers in the matrix represent the correlation coefficients, and asterisks indicate whether the correlation is statistically significant at the 0.05 level. The correlation coefficient between any two variables has positive values which means that as one variable increases, the other tends to increase as well. Most correlation coefficients are significant at the 0.05 level except for the correlation coefficients between Big Data Collaboration and Routinising, Big Data Collaboration and Big Data Decision Making Quality, and Big Data Collaboration and Decision-Making Efficiency.

Next, a multicollinearity test was conducted by calculating variance inflation factors (VIF) for each regression coefficient. The Variance Inflation Factor (VIF) statistic is used for formative constructs to investigate whether the formative measures are highly correlated (Petter, Straub, & Rai, 2007). Table 3 below shows the summary of VIF values of each variable for all four regression models. All variables and overall VIF for the model have VIF values below the stringent threshold value of 10 (Johnston, Jones and Manley, 2017), which suggests that there is likely no multicollinearity between any of the variables of the models.

Table 4 Correlation Matrix

Sr.	Variable	1	2	3	4	5	6	7	8	9
1	BDKE	1								
2	BDC	0.4698*	1							
3	PI	0.5186*	0.3580*	1						
4	R	0.4378*	0.2524	0.5875*	1					
5	FIBD	0.3406*	0.3054*	0.6338*	0.6190*	1				
6	LFBD	0.6568*	0.5204*	0.4674*	0.5469*	0.4312*	1			
7	BDDMQ	0.3792*	0.2628	0.4339*	0.3394*	0.3170*	0.3678*	1		
8	DMEF	0.5378*	0.2806*	0.3907*	0.4150*	0.4680*	0.3823*	0.4392*	1	

9	DMEFF	0.3657*	0.1851	0.5010*	0.4565*	0.4022*	0.3931*	0.4647*	0.4336*	1

* Correlation is significant at 0.05 level.

Table 5 Summary table of VIF Values for each model

Model 1		Model 2		Model 3		Model 4	
Variable	VIF	Variable	VIF	Variable	VIF	Variable	VIF
BDKE	2.91	BDKE	2.91	DMEFF	1		1
BDC	2.27	BDC	2.27				
PI	2.58	PI	2.58				
R	3.35	R	3.35				
FIBD	2.89	FIBD	2.89				
LFDB	4.39	LFDB	4.39				
MEAN VIF	3.07		3.07		1		1

4.3 Regression Analysis

OLS regression, or Ordinary Least Squares regression, is a statistical technique used to estimate the relationship between one or more independent variables and a dependent variable. It is commonly used to predict outcomes based on input variables. OLS regression minimizes the sum of the squared differences between the observed and predicted values, making it a widely used optimisation strategy for linear regression models. The steps to run the regression analysis were to first enter the dependent variable and then the independent variables. The same steps were repeated four times to run the regression for each research question.

Table 3 below shows the OLS regression model summary of the first research question (R2). The dependent variable for this model is Decision Making Effectiveness. The independent variables are Big Data Knowledge Exchange, Big Data Collaboration, Process Integration, Routinising, Flexible Infrastructure for Big Data, and Leadership Focus on Big Data. The control variables are the type of industry and the number of employees. The baseline model for Model 1 is as shown below:

$$\text{Decision Making Effectiveness}_{it} = \beta_0 + \beta_1 \times \text{BDKE}_{it} + \beta_2 \times \text{BDC}_{it} + \beta_3 \times \text{PI}_{it} + \beta_4 \times \text{R}_{it} + \beta_5 \times \text{FIBD}_{it} + \beta_6 \times \text{LFBD}_{it} + \beta_j \times \text{Industry}_{it} + \beta_k \times \text{NoofEmployees}_{it} + \varepsilon_{it}$$

Model 1 has an F-statistic of 0.6113 with a probability (Prob > F) of 0.0196. As this probability is less than 0.05, the model is statistically significant at the 5% level. This suggests that the group of independent variables show a statistically significant relationship and reliably predict the dependent variable, Decision Making Effectiveness. This also suggests that at least one of the independent variables in the model is a significant predictor of the dependent variable. The value of R-squared for R2 is 0.6372, which means 63.7% of the variance can be explained from the independent variables. The coefficients of independent variables that are statistically significant at 0.05 significance level are Big Data Knowledge Exchange and Flexible Infrastructure for Big Data, with p-values of 0.01 and 0.038 respectively.

Big Data Knowledge Exchange has a positive effect and statistically significant impact on an individual's Decision-Making Effectiveness, all other variables being equal (β= 0.3423, p-value <0.05). One unit increase in Big Data Knowledge Exchange increases an individual's decision-making effectiveness by an average of 0.342. Thus, it can be rejected that this impact equals to zero (i.e., the null hypothesis) at a 0.05 significance level or with more than 95% probability. Flexible Infrastructure for Big Data has a positive effect on an individual's Decision-Making Effectiveness (β= 0.3063, p-value <0.05), *ceteris paribus*. This impact is also statistically significant at a 0.05 significance level. A one unit increase in Flexible Infrastructure for Big Data increases an individual's Decision-Making Effectiveness by an average of 0.3063.

Table 6 Model Summary for Research Question 1

Prob > F	0.0134	R-squared	0.6372	Adj. R-squared	0.3712	
Variable	Coefficients	Standard Error	t	p	[95% conf. interval]	
BDKE	0.3423	0.1249	2.74	0.01	0.0871	0.5976
BDC	0.5206	0.0926	0.56	0.579	-0.1376	0.2417
PI	-0.0072	0.1201	-0.06	0.952	-0.2526	0.238
R	0.0177027	0.1379	0.13	0.899	-0.2639	0.2993
FIBD	0.3063	0.1413	2.17	0.038	0.0176	0.5949
LFBD	-0.01468	0.1296	-0.092	0.365	-0.4728	0.1792

Table 4 below shows the model summary for research question 2 (R2). The dependent variable for this model is Decision Making Effectiveness. The independent variables are Big Data Knowledge Exchange, Big Data Collaboration, Process Integration, Routinising, Flexible Infrastructure for Big Data, and Leadership Focus on Big Data. The control variables are the type of industry and the number of employees. The baseline model for Model 2 is as shown below:

$$\text{Decision Making Efficiency}_{it} = \beta_0 + \beta_1 \times BDKE_{it} + \beta_2 \times BDC_{it} + \beta_3 \times PI_{it} + \beta_4 \times R_{it} + \beta_5 \times FIBD_{it} + \beta_6 \times LFBD_{it} + \beta_j \times Industry_{it} + \beta_k \times NoofEmployees_{it} + \varepsilon_{it}$$

Model 2 has an F-statistic of 0.6113 with a probability (Prob > F) of 0.0196. As this probability is less than 0.05, the model is statistically significant at the 5% level. This suggests that the group of independent variables show a statistically significant relationship and reliably predict the dependent variable, Decision Making Efficiency. This also suggests that at least one of the independent variables in the model is a significant predictor of the dependent variable. The value of R-squared for R3 is 0.6230, which means 62.3% of the variance can be explained from the independent variables. The only coefficient of independent variables that is statistically significant at a 0.05 significance level is Process Integration.

Process Integration has a positive effect on Decision Making Efficiency and a statistically significant impact on an individual's Decision-Making Efficiency, all other variables being equal (β= 0.3612, p-value <0.05). One unit increase in Process Integration increases an individual's decision-making efficiency by an average of 0.3612. Thus, it can be rejected that this impact equals to zero (i.e., the null hypothesis) at a 0.05 significance level or with more than 95% probability.

Table 7 Model Summary for Research Question 2

Prob > F	0.0196	R-squared	0.623	Adj. R-squared	0.3465	
Variable	Coefficients	Standard Error	t	p	[95% conf. interval]	
BDKE	0.0501	0.1514	0.33	0.743	-0.2591	0.3593
BDC	0.0503	0.1124	0.45	0.658	-0.1794	0.28

PI	0.3612	0.1455	2.48	0.019	0.0641	0.6583
R	0.1927	0.167	1.15	0.258	-0.1485	0.5338
FIBD	-0.9073	0.1711	-0.53	0.6	-0.4406	0.2588
LFBD	-0.0797	0.1933	-0.41	0.683	-0.4745	0.3881

Table 5 below shows the OLS regression of the third research question (R3). The dependent variable for this model is Big Data Decision Making Quality and the independent variable is Decision Making Effectiveness. The baseline model for Model 3 is as shown below:

$$\text{Big Data Decision Making Quality}_{it} = \beta_0 + \beta_1 \times \text{DMEF}_{it} + \varepsilon_{it}$$

The R-squared value for R3 is 0.1929, indicating that approximately 19.29% of the variance in the dependent variable is explained by the independent variable in the model. The probability associated with the F-statistic is 0.001. As this probability is less than 0.05, the model is statistically significant at the 5% level. The individual coefficient for the variable Decision-Making Effectiveness is statistically significant. The coefficient for the variable Decision-Making Effectiveness is 0.601 and is statistically significant at a 0.05 significance level, with a p-value of 0.001, which is less than 0.05. Decision-Making Effectiveness has a positive effect and a statistically significant impact on an individual's Big Data Decision Making Quality. One unit increase in Decision-Making Effectiveness increases an individual's Big Data Decision Making Quality by an average of 0.601. Thus, it can be rejected that this impact equals to zero (i.e., the null hypothesis) at a 0.05 significance level or with more than 95% probability.

Table 8 Model Summary for Research Question 3

Prob > F	0.001	R-squared	0.1929	Adj. R-squared	0.1771	
Variable	Coefficients	Standard Error	t	p	[95% conf. interval]	
DMEF	0.601	0.172	3.49	0.001	0.255	0.946

Table 6 below shows the OLS regression of the fourth research question (R4). The dependent variable for this model is Big Data Decision Making Quality and the independent variable is Decision Making Efficiency. The baseline model for Model 4 is as shown below:

$$\text{Big Data Decision Making Quality}_{it} = \beta_0 + \beta_1 \times \text{DMEFF}_{it} + \varepsilon_{it}$$

The R-squared value for R4 is 0.2159, indicating that approximately 21.5% of the variance in the dependent variable is explained by the independent variable in the model. The probabilitiy associated with the F-statistic is 0.0005. As this probability is less than 0.05, the model is statistically significant at the 5% level. The individual coefficient for the variable Decsion-Making Efficiency is statistically significant. The coefficient for the variable Decision-Making Efficiency is 0.535 and is statistically significant at a 0.05 significance level. Decision-Making Efficiency has a positive effect and a statistically significant impact on an individual's Big Data Decision Making Quality. One unit increase in Decision-Making Efficiency increases an individual's Big Data Decision Making Quality by an average of 0.535. Thus, it can be rejected that this impact equals to zero (i.e., the null hypothesis) at a 0.05 significance level or with more than 95% probability.

Table 9 Model Summary for Research Question 4

Prob > F	0.0005	R-squared	0.2159	Adj. R-squared	0.201	
Variable	Coefficients	Standard Error	t	p	[95% conf. interval]	
DMEF	0.535	0.142	3.75	0	0.248	0.821

Analysis and Discussion

5.1 Theoretical Contribution

This study investigated how an individual's data-driven decision-making can be influenced by their big data analytics capabilities by gathering data from 53 working individuals. The findings of this dissertation contribute significantly to the understanding of the relationship between various factors related to the big data capabilities of an individual and their impact on decision-making effectiveness, decision-making efficiency, and big data decision-making quality within organisational settings.

The findings of this study suggest that there is a significant positive association between an individual's big data analytics capability and the decision-making effectiveness of an individual. The results of these findings align with the view of Redman (2008), who states that leveraging data and information is the most effective way for anyone to make and execute better and faster decisions. Similarly, as Yiu (2012) suggested, with the help of big data, organisations can enhance their data-driven decision-making in terms of effectiveness and efficiency. In particular, big data knowledge exchange and flexible infrastructure for big data have the most significant positive association with decision-making effectiveness among all the big data analytics capabilities measured. As Abubakar (2019) states, the role of knowledge in resources is important in decision-making. Also, as Mikalef et al. (2020) state analytical applications have gained significance in evidence-based decision-making, given the growing reliance on BD in organisational decision-making. The results from this study support the views of both Abubakat et al. (2019) and Mikalef et al. (2020). The positive impact of big data knowledge exchange on decision-making effectiveness could imply that decision-makers who are actively engaged in sharing and acquiring knowledge about data are more likely to make effective decisions. Moreover, the significance of flexible infrastructure for big data indicates that having access to a flexible and adaptive infrastructure, that allows for a better big data analysis of an individual, positively impacts an individual's decision-making effectiveness. Individuals are able to speed up their decision-making process and make more informed decisions by effectively utilising the power of big data tools when they are in a technologically supportive environment. The positive association of big data knowledge exchange and the flexible infrastructure of big data may suggest that decision-makers with a higher level of knowledge exchange of big and access to a more flexible infrastructure of big data are more likely to demonstrate more effective decision-making skills. The study also sheds light on the relationship between an individual's big data analytics capability and their decision-making efficiency. There is a significant positive association between an individual's big data analytics capability and their decision-making efficiency. Notably, process integration is the sole significant contributor to an individual's decision-making efficiency. Individuals who are able to understand the process involved in the big data chain better showcase a higher decision-making efficiency. When an individual understands the process involved in the big data chain, it allows them to navigate through tasks more efficiently, thus leading to quicker decision-making. The findings can be supported by Shamin et al's (2019) view which states that having access to better data and information in advance and utilising them, allows individuals to make prompt and informed decisions, and thus, operate with greater efficiency.

The findings of this study can also be supported by Ahmed's et al. (2022) views where the author states that efficiency and effectiveness have a direct impact on the quality of decisions made. The relationship between an individual's decision-making effectiveness and their big data decision-making quality was also studied. There is a positive association between decision-making effectiveness and big data decision-making quality. When there is an increase in an individual's decision-making effectiveness, there is a corresponding improvement in the quality of their decisions in the context of big data. The positive impact of the decision-making effectiveness on big data decision-making quality could suggest that effective decision-making contributes significantly to the overall quality of big data-driven decisions on an individual level.

In this study, it was also established that an individual's efficient decision-making and their big data decision-making quality were also studied. There is a positive association between decision-making efficiency and big data decision-making quality. This finding suggests that the positive impact of a decision-maker's efficient decision-making on the quality of their decision-making in the context of big data highlights the importance of fast and informed decision-making processes in producing high-quality results. Efficient decision-making ensures that decisions are made promptly in the context of big data. This also suggests that not only is effective decision-making important, but the efficiency with which decisions are made also plays a crucial role in enhancing the overall quality of big data-driven decisions. This aligns with the argument made by Shamin et al. (2019) who state that the achievement of better results from data-driven decision-making and its effectiveness, is the outcome of quality decision-making.

The importance of an individual's big data analytics capabilities, and decision-making measures in terms of effectiveness, efficiency, and quality, highlights how connected these elements are in the field of big data analytics. All the findings point to the necessity of a comprehensive strategy in which human strengths and effective procedures combine to improve decision-making efficiency and effectiveness, ultimately raising the calibre of big data-driven choices.

The findings of this study contribute to theoretical frameworks that emphasise on the relationship between an individual's big data analytic capabilities, decision-making effectiveness, decision-making efficiency, and the quality of big data-driven decision-making. This study also addresses the importance of considering big data analytics capabilities at an individual level that aligns with the growing body of literature which acknowledges the multifaceted dynamics of decision-making in the current digital era. It bridges the gap of the lack of research by demonstrating the big data analytics capabilities of individuals in influencing their effectiveness and efficiency when making a decision. With most of the current literature being studied from an organisational level, integrating the perspective at an individual level could provide a deeper understanding of decision-making in the era of big data.

5.2 Practical Implications

The results of this study hold practical implications for organisations that are seeking to leverage big data in their decision-making for their employees. By emphasising on the importance of big data knowledge exchange and ensuring a flexible technological infrastructure for big data can enhance an individual's decision-making effectiveness. Organisations should encourage their employees' knowledge exchange on big data and provide training programs for employees to understand data knowledge exchange in their organisations. Fostering the culture of big data knowledge exchange within the company would help enhance an individual's big data analytics capabilities and hence, decision-making effectiveness in any data-driven environment. Also, the significance of flexible infrastructure for big data highlights the importance of a well-rounded technological environment for any organisation.

Organisations should invest in flexible and adaptive technological infrastructures which allow the seamless integration of big data analytic tools for employees to enhance the effectiveness of their decision-making. Moreover, process integration is crucial for improving an individual's decision-making efficiency and hence, has an impact on the quality of big data-driven decisions. In order to improve the decision-making efficiency of the employees within an organisation, these organisations should focus on optimising and integrating the process of their big data chain. Their employees should have a good understanding of how these processes work which will allow them to have seamless communication and coordination with other decision-makers in their organisations. Through these efforts, organisations are able to reap the benefits of the improvement in their employees' decision-making efficiency. Overall, organisations can leverage the findings to design targeted training programs that enhance individual capabilities in big data analytics to improve decision-making efficiency and effectiveness.

Fostering the culture of big data knowledge exchange, investing in technologies that allow flexible infrastructure for big data, and optimising process integration can enhance both the effectiveness and efficiency of an individual's decision-making altogether. Both factors, decision-making effectiveness, and decision-making efficiency, have an influence on an individual's big data decision-making quality. Organisations that are seeking to develop and nurture these capabilities among their employees to improve the overall quality of decision-making in the context of big data. Organisations should prioritise initiatives that would enhance these factors among employees. This could include having mentorship programs with experienced decision-makers who can provide support to other decision-makers which cultivates an environment that nurtures collaboration among decision-makers and make fast decisions. Organisations aiming to improve the quality of big data decisions made by their employees should not only consider strategies that only enhance decision-making effectiveness but also their decision-making efficiency. This includes investing in technologies that can complete data analysis and tasks quickly, and can provide employees with the necessary tools to advance their decision-making process.

The theoretical and practical implications resulting from this study highlight the need for organisations to adopt a dynamic and integrated approach to decision-making in the era of big data. There is a need for an organisational strategy that is comprehensive and integrates investments in process optimisation, technology infrastructure, and knowledge exchange.

Organisations will be in a better position to handle the intricacies of big data and use it to make smarter decisions when incorporating these components into their strategy planning. This entails not just offering the required tools and technology but also fostering an attitude that prioritises constant learning, and flexibility among employees. This culture would guarantee that decision-making outcomes would stay high-quality, efficient, and effective. By improving the big data analytics capabilities of their employees, it would improve their decision-making effectiveness and efficiency, and in turn, improve the quality of big data decision-making. Not only does this approach enhance an individual's decision results, but it also contributes to the overall performance and competitiveness of organisations as a whole in the current big data era.

Conclusion

6.1 Conclusion

This study focused on big data analytics capabilities and their impact on the data-driven decision-making of individuals in an organisation. This paper offers insights into the interconnectedness between the big data analytics capability, decision-making effectiveness, decision-making efficiency, and the quality of big data decision-making among individuals of various organisations. The results of this paper provide an in-depth comprehension of the various factors that impact an individual's decision-making process in the context of big data. The finding highlights the positive influence of big data analytics capabilities on data-driven decision-making in terms of its effectiveness and efficiency, and in turn, the quality of big data decision-making. The study provides empirical evidence supporting the significance of big data analytics capabilities in boosting an individual's decision-making effectiveness, efficiency, and the quality of big data-driven decisions. The identified big data analytic capabilities from the results can operate as a guide for organisations that are facing the challenges of data-driven decision-making. Organisations should seek to harness the full potential of big data for improved decision outcomes. It enables organisations to create new products and services, enhance existing ones, and create new business models. Optimising one's big data analytic capability can help individuals and organisations improve the overall quality of decisions made in the field of big data analytics. By doing so, organisations are able to improve performance and have a competitive advantage in their industry and the economy. The theoretical and practical implications that were presented in this paper further contribute to the existing literature on the evolving world of big data analytics and data-driven decision-making. By addressing the gaps in existing literature and offering useful perspectives, this study will help future research endeavours to better understand the intricacies of big data and data-driven decision-making processes within organisations.

6.2 Limitations and Future Research Areas

Despite the valuable insights gained, this study is not without limitations. The reliance on self-reported data and the cross-sectional nature of the study imposes limitations on causal inferences. Future research could employ longitudinal designs and explore additional contextual factors to further enrich the understanding of the dynamics between big data and decision-making among individuals in organisations. One significant limitation of this study is the reliance on a data set characterised by a comparatively low sample size. Only 53 responses were collected for this study. The limited number of observations may constrain the generalisability of the findings and compromise the statistical power of the analyses. While every effort has been made to draw meaningful conclusions from the available data, the constrained sample size poses challenges in achieving a comprehensive understanding of the topic of this study. A longer time frame to collect data could also be implemented to achieve more results to obtain a more diverse and accurate representation. Future research endeavours with larger and more diverse samples are recommended to enhance the robustness and external validity of the study's outcomes. A larger sample size would help improve the accuracy of the model.

Another limitation of this study is the exclusive reliance on quantitative research methods. While quantitative approaches provide valuable statistical insights into relationships and patterns within the data, they inherently lack the depth and nuance that qualitative methods can offer. The utilisation of only quantitative measures may overlook valuable contextual information and fail to capture the intricacies of this study. The absence of qualitative data also limits the ability to explore participants' views, experiences, or other unforeseen factors that could contribute to a more comprehensive understanding of this study. Future research endeavours might consider incorporating qualitative methodologies alongside quantitative approaches to provide a more holistic and nuanced exploration of the four research questions proposed, cultivating a broader interpretation of the study's outcomes.

In the future, the same analysis can be performed with more in-depth questions based on different regions of the world. As this research was only done in the United Kingdom (UK), it only represents the. Results could differ based on different countries or regions of the world.

Reference List

Abubakar, A.M., Elrehail, H., Alatailat, M.A. and Elçi, A. (2019). Knowledge management, decision-making style and organizational performance. *Journal of Innovation & Knowledge*, [online] 4(2), pp.104–114. doi:https://doi.org/10.1016/j.jik.2017.07.003.

Ahmed, R., Shaheen, S. and Philbin, S.P. (2022). The Role of Big Data Analytics and Decision-Making in Achieving Project Success. *Journal of Engineering and Technology Management*, 65(65), p.101697. doi:https://doi.org/10.1016/j.jengtecman.2022.101697.

Awan, U., Shamim, S., Khan, Z., Zia, N.U., Shariq, S.M. and Khan, M.N. (2021). Big data analytics capability and decision-making: The role of data-driven insight on circular economy performance. *Technological Forecasting and Social Change*, 168, p.120766. doi:https://doi.org/10.1016/j.techfore.2021.120766.

Batistič, S. and der Laken, P. (2019). History, Evolution and Future of Big Data and Analytics: A Bibliometric Analysis of Its Relationship to Performance in Organizations. *British Journal of Management*, 30(2), pp.229–251. doi:https://doi.org/10.1111/1467-8551.12340.

Björkman, F., Franco, S. and Caesarius, L. (2017). *How Big Data Analytics Affect Decision-Making: A Study of the Newspaper Industry*. [online] Available at: https://www.diva-portal.org/smash/get/diva2:1110878/FULLTEXT01.pdf.

Boyd, D. and Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15(5), pp.662–679.

Cavlak, N. and Cop, R. (2021). *The Role of Big Data in Digital Marketing*. [online] https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-7998-8003-5.ch002. Available at: https://www.igi-global.com/gateway/chapter/284279.

Chatterjee, S., Chaudhuri, R., Gupta, S., Sivarajah, U. and Bag, S. (2023). Assessing the impact of big data analytics on decision-making processes, forecasting, and performance of a firm.

Technological Forecasting and Social Change, 196, p.122824. doi:https://doi.org/10.1016/j.techfore.2023.122824.

Chen, H., Chiang, R.H.L. and Storey, V.C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, [online] 36(4), p.1165. doi:https://doi.org/10.2307/41703503.

Clark, T.D., Jones, M.C. and Armstrong, C.P. (2007). The Dynamic Structure of Management Support Systems: Theory Development, Research Focus, and Direction. *MIS Quarterly*, [online] 31(3), pp.579–615. doi:https://doi.org/10.2307/25148808.

Dell Technologies. (2020). *Dell Technologies Research: Cyber Attacks and Disruptive Events Are on the Rise, Affecting 82% of Organizations Surveyed | Dell Technologies*. [online] Available at: https://investors.delltechnologies.com/news-releases/news-release-details/dell-technologies-research-cyber-attacks-and-disruptive-events [Accessed 5 Nov. 2023].

Di Berardino, D. and Vona, S. (2023). Discovering the Relationship between Big Data, Big Data Analytics, and Decision Making: a Structured Literature Review. *European Scientific Journal, ESJ*, 19(19), pp.1–1. doi:https://doi.org/10.19044/esj.2023.v19n19p1.

Dietrich, D., Heller, B. and Yang, B. (2021). *Data Science & Big Data Analytics : Discovering, Analyzing, Visualizing and Presenting Data*. Indianapolis, In: Wiley.

Dubey, R., Gunasekaran, A., Childe, S.J., Roubaud, D., Fosso Wamba, S., Giannakis, M. and Foropon, C. (2019). Big data analytics and organizational culture as complements to swift trust and collaborative performance in the humanitarian supply chain. *International Journal of Production Economics*, 210, pp.120–136. doi:https://doi.org/10.1016/j.ijpe.2019.01.023.

Economist Intelligence Unit. (2012). The deciding factor: Big data & decision making. In

C.G. reports (Ed.), (pp. 24): Economist Intelligence Unit.

Elgendy, N. and Elragal, A. (2014). Big Data Analytics: a Literature Review Paper. *Advances in Data Mining. Applications and Theoretical Aspects*, [online] 8557(1), pp.214–227. doi:https://doi.org/10.1007/978-3-319-08976-8_16.

Elgendy, N. and Elragal, A. (2016). Big Data Analytics in Support of the Decision Making Process. *Procedia Computer Science*, [online] 100, pp.1071–1084. doi:https://doi.org/10.1016/j.procs.2016.09.251.

Gandomi, A. and Haider, M. (2015). Beyond the Hype: Big Data Concepts, Methods, and Analytics. *International Journal of Information Management*, 35(2), pp.137–144.

Garmaki, M., Boughzala, I. and Fosso Wamba, S. (2016). *THE EFFECT OF BIG DATA ANALYTICS CAPABILITY ON FIRM PERFORMANCE*.

Grušovnik, D., Kavkler, A. and Uršič, D. (2017). Dimensions of Decision-Making Process Quality and Company Performance: A Study of Top Managers in Slovenia. *Naše gospodarstvo/Our economy*, 63(4), pp.66–75. doi:https://doi.org/10.1515/ngoe-2017-0024.

Ikegwu, A.C., Nweke, H.F., Anikwe, C.V., Alo, U.R. and Okonkwo, O.R. (2022). Big Data Analytics for data-driven industry: a Review of Data sources, tools, challenges, solutions, and Research Directions. *Cluster Computing*, 25(3343–3387). doi:https://doi.org/10.1007/s10586-022-03568-5.

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

Jeble, S., Kumari, S. and Patil, Y. (2018). Role of Big Data in Decision Making. *Operations and Supply Chain Management: An International Journal*, 11(1), pp.36–44. doi:https://doi.org/10.31387/oscm0300198.

Johnston, R., Jones, K. and Manley, D. (2017). Confounding and Collinearity in Regression analysis: a Cautionary Tale and an Alternative procedure, Illustrated by Studies of British Voting Behaviour. *Quality & Quantity*, 52(4), pp.1957–1976. doi:https://doi.org/10.1007/s11135-017-0584-6.

Kaltoft, M., Cunich, M., Salkeld, G. and Dowie, J. (2013). Assessing decision quality in patient-centred care requires a preference-sensitive measure. *Journal of Health Services Research & Policy*, 19(2), pp.110–117. doi:https://doi.org/10.1177/1355819613511076.

Khine, P.P. and Shun, W.Z. (2017). Big Data for Organizations: A Review. *Journal of Computer and Communications*, [online] 5(3), pp.40–48. doi:https://doi.org/10.4236/jcc.2017.53005.

Kościelniak, H. and Puto, A. (2015). BIG DATA in Decision Making Processes of Enterprises. *Procedia Computer Science*, 65(65), pp.1052–1058. doi:https://doi.org/10.1016/j.procs.2015.09.053.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S. and Kruschwitz, N. (2010). *Big Data, Analytics and the Path From Insights to Value*. [online] MIT Sloan Management Review. Available at: https://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/.

Malakooti, B. (2010). Decision making process: typology, intelligence, and optimization. *Journal of Intelligent Manufacturing*, 23(3), pp.733–746. doi:https://doi.org/10.1007/s10845-010-0424-1.

Manyika, J. and Chui, M. (2011). *Big data: The next frontier for innovation, competition, and productivity | McKinsey*. [online] www.mckinsey.com. Available at: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation.

McAfee, A. and Brynjolfsson, E. (2012). *Big Data: The Management Revolution*. [online] Harvard Business Review. Available at: https://hbr.org/2012/10/big-data-the-management-revolution.

Mishra, B.P., Bibhuti Bhusan Biswal, Ajay Behera and Harish Chandra Das (2023). Effect of Big Data Analytics on Performance and Consumer Loyalty of Manufacturing Industries: An Empirical Research. *Journal of Industrial and Integration and Management*, 8(3), pp.341–355. doi:https://doi.org/10.1142/s2424862223500112.

Muniswamaiah, M., Agerwala, T. and Tappert, C. (2019). Big Data in Cloud Computing Review and Opportunities. *International Journal of Computer Science and Information Technology*, 11(4), pp.43–57. doi:https://doi.org/10.5121/ijcsit.2019.11404.

Neirotti, P., Pesce, D. and Battaglia, D. (2021). Algorithms for operational decision-making: An absorptive capacity perspective on the process of converting data into relevant knowledge.

Technological Forecasting and Social Change, 173, p.121088. doi:https://doi.org/10.1016/j.techfore.2021.121088.

Olszak, C.M. (2016). Toward Better Understanding and Use of Business Intelligence in Organizations. *Information Systems Management*, 33(2), pp.105–123. doi:https://doi.org/10.1080/10580530.2016.1155946.

Redman, T.C. (2008). *Data Driven : Profiting from Your Most Important Business Asset*. Boston, Mass.: Harvard Business Press, Cop, pp.36, 89–90.

Roy, K. (2021). *7 Steps to Convert Big Data Into Actionable Intelligence*. [online] DataToBiz. Available at: https://www.datatobiz.com/blog/big-data-into-actionable-intelligence/.

Sabharwal, R. and Miah, S.J. (2021). A new theoretical understanding of big data analytics capabilities in organizations: a thematic analysis. *Journal of Big Data*, 8(1), pp.1–17. doi:https://doi.org/10.1186/s40537-021-00543-6.

Sabharwal, S., Gupta, S. and Thirunavukkarasu, K. (2016). *Insight of Big Data Analytics in Healthcare Industry*. [online] IEEE Xplore. doi:https://doi.org/10.1109/CCAA.2016.7813696.

Saggi, M.K. and Jain, S. (2018). A survey towards an integration of big data analytics to big insights for value-creation. *Information Processing & Management*, 54(5), pp.758–790. doi:https://doi.org/10.1016/j.ipm.2018.01.010.

Shamim, S., Zeng, J., Shariq, S.M. and Khan, Z. (2019). Role of big data management in enhancing big data decision-making capability and quality among Chinese firms: A dynamic capabilities view. *Information & Management*, 56(6), p.103135. doi:https://doi.org/10.1016/j.im.2018.12.003.

Thirathon, U. (2016). PERFORMANCE IMPACTS OF BIG DATA ANALYTICS. *PACIS 2016 Proceedings*. [online] Available at: https://aisel.aisnet.org/pacis2016/187/ [Accessed 28 Nov. 2023].

Tseng, H.-T., Aghaali, N. and Hajli, D.N. (2022). Customer agility and big data analytics in new product context. *Technological Forecasting and Social Change*, [online] 180(C). Available at: https://ideas.repec.org/a/eee/tefoso/v180y2022ics0040162522002177.html [Accessed 7 Jan. 2024].

Visinescu, L.L., Jones, M.C. and Sidorova, A. (2016). Improving Decision Quality: The Role of Business Intelligence. *Journal of Computer Information Systems*, 57(1), pp.58–66. doi:https://doi.org/10.1080/08874417.2016.1181494.

Vittinghoff, E. (2012). *Regression Methods in Biostatistics : linear, logistic, survival, and Repeated Measures Models*. New York: Springer.

Waller, M.A. and Fawcett, S.E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34(2), pp.77–84.

Wamba, S.F., Gunasekaran, A., Akter, S., Ren, S.J., Dubey, R. and Childe, S.J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70(1), pp.356–365. doi:https://doi.org/10.1016/j.jbusres.2016.08.009.

Wang, L. and Alexander, C.A. (2019). Big Data Analytics in Healthcare Systems. *International Journal of Mathematical, Engineering and Management Sciences*, 4(1), pp.17–26. doi:https://doi.org/10.33889/ijmems.2019.4.1-002.

Woerner L.S., Wixom H.B., (2015). "Big data: extending the business strategy toolbox." Journal of Information Technology, 30, 60-62

Xie, Z. (2021). Development Model of Business Management Based on Big Data. *E3S Web of Conferences*, 275, p.01041. doi:https://doi.org/10.1051/e3sconf/202127501041.

Yousuf, H. (2020). Quantitative Approach in Enhancing Decision Making Through Big Data as An Advanced Technology. *Advances in Science, Technology and Engineering Systems Journal*, 5(5), pp.109–116. doi:https://doi.org/10.25046/aj050515.

Appendix

A. Questionnaire

Welcome!

This research study is being done to understand the factors that contribute to big data analytics capability and data driven decision making. Taking part in the study will take about 5 minutes or less. You cannot take part in this study if you are under 18.

What will I be asked to do if I am in this study?
If you take part in this study, you will be asked to answer questions that relate to big data analytics capability and data driven decision making.

Are there any benefits to me if I am in this study?
There is no direct benefit to you from being in this study.

Are there any risks to me if I am in this study?
The potential risks from taking part in this study are limited to minimal inconveniences associated with answering multiple questions.

Will my information be kept private?
The data for this study will be kept confidential to the extent allowed by law. The results of this study maybe published in academic journals or presented at professional meetings, but the identities of all research participants will remain anonymous.

Are there any costs or payments for being in this study?
There will be no costs to you for taking part in this study.

Who can I talk to if I have questions?
If you have any questions regarding this study, please contact the researcher (Hanis Rizal): hm01440@surrey.ac.uk

What are my rights as a research study volunteer?
Your participation in this research study is completely voluntary. You may choose not to be a part of this study. There will be no penalty to you if you choose not to take part. You may choose not to answer specific questions or to stop participating at any time without needing to justify your decision, without prejudice, and without your legal rights and studies/employment being affected.

What does clicking "Yes" after the statement of consent mean?
Clicking "Yes" after the statement of consent means that:
You understand the information given to you in this form.
You believe you understand the research study and the potential benefits and risks that are involved.

* Required

1. Statement of Consent *

- ◯ Yes
- ◯ No

2. Age *

- 18-25
- 26-35
- 36-50
- Above 50

3. Gender *

- Male
- Female
- Other

4. Education *

- Primary School Qualification
- Secondary School Qualification
- College (Diploma/Certificate)
- Undergraduate Degree
- Postgraduate Degree (Masters/PhD)

5. Industry *

- ◯ Accommodation and Food Services 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 activites
- ◯ 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 and motorcycles
- ◯ Other Services activities

6. Number of Employees *

- ◯ Less than 50
- ◯ 50-200
- ◯ 201-1000
- ◯ 1001-2000
- ◯ Above 2000

7. Big Data Knowledge Exchange *

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
I transfer my knowledge about data.	○	○	○	○	○
Knowledge about how data are collected is exchanged within my organisation.	○	○	○	○	○
Knowledge about how data are used is exchanged within my organisation.	○	○	○	○	○
Knowledge about how data are processed is exchanged within my organisation.	○	○	○	○	○
The exchange of knowledge makes it easy for me to analyse data.	○	○	○	○	○

8. Big Data Collaboration *

	Strongly Disgree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
As decision maker, I collaborate with big data analysts and big data providers.	○	○	○	○	○
There is collaboration among myself and big data analysts.	○	○	○	○	○
There is collaboration among myself and big data providers.	○	○	○	○	○

9. Process Integration *

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
I have the ability to understand the process involved in the big data chain (ie. data collection, preparation, analysis and decision making).	◯	◯	◯	◯	◯
The understanding of the process involved in the big data chain reduces the cost of big data use.	◯	◯	◯	◯	◯
The understanding of the process involved in the big data chain reduces the efforts necessary to analyse big data myself.	◯	◯	◯	◯	◯

10. Routinising *

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
The big data chain is a routine matter for me (ie data collection, preparation, analysis and decision making).	◯	◯	◯	◯	◯
Routinising big activities improves big data velocity.	◯	◯	◯	◯	◯
Routinising big data activities helps me to make real time decisions.	◯	◯	◯	◯	◯

11. Flexible Infrastructure For Big Data *

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
My company's big data management infrastructure is flexible.	○	○	○	○	○
A flexible infrastructure helps me to enhance my ability to handle big data.	○	○	○	○	○
A flexible infrastructure helps me to enhance my ability to process big data.	○	○	○	○	○
Because of a flexible big data infrastructure, my decision making is quick.	○	○	○	○	○

12. Big Data Decision Maker Quality *

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
I am able to interpret the outcomes of big data analytics.	○	○	○	○	○
I understand the implications of big data analytics outcomes.	○	○	○	○	○
I am an experienced decision maker.	○	○	○	○	○
I have the ability to make quick decisions.	○	○	○	○	○

13. Leadership Focus on Big Data *

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
My leader provides a clear vision.	◯	◯	◯	◯	◯
My leader sets clear goals.	◯	◯	◯	◯	◯
My leader encourages big data decision making.	◯	◯	◯	◯	◯
My leader shows great interest in the big data chain.	◯	◯	◯	◯	◯
My leader shows concern for the use of big data.	◯	◯	◯	◯	◯
My leader is very active in managing big data.	◯	◯	◯	◯	◯

14. Decision-Making Effectiveness *

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
I believe that I make good decisions	◯	◯	◯	◯	◯
The decisions that I make result in the desired outcomes.	◯	◯	◯	◯	◯
I am satisfied with the outcomes of my decisions.	◯	◯	◯	◯	◯
My decisions improves organisationa l performance	◯	◯	◯	◯	◯

15. Decision-Making Efficiency *

	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
My decision-making cost is very low.	◯	◯	◯	◯	◯
I spend a lot of time in making decisions.	◯	◯	◯	◯	◯
I am able to make timely decisions.	◯	◯	◯	◯	◯

B. Programming Codes

```
//exploratory analysis
graph bar (mean) BDKE1 (mean) BDKE2 (mean) BDKE3 (mean) BDKE4 (mean) BDKE5,
blabel(bar)
graph bar (mean) BDC1 (mean) BDC2 (mean) BDC3, blabel(bar)
graph bar (mean) PI1 (mean) PI2 (mean) PI3, blabel(bar)
graph bar (mean) R1 (mean) R2 (mean) R3, blabel(bar)
graph bar (mean) FIBD1 (mean) FIBD2 (mean) FIBD3 (mean) FIBD4, blabel(bar)
graph bar (mean) BDDMQ1 (mean) BDDMQ2 (mean) BDDMQ3 (mean) BDDMQ4,
blabel(bar)
graph bar (mean) LFBD1 (mean) LFBD2 (mean) LFBD3 (mean) LFBD4 (mean) LFBD5
(mean) LFDB6, blabel(bar)
graph bar (mean) DMEF1 (mean) DMEF2 (mean) DMEF3 (mean) DMEF4, blabel(bar)
graph bar (mean) BDKE (mean) BDC (mean) PI (mean) R (mean) FIBD (mean) BDDMQ
(mean) LFDB (mean) DMEF (mean) DMEFF, blabel(bar)


//correlation

pwcorr BDKE BDC PI R FIBD LFDB BDDMQ DMEF DMEFF, star(0.05)

//to generate dummy variables
encode Industry, generate(industry_dum)
encode NumberofEmployees, generate(employees_dum)

//regression first model
reg DMEF BDKE BDC PI R FIBD LFDB i.industry_dum i.employees_dum
vif

//regression second model
reg DMEFF BDKE BDC PI R FIBD LFDB i.industry_dum i.employees_dum
vif


//regression third model

reg BDDMQ DMEF
vif

// regression fourth model
reg BDDMQ DMEFF
vif
```

C. Regression Tables

Source	SS	df	MS			
Model	9.74449273	22	.442931488	Number of obs	=	53
Residual	5.5479601	30	.184932003	F(22, 30)	=	2.40
				Prob > F	=	0.0134
Total	15.2924528	52	.294085631	R-squared	=	0.6372
				Adj R-squared	=	0.3712
				Root MSE	=	.43004

DMEF	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
BDKE	.3423316	.1249982	2.74	0.010	.0870512	.5976121
BDC	.0520566	.0928581	0.56	0.579	-.137585	.2416982
PI	-.0072968	.1201052	-0.06	0.952	-.2525844	.2379908
R	.0177027	.1379072	0.13	0.899	-.2639413	.2993467
FIBD	.3063115	.1413297	2.17	0.038	.0176777	.5949454
LFDB	-.1468201	.1596256	-0.92	0.365	-.4728192	.179179
industry_dum						
Arts, entertainme..	.3505305	.604443	0.58	0.566	-.8839068	1.584968
Construction	1.163363	.7351673	1.58	0.124	-.3380495	2.664775
Education	-.027408	.5068034	-0.05	0.957	-1.062439	1.007623
Financial and ins..	.2786597	.6012011	0.46	0.646	-.9491567	1.506476
Human health and ..	-.1041067	.5216668	-0.20	0.843	-1.169492	.961279
Information and c..	-.0326682	.5195779	-0.06	0.950	-1.093788	1.028451
Manufacturing	.4260663	.6773412	0.63	0.534	-.957249	1.809382
Other Services ac..	.16233	.4852544	0.33	0.740	-.8286917	1.153352
Professional, sci..	.001541	.5113826	0.00	0.998	-1.042841	1.045924
Public Administra..	.2812187	.5130809	0.55	0.588	-.7666322	1.32907
Transportation an..	-.885432	.721315	-1.23	0.229	-2.358554	.5876897
Wholesale and Ret..	.7207751	.7524921	0.96	0.346	-.8160187	2.257569
employees_dum						
201-1000	-.2173642	.5010797	-0.43	0.668	-1.240705	.805977
50-200	.6951817	.4727847	1.47	0.152	-.2703736	1.660737
Above 2000	.5895668	.4700418	1.25	0.219	-.3703866	1.54952
Less than 50	.5802641	.4286623	1.35	0.186	-.2951811	1.455709
_cons	1.192062	.7962462	1.50	0.145	-.4340893	2.818214

Source	SS	df	MS			
Model	13.4493532	22	.611334236	Number of obs	=	53
Residual	8.13974535	30	.271324845	F(22, 30)	=	2.25
				Prob > F	=	0.0196
Total	21.5890985	52	.415174972	R-squared	=	0.6230
				Adj R-squared	=	0.3465
				Root MSE	=	.52089

DMEFF	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
BDKE	.0501294	.1514058	0.33	0.743	-.2590825	.3593414
BDC	.050334	.1124757	0.45	0.658	-.1793721	.28004
PI	.3611968	.1454791	2.48	0.019	.0640888	.6583048
R	.1926705	.167042	1.15	0.258	-.1484747	.5338157
FIBD	-.0907394	.1711876	-0.53	0.600	-.4403512	.2588723
LFDB	-.079665	.1933488	-0.41	0.683	-.4745359	.3152059
industry_dum						
Arts, entertainme..	-.0073828	.7321399	-0.01	0.992	-1.502612	1.487846
Construction	-.3604418	.8904816	-0.40	0.689	-2.179048	1.458164
Education	-.4596376	.6138727	-0.75	0.460	-1.713333	.7940576
Financial and ins..	.255683	.7282131	0.35	0.728	-1.231527	1.742893
Human health and ..	-1.09279	.6318761	-1.73	0.094	-2.383253	.1976736
Information and c..	-1.019778	.6293459	-1.62	0.116	-2.305074	.2655173
Manufacturing	.2285712	.8204389	0.28	0.782	-1.446989	1.904131
Other Services ac..	-.3330029	.5877711	-0.57	0.575	-1.533392	.8673858
Professional, sci..	-.8081139	.6194192	-1.30	0.202	-2.073137	.4569089
Public Administra..	-.2936452	.6214763	-0.47	0.640	-1.562869	.9755787
Transportation an..	.2014538	.8737027	0.23	0.819	-1.582885	1.985793
Wholesale and Ret..	-.7047467	.9114664	-0.77	0.445	-2.566209	1.156716
employees_dum						
201-1000	-.0581295	.6069397	-0.10	0.924	-1.297666	1.181407
50-200	.3675922	.572667	0.64	0.526	-.8019499	1.537134
Above 2000	.6252147	.5693446	1.10	0.281	-.5375421	1.787972
Less than 50	.5212598	.5192232	1.00	0.323	-.5391353	1.581655
_cons	1.911816	.9644643	1.98	0.057	-.0578831	3.881515

Source	SS	df	MS			
Model	5.5225607	1	5.5225607	Number of obs	=	53
Residual	23.1095148	51	.453127741	F(1, 51)	=	12.19
				Prob > F	=	0.0010
Total	28.6320755	52	.550616836	R-squared	=	0.1929
				Adj R-squared	=	0.1771
				Root MSE	=	.67315

BDDMQ	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
DMEF	.6009408	.172136	3.49	0.001	.2553633	.9465183
_cons	1.234115	.6689821	1.84	0.071	-.1089228	2.577152

Source	SS	df	MS
Model	6.18286203	1	6.18286203
Residual	22.4492134	51	.440180656
Total	28.6320755	52	.550616836

Number of obs	=	53
F(1, 51)	=	14.05
Prob > F	=	0.0005
R-squared	=	0.2159
Adj R-squared	=	0.2006
Root MSE	=	.66346

BDDMQ	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
DMEFF	.5351525	.1427901	3.75	0.000	.2484893	.8218156
_cons	1.632064	.5190542	3.14	0.003	.5900197	2.674109

Variable	VIF	1/VIF
BDKE	2.91	0.343151
BDC	2.27	0.439939
PI	2.58	0.388066
R	3.35	0.298666
FIBD	2.89	0.345963
LFDB	4.39	0.227702
industry_dum		
2	3.80	0.263013
3	2.87	0.348748
4	6.29	0.159000
5	5.53	0.180783
6	7.83	0.127718
7	6.61	0.151278
8	2.43	0.410837
9	13.12	0.076235
10	8.59	0.116397
11	6.45	0.155133
12	2.76	0.362272
13	5.89	0.169701
employees_~m		
2	5.02	0.199167
3	12.45	0.080310
4	15.37	0.065048
5	8.66	0.115456
Mean VIF	6.00	

Variable	VIF	1/VIF
BDKE	2.91	0.343151
BDC	2.27	0.439939
PI	2.58	0.388066
R	3.35	0.298666
FIBD	2.89	0.345963
LFDB	4.39	0.227702
industry_dum		
2	3.80	0.263013
3	2.87	0.348748
4	6.29	0.159000
5	5.53	0.180783
6	7.83	0.127718
7	6.61	0.151278
8	2.43	0.410837
9	13.12	0.076235
10	8.59	0.116397
11	6.45	0.155133
12	2.76	0.362272
13	5.89	0.169701
employees_~m		
2	5.02	0.199167
3	12.45	0.080310
4	15.37	0.065048
5	8.66	0.115456
Mean VIF	6.00	

Variable	VIF	1/VIF
DMEF	1.00	1.000000
Mean VIF	1.00	

Variable	VIF	1/VIF
DMEFF	1.00	1.000000
Mean VIF	1.00	

D. Ethics Form

SAGE-HDR (v3.8 24/04/23)

Response ID	Completion date
1046015-1045997-115579018	31 Aug 2023, 18:21 (BST)

1	**Applicant Name**	Hanis Qistina Binte Mohammad Rizal
1.a	**University of Surrey email address**	hm01440@surrey.ac.uk
1.b	**Level of research**	Postgraduate Taught (Masters)
1.b.i	**Please enter your University of Surrey supervisor's name. If you have more than one supervisor, enter the details of the individual who will check this submission.**	Dr Saqib Shamim
1.b.ii	**Please enter your supervisor's University of Surrey email address. If you have more than one supervisor, enter the details of the supervisor who will check this submission.**	s.shamin@surrey.ac.uk
1.c	**School or Department**	Surrey Business School
1.d	**Faculty**	FASS - Faculty of Arts and Social Sciences

5	**Are you making an amendment to a project with a current University of Surrey favourable ethical opinion or approval in place?**	NO

6	**Does your research involve any animals, animal data or animal derived tissue, including cell lines?**	NO

8	**Does your project involve human participants (including human data and/or any human tissue*)?**	YES

		acknowledge that my SAGE answers and research project will be subject to audit and inspection by the RIGO team at a later date to check compliance.

31	**If I am conducting research as a student:**	• I confirm that I have discussed my responses to the questions on this form with my supervisor to ensure they are correct. • I confirm that if I am handling any information that can identify people, such as names, email addresses or audio/video recordings and images, I will adhere to the security requirements set out in the relevant Data Protection Policy

30	**Declarations**	• I confirm that I have read the University's Code on Good Research Practice and ethics policy and all relevant professional and regulatory guidelines applicable to my research and that I will conduct my research in accordance with these. • I confirm that I have provided accurate and complete information regarding my research project • I understand that a false declaration or providing misleading information will be considered potential research misconduct resulting in a formal investigation and subsequent disciplinary proceedings liable for reporting to external bodies • I understand that if my answers to this form have indicated that I must submit an ethics and governance application, that I will NOT commence my research until a Favourable Ethical Opinion is issued and governance checks are cleared. If I do so, this will be considered research misconduct and result in a formal investigation and subsequent disciplinary proceedings liable for reporting to external bodies. • I understand that if I have selected 'YES' on any governance risk questions and/or have selected any options on the higher, medium or lower risk criteria then I MUST submit an ethics and governance application (EGA) for review before conducting any research. If I have NOT selected any governance risks or selected any of the higher, medium or lower ethical risk criteria, I understand I can proceed with my research without review and

27	**Does your research involve any of the following individuals or higher-risk methodologies? Select all that apply or select 'not applicable' if no options apply to your research. Please note: the UEC reviewers may deem the nature of the research of certain high risk projects unsuitable to be undertaken by undergraduate students**	NOT APPLICABLE - none of the above high-risk options apply to my research.

28	**Does your research involve any of the following individuals or medium-risk methodologies? Select all that apply or select 'not applicable' if no options apply to your research.**	NOT APPLICABLE - none of the above medium-risk options apply to my research.

29	**Does your research involve any of the following individuals or lower-risk methodologies? Select all that apply or select 'not applicable' if no options apply to your research.**	NOT APPLICABLE - none of the above lower-risk options apply to my research.

24	**Does your research involve any of the above statements? If yes, your study may require external ethical review or regulatory approval**	NO

25	**Does your research involve any of the above? If yes, your study may require external ethical review or regulatory approval**	NO

26	**Does your project require ethics review from another institution? (For example: collaborative research with the NHS REC, the Ministry of Defence, the Ministry of Justice and/or other universities in the UK or abroad)**	NO

22	**Does your project process personal data2? Processing covers any activity performed with personal data, whether digitally or using other formats, and includes contacting, collecting, recording, organising, viewing, structuring, storing, adapting, transferring, altering, retrieving, consulting, marketing, using, disclosing, transmitting, communicating, disseminating, making available, aligning, analysing, combining, restricting, erasing, archiving, destroying.**	NO

23	**Are you using a platform, system or server external to the University approved platforms (Outside of Microsoft Office programs, Sharepoint, OneDrive Qualtrics, REDCap, JISC online surveys (BOS) and Gorilla)**	NO

19	**Are you undertaking security-sensitive research, as defined in the text below?**	NO

20	**Does your project require the processing of special category1 data?**	NO

21	**Have you selected YES to one or more of the above governance risk questions on this page (Q10-Q20)?**	NO

14	**Will you be working with any collaborators or third parties to deliver any aspect of the research project?**	NO

15	**Are you conducting a service evaluation or an audit? Or using data from a service evaluation or audit?**	NO

16	**Does your funder, collaborator or other stakeholder require a mandatory ethics review to take place at the University of Surrey?**	NO

17	**Does your research involve accessing students' results or performance data? For example, accessing SITS data.**	NO

18	**Will ANY research activity take place outside of the UK?**	NO

11	**Does your research involve exposure of participants to any hazardous materials e.g. chemicals, pathogens, biological agents or does it involve any activities or locations that may pose a risk of harm to the researcher or participant?**	NO

12	**Will you be importing or exporting any samples (including human, animal, plant or microbial/pathogen samples) to or from the UK?**	NO

13	**Will any participant visits be taking place in the Clinical Research Building (CRB)? (involving clinical procedures; if only visiting the CRB to collect/drop-off equipment or to meet with the research team (i.e. for informed consent/discussion) select 'NO').**	NO

9	**Will you be accessing any organisations, facilities or areas that may require prior permission? This includes organisations such as schools (Headteacher authorisation), care homes (manager permission), military facilities, closed online forums, private social media pages etc. This also includes using University mailing lists (admin permission). If you are unsure, please contact ethics@surrey.ac.uk.**	NO

10	**Does your project involve any type of human tissue research? This includes Human Tissue Authority (HTA) relevant, or non-relevant tissue (e.g. non-cellular such as plasma or serum), any genetic material, samples that have been previously collected, samples being collected directly from the donor or obtained from another researcher, organisation or commercial source.**	NO

		employees across various companies. The significance of this study lies in promoting understanding and practice of big data analytics and data-driven decision-making in organizations. The insights gained can help organizations harness data effectively, optimize decision-making processes, enhance efficiency, and embrace innovation. By addressing challenges in developing big data capabilities and promoting a data-driven culture, this research equips businesses to succeed in a dynamic digital environment.

4	**Are you planning to join on to an existing Standard Study Protocol (SSP)? SSPs are overarching pre-approved protocols that can be used by multiple researchers investigating a similar topic area using identical methodologies. Please note, SSPs are only being used by 3 schools currently and cannot be used by other schools. Using an SSP requires permission and sign-off from the SSP owner**	NO

analytics, and understanding this variability enables tailored recommendations.

3. Impact on Data-Driven Decision Making: The research assesses how the adoption and effective implementation of big data analytics influence data-driven decision-making within organizations. This examination helps organizations improve outcomes and competitiveness.

4. Performance Outcomes of Data-Driven Decisions: The study identifies specific performance outcomes resulting from data-driven decision-making. Understanding this correlation provides insights into the benefits and allows organizations to set measurable objectives.

5. Organizational Culture's Influence: The research investigates how organizational culture affects the integration and adoption of big data analytics capabilities. Identifying barriers helps design strategies for fostering a data-driven culture among employees.

6. Strategies for a Data-Driven Culture: The study explores strategies to cultivate a data-driven culture supporting evidence-based decision-making across all levels of the organization. Successful case studies and practices offer insights for organizations to develop their own approaches.

The research employs a quantitative approach, collecting data through structured questionnaires distributed to

2	**Project title**	Determinants of big data analytics capability and data driven decision making.

3	**Please enter a brief summary of your project and its methodology in 250 words. Please include information such as your research method/s, sample, where your research will be conducted and an overview of the aims and objectives of your research.**	In the modern business landscape, organizations are increasingly utilizing big data analytics to gain insights, make informed decisions, and remain competitive. This research focuses on uncovering the factors that shape the development of big data analytics capabilities within organizations and how these capabilities impact data-driven decision-making processes. The term "big data" refers to large, fast-moving, and diverse datasets that challenge traditional handling methods. The study aims to contribute to understanding how organizations can effectively use big data to drive success. The research's objectives are as follows: 1. Determinants of Big Data Analytics Capability: The study seeks to identify key factors influencing the development of big data analytics capabilities within organizations. This helps address challenges in building analytical capacity, optimizing data infrastructure, and allocating resources efficiently. 2. Variability Across Industries and Sizes: The research explores how these determinants vary across industries and organizational sizes. Different sectors and scales encounter unique challenges and resources when adopting big data