Bridging Big Data Analytics skills gap in the financial services sector

Reece Johnson

Student number: 2414115

Student email: 2414115@student.wits.ac.za

Mobile: +27 82 878 6394

Supervisor: Dr. Euphemia Godspower-Akpomiemie

A research report submitted to the Faculty of Commerce, Law and Management, University of the Witwatersrand, in partial fulfilment of the requirements for the degree of Master of Business Administration

Johannesburg, February 2024

DECLARATION

I, Reece Johnson, declare that this research report is my work except as indicated in the references and acknowledgments. It is submitted in partial fulfilment of the requirements for the degree of Master of Business Administration at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in this or any other university.

Signed by:Reece Alwyn Johnson
Signed at:2024-05-29 20:11:02 +02:00
Reason:Witnessing Reece Alwyn Johnso

29/05/24

Reece Alwyn Johnson

Acknowledgements

I express my gratitude to the Divine for providing guidance, strength, mental acuity, and expertise necessary to accomplish this research. I also want to extend my heartfelt thanks to Dr. Euphemia Godspower-Akpomiemie, my supervisor, for her consistent support, encouragement, enthusiasm, and extensive knowledge in the domain of Big Data analytics. Her guidance was instrumental throughout the research and writing phases of this dissertation.

I express my appreciation to my colleagues in the syndicate group for engaging discussions, shared efforts during sleepless nights to meet deadlines, and the enjoyable moments we experienced together over the past year.

Finally, I want to express my gratitude to my family, specifically my beloved wife Lyndal Johnson, and my children Kayley, Tyra, and Olivia. Their patience and unwavering support, both spiritually and emotionally, have been invaluable throughout this challenging yet fulfilling journey.

Abstract

As the significance of big data analytics continues to escalate within the banking industry, there is a parallel surge in demand for adept professionals in this domain. Notably, one of the foremost challenges in the era of data is the scarcity of individuals possessing the requisite skill set to transform raw data into actionable insights that generate business value. Addressing this pressing concern, this study sets out to ascertain the technical and business-oriented data analytics skills crucial in the banking sector, pinpointing the most pivotal skills required both presently and in the forthcoming years.

This qualitative research used collected data from banking institutions in South Africa and an academic institution via semi-structured interviews. A purposive sampling method was applied to select fourteen participants including executives and senior managers who have decision-making authority. A thematic data analysis was used as a lens to determine the critical skills required to effectively execute Big data and analytics in the banking industry.

The findings revealed that critical technical, business, and interpersonal skills are in short supply among graduates and professionals in the South African banking industry. The study concludes and recommends that institutions of higher learning must enhance their curricula. This includes the incorporation of simulations and experiential learning, which are deemed particularly paramount.

Table of Contents

1. Introduction

1.1 Statement of purpose

This study adopted the viewpoint that Big data analytics (BDA) is crucial for providing valuable insights beneficial for organizational decision-making and developing a competitive edge. Previous studies indicate that the financial services sector has experienced substantial digital evolution, driven by the emergence of novel technologies and the growing significance of data-based decision-making. As a result, there is an increased need for specialists with Big data analytics skills in this industry. However, the current talent pool does not seem to be adequately equipped with the necessary skills, leading to a data analytics skills gap. The implementation of BDA within the banking industry has led to a significant increase in the need for enhanced insights, prompting the acquisition of additional competencies and the integration of cutting-edge technologies to improve business performance and create a competitive advantage.

The objective of this research is to evaluate the shortage of Big data analytics abilities in the financial services industry of South Africa. This research intends to identify the capabilities necessary for the banking sector to thrive in the new digital landscape. Additionally, the aim is to obtain knowledge about Big data analytics skills and competencies necessary for the banking industry to improve its business functions in the modern economy. More so, this study aims to propose methods on how the skill gap can be mitigated.

1.2 Background to the Study

The global financial system is supported by the banking industry, as banks play a vital role in the economy by engaging in some important economic activities, such as investment and capital management, project finance, and loan distribution. In the process of engaging in such activities, banks generate millions of business records due to data-driven transactions and operations. The volume, velocity, and variety of data in banks' possession significantly become so huge that it becomes a change due to capacity shortage in Big data handling. More so, the high volume, velocity, and variety of transactional data from various bank business channels become a burden to banks, and new risk management strategies are required with "Big data" analytics providing solutions to such problems and other data management issues.

Big data utilization has been witnessed across many industries, including retail, marketing, transportation, entertainment, healthcare, financial services, and security. Furthermore, numerous business functions, including but not limited to marketing, sales, finance, human resources, and financial markets, have embraced the integration of big data analytics. Big data analysts have been deploying, analysing, and visualizing Big data sets despite the complexity, and making data-driven decisions (DDD), which include customer segmentation, customer management, product design, and redesign, product/service segmentation, and provision based on customer preferences. The banking industry can enhance profitability and gain customer insights by utilizing data-driven analytics and conducting "real-time evaluations (Shakya & Smys, 2021). Data-driven business enables banks to with internal and external partners to create value, provide quick products and services in real-time, as well safeguard consumer information (Dosso et al., 2021).

Due to the newness and importance of Big data analytics (BDA), academics and industry professionals are increasingly interested in the field, which is widely regarded as the most profound technical shift since the advent of the Internet. However, there has been a shortage of skilled qualified Big data analytics resources due to the high level of technological advancements it requires. There is an immediate need to invest in human capital to ensure continued productivity and economic advancement in a technologically propelled economy (Maruyama et al., 2015). Accenture predicted that if the data analytics skills gap is not bridged by 2028, the G20 nations may experience a collective reduction of $11.5 trillion in their total gross domestic product (GDP) growth. This equates to a loss of around 1.1 percentage points in combined GDP growth across 14 countries (Dosso et al., 2021). To fully leverage data-driven decision-making and stay competitive in the digital marketplace, it is crucial to tackle the skills gap in big data analytics within the banking industry. Banks largely depend on big data analytics to optimise their operations, customise client experiences, and manage risks such as fraud and cybersecurity threats. Nevertheless, the lack of skilled experts with knowledge in data analytics obstructs these endeavours, resulting in inefficiencies and missed chances. In addition, strict regulatory mandates require a highly skilled workforce that can ensure adherence to regulations and successfully handle large volumes of sensitive data. By addressing this disparity in expertise, financial institutions can bolster their analytical capacities, stimulate innovation, optimise operational effectiveness, and establish resilient risk mitigation strategies, ultimately resulting in enhanced financial performance and heightened consumer contentment.

1.3 Research Problem

This study explores the skills gap in Big data analytics. Innovation is the most important factor in the banking industry's sustained top-and bottom-line growth and South African banks are left behind in this quest for innovation. The South African banking sector is experiencing extraordinary innovation due to the fourth industrial revolution. As a result, there is a shift in the skills required by employers to adapt and remain competitive for business performance (Mikalef & Krogstie, 2019). However, banks are no longer viewed as the destination of choice for great software developers. Therefore they are required to organize and operate differently, preserve and cultivate talent, become agile in their development processes, and be open to collaboration with external institutions (Mikalef & Krogstie, 2019). This has a detrimental influence on banking, as the scarcity of appropriate competencies and skills will have an adverse impact on all organizational levels and performance. In the financial services industry, the demand for professionals with data analytics abilities has increased. As the sector becomes increasingly digital, there is a need for people who can analyse, interpret, and derive business value from the huge volumes of data being produced furthermore use them to drive business choices (Walker & Brown, 2019).

The research has identified technology as a key antecedent of Big data analytics (BDA). BDA preparedness is also heavily influenced by organizational characteristics, processes, and the skills and capabilities of the workforce. However, much research examines the tools, technologies, and advantages of using analytical techniques and ignores other factors. This has made it difficult for companies in developing nations to identify constraints and other factors requiring attention (Kalema & Mokgadi, 2017). The benefits of well-executed and meticulously planned BDA for businesses are well-documented. However, as with any new technology advance, various obstacles may prevent businesses from reaching Big data analytics full potential.

In line with the South African unemployment rate, the graduate labour force is growing. Higher education institutions are making an effort to modify curricula to meet industry demands and rapidly evolving systems, but there is still a disconnect between the skills produced in the classroom and those that are needed in the workforce. Earlier research undertaken across the South African banking industry revealed a lack of skills, the excellence of institutions that graduates attended, and discrepancies between employer and graduate expectations attributed to graduate unemployment in this industry (Oluwajodu et al., 2015). Understanding this new and attractive skill of Big data analytics will enable higher education institutions to coordinate and alter their curricula to meet these demands. According to Maruyama (2015), individuals employed in the domain of data analytics are typically self-taught based on the skills necessary for their jobs or obtained through on-the-job training. This indicates that there is insufficient training (in a systematic approach) in developing the appropriate skills needed in Big data analytics.

1.4 Research Objectives

Drawing from the identified problems and underlying motivation, the aims of this research can be succinctly outlined as follows:

- To investigate the skills and competencies for Big data analytics during digital transformation in the banking sector.
- To assess the skills gap in Big data analytics in the banking sector.
- To explore the roles of educational institutions and industry in closing the Big data skills gap between academic theory and industry requirements.

1.5 Significance of the Study

The findings of this research will contribute valuable insights to higher education institutions ensuring that their curriculum is current and in line with industry demands. The findings will also enable the banking sector to source and train the correct population to champion the Big data-driven decision segment of the industry. As a result, the banking industry will enhance risk management, detect fraud, tailor client experiences, optimize marketing and sales activities, and enhance operational efficiency, resulting in higher profitability and expansion.

Additionally, this study's findings can provide regulatory agencies such as SETA (Sector Education Training Authority) with valuable information about the existing skills environment and the specific abilities needed to effectively manage sophisticated data analytics. Regulators can create more specific norms and frameworks that promote the acquisition of essential skills inside the business.

Lastly, this study aims to identify knowledge gaps by precisely identifying areas where skills are weak. This will enable academic institutions and research to focus on these gaps, leading to the creation of new approaches and tools in big data analytics. Creation of novel insights and advancements in the field of data analytics specifically designed for financial services. Furthermore, the study's recommendations can have a significant impact on the creation of curricula in higher education and vocational training institutions, ensuring that they are better tailored to meet the needs of the business. Improved synchronisation between academic production and industry demands, resulting in a workforce that is better prepared for employment.

1.6 Delimitations of the Study

The scope of this study is restricted to matters specifically related to the financial services industry in South Africa. The sample comprises both institutions of higher education and the banking sector.

1.7 Operational Definitions

The below operational definitions have been adopted in the context of the skills gap of Big Data Analytics in banking:

- Data Science

According to Provost and Fawcett (2013), at its most fundamental level, data science may be understood as supporting and guiding the systemic extraction of information and knowledge from data are fundamental principles. Data science provides practitioners with a framework, structure, and principles to methodically address business problems by extracting intelligence from data for improved decision-making.

- Data Engineering

Sadiku (2018) refers to data engineering as a multidisciplinary field that involves expertise in areas such as design implementation, statistics, database management, visualisation, optimisation, and information theory.

- Analytics

Fisher (2012) refers to analytics applying expertise in data mining, statistics, machine learning and visualization to respond to complex business problems posed by leaders using business information.

- Cloud Computing

Cloud computing is the integration of several hardware computers, which share software and resources, and are enhanced by IT services. This collaborative infrastructure is intended to be installed with minimal exertion or participation from the vendor (Botta, et al., 2016).

1.8 Structure of Dissertation

The first chapter emphasizes the need for understanding and uncovering the skills and abilities required by banks to adapt to a changing environment from both academic and business perspectives, including the influence of education on the Big data analytics skills gap. The second chapter presents a comprehensive analysis of the theoretical and empirical foundations of Big data, examining relevant literature. The third chapter provides a detailed research methodology, research design, and the process adopted to achieve the objectives. In the fourth chapter, the research findings and themes are identified and presented from the research undertaken. The study culminates in the fifth chapter, where it concludes by offering recommendations and suggestions to stakeholders. Possible future research on the topic is also recommended.

2. Literature Review

2.1 Introduction

This chapter presents the impact and benefit of Big data analytics skills, specifically, reviews the current condition of the BDA skills gap in the banking sector, the aspects that contribute to the gap, and the proposed and implemented strategies to address the gap. This review provides insights and recommendations for addressing the Big data analytics skills divide in the banking industry by synthesizing the findings of previous studies (Mishi & Mushonga, 2024).

2.2 The Concept of Big Data

Big data is a vast gathering of data that cannot be handled, collected, or evaluated with conventional technologies as this data might also be found in structured, semi-structured, and unstructured formats. The structured data may take the form of text and numeric, at the same time, semi-structured data could resemble another type of data, like an email, video, images, audio, and non-structural take the form of social tweets and comments (Smaya, 2022). According to the European Commission, Big data is classified to have these four characteristics: Large volume of data, various types of data, high velocity of data, and various sources from which the data is received (Malaka and Brown, 2015).

According to Smaya (2022), Big Data in banking is based on five parameters (5 V's) which are: Volume, Variety, Velocity, Value, and Veracity, though sometimes it could be extended to 7V's to include visualization and value. Volume is the distinguishing factor between unstructured and structured data. Big data volume represents the amount of data and is measured in terabytes and petabytes. The term "variety" refers to the various kinds of data stored in relational database storage systems, such as structured, unstructured, and semi-structured data. The format of the data may take the shape of documents, social media texts, music, video, emails, graphics, graphs, and the output of all different kinds of computer-generated data from different sensors, technical devices, machine logs, and GPS signals. Velocity pertains to the rate at which data is produced and disseminated. In the current scenario the rate at which credit card transactions are processed and verified for fraudulent activities. Value denotes the significance or worth of the data that is being generated and employed. Veracity ensures that the data distributed is accurate (Quasim et al., 2019).

Big data is increasingly being acknowledged by management as the foundation of corporate operations. According to Fosso (2015), Big data's ultimate objective is to enable insights, provide business value, and produce a competitive edge.

2.3 Big data - Definitions

Numerous researchers have provided definitions for Big data, however, all the definitions lean toward the many V's that explain the characteristics of Big data. According to Malaka and Brown, (2015), Big data analytics can be specified as large volume, velocity, and variety of data assets, all of which require efficient and innovative approaches to information processing in order to improve analytical capabilities and the quality of decisions. This indicates the substantial amount of data that is received from diverse data sources including video, audio, documents, web traffic, e-mail, unconnected sources, call data records, files, and social media including relational databases.

Doko and Miskovski (2019) define BDA as the collection and analysis of massive volumes of both structured and unstructured data, with the possibility of doing so in real-time to produce value for businesses. A specific focus on the properties of 5 Vs in the context of the banking sector. Volume, Variety, Velocity, Veracity, and Value. According to Henry and Venkatraman, (2015), the business problem is to quickly evaluate these mountains of information to produce useful insights because Big data is not only unstructured but also rapidly generated and huge in volume. Big data's lack of direct control at the point of generation also raises concerns about the data's validity (quality, correctness), which occasionally casts doubt on the outcomes of data analysis. Big data also has a validity element, which refers to how valid the data is for problem analysis and whether it has an expiration date beyond which it loses some of its value.

De Mauro et al., (2016) defines Big data as the data asset whose High Volume, Velocity, and Diversity necessitate the application of specialized Technologies and Analytical Techniques for conversion into business value. De Mauro et al., (2016) further state that the Technology and Analytical elements specify the conditions necessary to make use of such information and the Value describes the conversion of data into actionable visions that can generate financial benefit for businesses and communities to make informed decisions.

The definitions provided describe Big data as an important collection of information that is characterized by three primary elements: high volume, velocity, and diversity. The term "High volume" indicates the large quantities of data generated and accumulated through various mechanisms that surpass the capacity of traditional data processing systems. The term velocity denotes the speed at which data is generated and must be analysed from various sources in real-time requiring efficient methods to analyse in a timely manner. The term diversity refers to the variety of data sources and types within the data landscape that includes numbers, tables, texts, images, and videos from many sources such as sensors, social media, and online platforms. Big data necessitates the use of specialized technologies and analytical methods due to these characteristics. Because of the size, complexity, and speed of Big data, conventional data processing tools and techniques may not be sufficient. Consequently, specialized approaches and tools, such as sophisticated data storage systems, distributed computing, machine learning, and data visualization techniques, are required to extract valuable insights and derive meaningful data from Big data sets.

Gandomi and Haider (2015) defines Big data, as a term used to define enormous quantities of rapid-moving, complex, and variable data which call for cutting-edge methods and tools in order to be captured, distributed, managed, and analysed. BDA is the frequently challenging method of analysing enormous quantities of data to find industry trends, latent patterns, and customer insights that might help organisations make improved decisions (Smaya, 2022a).

2.4 The Evolution of Big Data

The presence of Big data has been prevalent since the early 1950s, but grew gradually because of the significant cost of computers, storing of data, and network applications (Lee, 2017). The arrival of the World Wide Web (WWW) in the early 1990s led to a significant increase in the growth of data and the development of Big data analytics. Big data and data analytics have evolved through the creation of relational database management systems (RDBMS) in the 1980s which made it possible to store and retrieve vast quantities of structured data effectively. Large corporations employed these systems at first to handle their transactional data (Hugh, 2019).

Big data 1.0 (1994 to 2004) coincided with the introduction of e-commerce in 1994, and Web mining methods were devised to examine the online activities of users. Web usage mining is the application of data mining techniques to identify user access designs and patterns for webpages, while web structure mining is the process of identifying the relationship of a website or a web page linked by information. Text mining involves extracting information from unstructured text and relies on methods derived from information retrieval (IR) and natural language processing (NLP) to accomplish this task. The focus of web content mining involves the retrieval of data from web pages, grouping similar web pages through clustering, and categorizing web pages into various types such as cyberterrorism, email fraud, and filtering spam emails. (Lee, 2017).

Big data 2.0 is propelled by Web 2.0 and the social media trend, which enabled web users to interact with websites and contribute their own content. Text mining involves extracting information from unstructured text and relies on methods derived from IR and NLP to accomplish this task. Researchers analyse and interpret human behaviours on social media platforms, extracting insights and drawing conclusions based on factors such as a consumer's interests, browsing habits, social connections, sentiments, profession, and opinions (Lee, 2017). Lee (2017) further elaborates on the evolution of Big data 3.0 incorporating data from Big data 1.0 and Big data 2.0. The Internet of Things (IoT) applications that produce data in the form of images, audio, and video are the principal contributors to Big data 3.0. The Internet of Things (IoT) is a technological ecosystem where devices and sensors possess distinct identifiers and can interact and exchange data over the internet autonomously, without the need for human intervention. As the Internet of Things rapidly expands, devices that are connected and equipped with sensors will surpass social media and e-commerce websites as the primary sources of Big data.

Internet of Things (IoT)-based sensors are increasingly being used to analyse information from equipment deployed for banking operations such as ATM’s and smart devices. The development is causing the emergence of a fresh domain called streaming analytics. This field involves the ongoing extraction of data from streaming sources. In addition to overseeing current conditions, streaming analytics can forecast future occurrences. They offer immense possibilities in sectors where streaming data is created through sensor data, machine data, or human activities. For instance, healthcare providers can utilize streaming analytics to keep track of patient's physical and behavioural alterations, and to inform caregivers of urgent medical concerns.

Descriptive analytics tools including reporting, dashboards/scorecards, and alerts have received increased attention from businesses. With the advancement of the data analytics maturity model and create predictive algorithms, their focus has started to broaden. The possibility to gather, store, and analyse massive amounts of information as well as improvements in computer power, algorithms, and data analysis software are major drivers of this interest (Hugh, 2019).

Banks have been under immense pressure to stabilise and improve the level of service quality since the 2008 financial crisis. Banks are now obligated to be more transparent and seek permission to access billions of records belonging to their customers. because of new laws like Basel III and the Payments Service Directive (PSD) in digital markets (Delgosha et al., 2021). Big data offers businesses significant possibilities for starting new ventures, generating innovative goods and services, and enhancing corporate processes. Employing Big data analytics can lead to business value such as reduced costs, enhanced decision-making, and improved quality of products and services.

2.5 Factors driving Big Data Analytics

2.5.1 The Growth of Digital Volume

Recently, there has been a surge in the quantity of data due to the advancement in technology, and as a result, the number of emails, video clips, social media posts, and added text has grown to more than a few billion entries per day. The traditional storage technologies are unable to handle structured and unstructured data so more advanced technologies are best suited (Quasim et al., 2019). According to Dachyar (2019), the Internet of Things (IoT) has led to a rise in disorganized data. It is anticipated that IoT devices may increase to 43 million by the year 2023. The Big data technology and services industry is projected to expand at a CAGR of 22.6% between 2015 and 2020, reaching $58.9 billion by 2020 (Dachyar et al., 2019). According to Kubina et al. (2015), the volume of data is increasing rapidly, as evidenced by Google which receives 2 million search queries every minute, and Facebook users publishing around 700 thousand quantities of content simultaneously. IBM estimates that we generate 2.5 quintillion (2,5 x 1018) bytes of data every day, meaning ninety percent of the existing data was produced in the previous two years (Kubina et al., 2015).

2.5.2 Availability of powerful data analytics tools

In the digital age, data and analytics are crucial for companies to not only survive but also to flourish, so it is essential to drive them into the foundation of their business (Dachyar et al., 2019). Data variety includes various types and formats of data that are produced from a variety of sources, including text, photos, audio, and video as well as organized and unstructured data in real-time (Smaya, 2022). This data needs to be organised, arranged, and segmented for analytical queries by a data professional for decision-making purposes (Smaya, 2022a). Standard tools and competencies have been effective in handling organized data and are inadequate when processing unstructured data. Provost and Fawcett (2013) argue that despite having access to substantial amounts of structured, semi-structured, and unstructured data through the monetary transactions they oversee, the banking sector typically fails to utilize this extensive collection of big data for purposes such as targeted marketing and predictive analytics. Consequently, it is necessary to utilize and implement more advanced analytical techniques to process data, enabling us to obtain better insights into our business, customers, and suppliers (Kubina et al., 2015). Data science and machine learning engineers are in short supply, and businesses are increasingly searching for Big data analytics capabilities that can bridge the skills gap (Provost & Fawcett, 2013). Although the Big data revolution is widely acknowledged as a potential means of obtaining a competitive advantage, organizations lack the capacity to harness its power effectively.

2.5.3 The Rise of Cloud Computing

The term "Big data" was introduced as a solution to manage and process the rapidly expanding amount of data worldwide. It is a technology capable of storing and managing vast and varied volumes of data, providing companies with a comprehensive understanding of their customers. Cloud computing offers a reliable, fault-tolerant, accessible, and adaptable platform for hosting distributed management systems that handle massive volumes of data. (Neves et al., 2016). Cloud computing provides a flexible infrastructure that enables on-demand delivery of computing resources. According to Jagani (2021), Big data applied with Cloud computing provides improved performance, increased input data, enhanced insights, and more dependable and secure platforms at comparatively cheaper costs for customers. Jagani further states that Big data provides end users with visualization and analytics, making it more manageable, monitorable, and reportable. Cloud computing can be difficult to comprehend and manage because it involves a vast array of technologies and services. The technology of cloud computing is continuously evolving, and new features and services are introduced frequently. Keeping up with these changes can be difficult, and organizations need experienced professionals who are familiar with the most recent cloud technologies and can assist them in navigating the complexities of Cloud computing.

2.5.4 The need for faster Decision-making

Big data offers many advantages by enabling higher transparency of information, broader, enhanced, and more precise understanding, improving decision-making, and allowing organizations to develop more intricate and tailored products and services (Kubina et al., 2015b). Data sourced from various sources are meaningless if there is no way to identify structure or solutions to specific problems within them, and this cannot be done by people due to the volume, variety, and velocity of data (Huttunen et al., 2019). Over time, managers have been accountable for making decisions based on their proficiency, expertise, previous experiences, personal predispositions, natural intelligence, and gut feelings. There has also been a significant element of luck involved due to the high level of uncertainty surrounding external factors. Companies hire managers with the best track records and expect them to leverage their previous successful decisions to make sound future decisions. Managers have the benefit of utilizing extensive datasets and data-driven tools that can aid them in making superior decisions based on more impartial and factual information (which is represented by real-time Big data streams) and gaining comprehensive knowledge about them through the application of Big data (Henry & Venkatraman, 2015).

2.5.5 The Desire for Competitive Advantage

A company has a competitive edge when it outperforms its rivals (Barham, 2017b). Companies can gain a competitive advantage by leveraging Big data, which presents greater opportunities for growth than conventional technologies. and assist in gaining a leading position in the market. Big data has enabled companies to put up experiments by generating and analysing enormous quantities of data from process change experimentations to identify possible improvements in performance (Kubina et al., 2015). According to Kubina (2015), insurance companies use Big data analysis to assist with identifying hidden correlations and risks by using fraud analysis engines.

2.5.6 Customer Insights

According to Poornima and Pushpalatha (2016), different types of unstructured data make up nearly 67% of the Big data generated by business applications. As a result, those companies are passing up the chance to fully utilize the value of knowledge discovery from those data, which could be helpful to increase their profit and production. As a result, the increased complexity of business processes and innovation requires the introduction of new digital technologies to simplify things by automating these processes hence analytics assists in forecasting based on historical events which will enable us to determine the actions to be taken.

2.5.7 Redesigning of products and services

By leveraging Big data from multiple sources, businesses have the ability to send personalized product/service recommendations, coupons, and various promotional offers. Prominent retailers such as Macy's and Target utilize Big data to evaluate customer preferences and sentiments, aiming to enhance the overall shopping experience. Similarly, innovative financial technology firms have started leveraging social media data to assess the credit risk and financing needs of prospective customers, providing them with novel financial solutions. Big data analysis is being used by banks to raise sales, improve client retention, and provide better customer service (Lee, 2017).

2.6 Challenges with Implementing Big Data Analytics

The financial services sector is a data-intensive industry that handles vast amounts of confidential information. The financial services industry relies heavily on extensive data analytics to assist both consumers and financial service providers. According to He et al. (2022), Big data methodologies encounter issues when handling large datasets with high dimensionality and generating meaningful findings without employing an adequate variable selection process for dimension reduction. With the increasing use of advanced analytic approaches, the outcomes they produce can pose challenges for individuals lacking professional expertise in comprehending them. According to Hasan et al. (2023), commercial banks do not have a well-developed data analysis team that possesses advanced skills, expertise, and experience. As a result, there are obstacles such as limited knowledge and confidence in handling large data, as well as an inadequate ability to identify undiscovered problems.

2.7 Theory underpinning Big data

The proposed theoretical framework of this research seeks to explore how Big data analytics capabilities influence business performance in the financial services sector. It employs the resource-based theory and emphasizes Big data analytics to be a crucial component to enhance business performance.

2.7.1 Resource-based View Theory

The resource-based view theory of the firm posits how a firm's acquisition of essential resources empowers it to establish a competitive edge and enhance performance as a result of the resources under its management (Chahal et al., 2020). Big data analytics capabilities (BDAC) have been interpreted as a company's capability to convert large amounts of data into valuable information and insight through the use of both tangible and intangible assets (Aziz et al., 2023). The theory also emphasizes imitability to be a source of exceptional performance and competitive advantage (Akter et al., 2016). Barham (2017b) proposed, as part of a firm's resource-based perspective, that it should seek out resources that can cement a first-mover advantage. Barham further states that to obtain a competitive advantage that is long-lasting and sustainable, valuable rare resources should not be imitable. These are resources that have accumulated some expertise in a specific business area that allows lower costs and higher quality that their competitors do not yet possess. Aziz et al (2023) argues that organisations must continuously evolve and reconfigure their resources and capabilities to obtain a competitive advantage.
In addition, businesses must respond swiftly and efficiently to both internal and external changes, which requires them to recognise both opportunities and obstacles for market continuity and expansion.

The fundamental elements of the resource-based view theory are resources and capabilities, which have been extensively studied in previous empirical research. Resources are unable to create any business value on their own, but strategic action is necessary to utilize them effectively. According to Mikalef et al. (2018), resources are classified into three categories: tangible resources such as financial and physical assets, human skills which include the abilities and knowledge of employees, and intangible resources such as organizational culture and learning. Mikalef et al. (2018) argue that the performance of the company is contingent on the characteristics and the quality of the competencies and resources it possesses. One of the characteristics of resources is that they are incapable of producing any kind of value for the company on their own, instead demand some kind of action to be utilized strategically. This is suggested by the classification of resources as nouns, as they can remain inactive until they are required, much like an unused plant or knowledge, and they can be named irrespective of their use. Therefore, a resource is not something that a company can do; rather, it is something that the company has access to. Organizations require specialized skills and competencies to utilize the resources, which include analysis of immense amounts of data and identifying difficult-to-identify patterns and trends. This can aid organizations in making more informed decisions regarding risk management, customer insights, and marketing strategies.

2.7.2 Tangible Resources

Tangible resources are resources that possess physical attributes, including equipment, and physical technologies (Michalisin et al., 1970). These resources are readily accessible in the open market are generally not limited in supply and are required to create capabilities. Barney (1991) states that organizations of similar size have relatively easy access to tangible resources, implying that tangible resources alone are unlikely to create a competitive advantage. (Mikalef et al., 2018) states that data is the core resource in developing a Big data analytics capability. The author further states Big data is described by its volume, variety, and velocity, but analysts are primarily concerned with the data's quality. Critical elements such as completeness, accuracy, format, timeliness, dependability, and perceived value define data quality. In a heavily data-driven economy, data resources with these characteristics are required for organisations to gain a competitive edge and improve performance. Companies that purchase data to supplement their analytics and obtain additional insight into their customers and processes recognize the significance of data availability and integration from multiple sources.

2.7.3 Intangible Resources

According to Michalisin et al. (1970), some resources do not qualify as sources of competitive advantage and business performance. Strategic assets refer to resources that serve as enduring sources of competitive advantage and enable organizations to achieve superior profits. Michalisin et al. (1970) further asserts that employees know how to be an important resource to an organisation’s competitive advantage and business accomplishment. Know-how refers to the collective knowledge, understanding, expertise, skills, and decision-making abilities of both the managers and non-managers within a firm.

2.7.4 Human Resources

A company's human resources include all of its employees' talents, skills, business savvy, problem-solving abilities, leadership qualities, and interpersonal relationships (Gupta & George, 2016). The author further asserts that previous research on Information Technology capability has indicated that technical and managerial skills are key dimensions of human resources that are crucial in the environment of Information Technology. Akhtar et al. (2019) Previous research on Big data analytics has underscored the significance of possessing particular skills and capabilities related to Big data to achieve organizational success. According to Gupta and George (2016), Big data technical skills encompass computing, mathematics, statistics, machine learning, and expertise in the business domain. These skills enable them to transform conventional business practices into data-driven insights, leading to actionable decisions that improve overall business performance. Additionally, decision-makers inside businesses must have the capability of successfully interpreting the information, ensuring the identification of both opportunities and risks. According to Provost and Fawcett (2013), the technical capabilities noted above are data science skills. A data scientist's major function is to advise and aid in the process of deriving actionable insights and information from data to assist data-driven decision-making (Provost & Fawcett, 2013). Furthermore, this position calls for the use of intuition, knowledge, creativity, and practical thinking, as well as proficiency in analytics and statistics, good data visualization, and the expression of business difficulties from a data-centric perspective (Provost & Fawcett, 2013).

It has been hypothesized that the scarcity of data science skills is a consequence of academic institutions being unable to promptly establish relevant data science programs to meet the increasing demand from industries (Provost & Fawcett, 2013). The shortage of Big data technical skills remains prevalent, as only a limited number of educational institutions provide relevant courses, despite an escalating demand for such skills (Gupta & George, 2016).

2.8 Components of Resource-based View Theory

Barney (1991) argues that the resource-based view theory advocates for recognizing knowledge as a valuable competitive asset. Resource-based view theory views organisations as aggregations of resources that can be strategically utilised to create a competitive edge (Gupta & George, 2016). Kozlenkova et al. (2013) asserted that the Resource-Based Theory is a crucial paradigm that is widely employed to elucidate and forecast an organization's competitive advantage and performance results. The resource-based view theory is largely acknowledged as the most resilient and influential framework for explaining and predicting organisational relationships, competitive advantage, and profitability (Gupta & George, 2016). Erevelles et al. (2015) argued that resource-based view theory is an appropriate framework for comprehending the advantages of big data and how businesses can more effectively exploit it to improve business performance. These traits are important in the research environment because understanding the current and necessary big data resources of organizations and finding ways to effectively use and benefit from them is crucial for gaining a competitive edge.

2.8.1 Technology

Technology is considered essential for the development of capabilities, even though it is not considered to be the only source of business performance for organizations (Barney, 1991). Resource-based view theory has been utilized extensively in the area of Information Technology (IT) to characterize the abilities and skills associated with IT resources. According to existing IT research, the ability to utilize IT resources effectively alongside other organizational resources is a crucial factor in obtaining a competitive advantage (Smith et al., 1996). Past research has utilized the idea of IT resources and Big data analytics capabilities to demonstrate how they can directly or indirectly impact an organization's performance (Raguseo, 2018). The fundamental assumption of these studies is that a company must have invested in all necessary resources to develop a robust Big data capability. As businesses discover new uses for the enormous amounts of data they possess, Big data technologies, practices, and approaches are continuously evolving (Raguseo, 2018). Various Big data architectures and techniques, involving the acquisition, handling, management, and analysis of enormous volumes of data across the enterprise, are continuously evolving requiring organisations to invest in upskilling and reskilling employees in these technologies.

2.8.2 Data

The correlation between data and Big data analytics capabilities and business performance is the capacity to obtain meaningful understandings from analysing data and use them to make growth-driving decisions. Responding to the increase in volume, variety, and velocity of data, organizations need sophisticated analytical capabilities to transform raw data into actionable insights. The data can offer valuable perspectives on emerging risks and assist organizations in making better-informed decisions. Big data analytics, organizations can recognise relationships, trends, and anomalies that are not readily apparent using conventional data analysis techniques. By utilizing these insights, businesses can optimize operations, develop new products and services, improve customer experiences, and obtain a competitive advantage. Therefore, the ability to utilize data effectively through Big data analytics capabilities is becoming increasingly crucial for business performance.

2.8.3 Technical Skills

Technical Big data skills pertain to the expertise needed to utilize emerging technologies for extracting valuable insights from large and complex datasets (Gupta & George, 2016). According to Ferreira et al., (2023), working with and analysing massive amounts of data, it is essential to possess the requisite technical skills. To analyse complex data, organizations require individuals with strong technical skills, such as data programming, data mining, and data visualization. According to Gupta and George (2016), Machine learning expertise, data mining, data cleansing, analysis of statistics, and experience with programming concepts like MapReduce are all examples of the kinds of expertise that may be useful. These competencies are essential because they enable professionals to access, manipulate, and analyse data in order to derive valuable insights that can inform business decisions (Provost & Fawcett, 2013). Without these technical skills, it may be challenging for financial institutions to extract insights from their data, which could result in missed opportunities and diminished competitiveness. Individuals and organizations that wish to remain competitive and make data-driven decisions must therefore possess the necessary technical skills for Big data.

Studies conducted outline that Big data analytics technical abilities contribute to business performance by enabling organizations to analyse and interpret large volumes of data effectively (Tokarcikova & Malichova, 2020). Professionals with proficiency in data programming, data mining, and data visualization can utilize Big data analytics tools and techniques to identify patterns, trends, and insights that can inform business decisions (Talwar et al., 2021). By leveraging these insights to optimize operations, improve products and services, and enhance customer experiences, businesses can obtain a competitive advantage and increase their overall business performance (Talwar et al., 2021). In addition, the ability to effectively manage and analyse data can result in cost reductions, an increase in revenue, and enhanced risk management, all of which contribute to enhanced business performance (Drydakis, 2022).

2.8.4 Managerial Skills

According to Ferreira et al., (2023), by offering a good management framework for making wise business decisions, Big data management capabilities play a significant role in Big data analytics capabilities. It consists of four main aspects that are seen to be essential for the implementation of Big data management capabilities: Big data analytics planning, investment, coordination, and control. The first phase in Big data management capabilities is Big Data analytics planning, where possible business prospects are evaluated, and Big data-based concepts are created to increase the performance of the company. According to Gupta and George (2016), leadership abilities are firm-specific and refined periodically by employees within the organization. The author argues that these skills are improved by solid departmental relationships. These skills are heterogeneous across organizations and are utilized by managers in their daily work and decision-making. Torres et al. (2018) Big data managers need to be data savvy in order to analyse the data collected by the technical teams. Big data managers must possess the crucial capacity to comprehend and anticipate the requirements of various business units, customers, and partners. This enables them to effectively apply the insights derived by their technical teams.

2.8.5 Organizational Learning

According to Widyaningrum (2016), a learning organization can be defined as one in which employees continuously improve their capacity to produce desired results, foster original and diverse thought patterns, unleash group ambition, and continuously hone their skills in collaborative learning. According to Gupta and George (2016), one limitation of the resource-based view theory is its failure to explain why certain firms outperform their competitors, particularly in dynamic and rapidly evolving market environments.

2.9 Chapter Summary

This study proposed an understanding of the components of Big data Analytics Capabilities (BDAC), namely managerial skills, data, technology, technical skills, human resources, and organisational learning, highlighting their interdependencies for achieving superior productivity and efficiency, leading to improved performance and long-term competitive advantage. Despite the significant investment required for Big data infrastructure and a shortage of qualified personnel, the literature identifies several significant obstacles (Torres et al., 2018). These obstacles originate from a lack of comprehension regarding the advancement of Big data analytics capabilities, an inadequate understanding of the distinct resource requirements, organizations that are unprepared for the execution of Big data, and the difficulty of adopting a data-centric culture.

The challenges in building Big data analytics capabilities arise from a lack of understanding of how to do so, a lack of understanding of the specific resource requirements, organisations not being prepared for Big data, and the inability to adopt a data-oriented culture. These challenges have been identified in the literature due to the substantial investments required for Big data technological infrastructure and the scarcity of skilled human resources (Erevelles et al., 2015; Gupta & George, 2016; Kiron et al., 2011; Lavelle et al., 2011; Torres et al., 2018).

3. Research Methodology and Design

3.1 Introduction

This chapter provides a detailed methodology adopted in this study. It outlines the study's targeted population, and sample size chosen to effectively investigate the research objectives. The methodology also encompasses a thorough discussion of the sampling method, data collection procedures, and data analysis procedures. The chapter concludes by elucidating the procedures taken to ensure the reliability and validity of the data and addressing the study's limitations.

3.2 Research Approach and Design

The design of qualitative research permits researchers to examine in-depth the experiences, perspectives, and viewpoints of individuals through techniques such as interviews, focus group discussions, and observations (Sofaer, 2002). Interviews provide the researcher with the ability to pose complex questions, deconstruct them, and include follow-up inquiries in order to obtain the necessary information.

This study adopted a qualitative approach as it predominantly suites the interpretivism domain. The philosophy is employed when the researcher requires an understanding of an individual's experiences, therefore, this method is substantially subjective. Saunders et al. (2023) asserts that interpretivism is key for understanding differences between humans as social factors. This is based on the nature of the questions which the study seeks to answer. According to Eriksson and Kovalainen (2015), qualitative business research is a method that seeks to achieve business objectives by employing methods that enable detailed interpretations of phenomena without relying on quantitative measurements. Emphasis was placed on unearthing underlying meanings and generating novel insights. The researcher was able to learn more about a little-explored subject area by using a qualitative, exploratory research method (Saunders et al., 2023). The inquiries of the research are based on personal perspectives and are influenced by the social environment, rather than concentrating on factors that can be easily measured or quantified. Qualitative research involves exploring and analysing non-numeric data, emphasizing the generation and examination of information that cannot be easily quantified (Bryman & Bell, 2011). The scope of this research used once-off interviews conducted over a short period of time to achieve the research objectives. Saunders et al. (2023) proposes that during the conduct of an exploratory study, interviews incorporating open-ended questions are more suited to be employed. The objectives and questions of the research necessitate the conduct of semi-structured interviews. Further investigation is necessary to explore and understand the under-researched skillset that the banking sector aims to utilize for improved decision-making and enhanced business performance. This approach enabled the researcher to obtain insights into the skills and competencies required in the banking industry to meet its shifting priorities. Various perspectives, including those of employers (executives) and higher education institutions, were examined to explore the skills that are currently in demand within the banking industry.

3.3 Population and Sample

3.3.1 Population

The study population is defined as the total number of members of a group, which can include not only individuals but also organizations and locations (Saunders et al., 2023). The aim of this research paper was to determine the skills required by the banking industry from the perspective of employers. Given that the overall aim of this study was to identify the skills that are required by banks from the employer's perspective, the study population comprised executives, senior managers, and other data specialists within banks in the South African banking sector and a stakeholder from an institution of higher learning. The population of the research comprised traditional major banks and the South African Reserve Bank.

In this study, the unit of analysis was the perceptions of individual executives, senior managers, and specialists employed in the banking sector in Gauteng. The selected sample provided the researcher with a comprehensive understanding and valuable insights into the skills and competencies necessary for the banking industry to address its evolving needs, in comparison to the skills imparted by institutions of higher learning in South Africa.

3.3.2 Sampling Method and Size

This research study employed non-probability purposeful sampling. Purposive sampling is a non-probability sampling technique that allows the researcher to select the cases that will allow him or her to respond to the research questions in the most effective manner (Saunders et al., 2023). According to Marshall (1996), purposive sampling is the prevailing sampling technique. The purposive sampling strategy was chosen due to the exploratory character of the research, the use of semi-structured interviews, and the expectation of small sample sizes.

The researcher's targeted sample size for the study was 20 respondents. A total of 14 out of 20 interviews were conducted with an estimated 13 senior managers, specialists, and heads of departments to achieve the first two research objectives. The third research objective is achieved by interviewing an executive from one institution of higher learning. South African banks were chosen because they play a critical role in the economic functioning and financial stability of a country. Additionally, these banks are currently experiencing extraordinary transformations as a result of the technological advancements taking place.

The interviewing procedure and data processing were conducted simultaneously until theoretical saturation was achieved. In this research, the point at which no new themes could be identified was reached during the tenth interview, however, the researcher continued with the interviews of the remaining four participants to ensure no new themes were identified. Dworkin (2012) posits that saturation occurs when collecting additional data does not yield new theoretical insights or uncover new aspects related to the core theoretical categories.

3.4 The Research Instrument

A semi-structured interview is a way of collecting data that involves asking questions based on a preset theme framework (Saunders et al., 2023). Individual, detailed, semi- structured interviews were conducted by the researcher using the interview guide outlined in **Appendix C**. The interview guide followed a semi-structured format which consisted of open-ended questions related to each of the primary research objectives. This provided the researcher with a deeper understanding of participants' perspectives, beliefs, behaviours, and experiences through the use of open-ended questions, which were designed to elicit rich and detailed information (Hyman, 2016). These questions did not restrict or limit the responses of the participants, allowing them to provide nuanced and detailed accounts that may disclose unexpected insights or uncover new perspectives.

3.5 Procedure for Data Collection

This study utilized a mono-method approach, employing in-depth, face-to-face, semi-structured interviews. In-depth semi-structured interviews are well-suited for investigating a specific idea or topic from various perspectives. They are particularly valuable when researchers aim to gather detailed information about a particular issue. These interviews enabled flexibility in question sequencing and allowed for spontaneous inquiries, thereby facilitating a thorough exploration of the chosen topic (Saunders et al., 2023). Semi-structured interviews enabled the elicitation of detailed responses and the further investigation and examination of the topics of interest. The interviews utilized open-ended, non-leading questions from **Appendix C**.

The study centered on the South African Banking industry, which predominantly has its headquarters in Johannesburg. Participants were initially identified and contacted through LinkedIn. Once the contract was established, the interviews were confirmed through electronic mail. The interviewees were invited to participate in the study utilising the invitation included in **Appendix A.** The interviews were voice recorded with the permission of the participants, as evidenced by a signed consent form. An example of the consent form is in **Appendix B.** The interviews were later transcribed for analysis.

3.6 Data Analysis and Interpretation

The qualitative data obtained from the interviews was captured through voice recordings and subsequently transcribed to uphold the level of confidentiality. Thematic analysis, a method of analysing data by identifying and organizing themes, was employed in this study. The researcher conducted manual coding by utilising categories and themes that arose from the data. The researcher employed coding units to establish connections between the data and the subjects, themes, and concepts, facilitating the sorting, organising, and categorising of the data. The analysis facilitated the researcher in discerning patterns and correlations in the assessment of the capabilities necessary for the banking industry to effectively utilise Big data and analytics. This approach allowed the researcher to gain a comprehensive perspective and comprehension of meanings, relationships, and themes derived from the transcripts. During the qualitative data analysis process, to the initial data was revisited as part of the iterative process. This ensured consistency and alignment between the analysis and the recorded data.

3.7 Limitations and Challenges of the Study

Qualitative research is inherently subjective and can be influenced by various biases, such as interviewer bias and response bias (Saunders et al., 2023). These biases are anticipated and acknowledged within the study. Biases are introduced because the researcher is not impartial during the data collection procedure. Biases arise due to the lack of impartiality on the part of the researcher during the process of collecting data. Due to the qualitative nature of the research, establishing the credibility and reliability of the data will require special care (Shenton, 2004).

Due to the utilization of semi-structured interviews for data collection, inherent biases such as interviewer and interviewee bias are introduced. The lack of complete independence of the researcher from the data collection process further contributes to the potential biases in the findings. While the study primarily involved participants from prominent banks, it is important to note that the conclusions are not extrapolated to the entire South African banking industry. Furthermore, the smaller sample size may introduce biases and limit the generalizability of the data analysis results. Additionally, the recruitment of participants exclusively from banks in Johannesburg introduces a geographic bias.

3.8 Quality Assurance

3.8.1 Reliability and Validity

In a qualitative investigation, validity refers to the degree to which the research findings accurately represent the studied context, while reliability indicates that if the study were to be repeated, the results would remain largely consistent with the original research outcomes (Richards & Morse, 2013). Within the domain of qualitative research, the assessment of reliability and validity encompasses various aspects, including interviewer bias, interpretation bias, and responder bias. Validity in the data was upheld by conducting interviews in a language comprehensible to the study participants. Standardization of data collection techniques and protocols was used to guarantee the reliability and validity of the measurement instrument. The interviews were conducted in a standardised manner, following the interview guide provided in **Appendix C**, with the goal of improving both validity and reliability. In order to mitigate researcher bias and enhance precision, interviews were recorded using a recording device, and the researcher simultaneously made notes throughout the interview session. The researcher was aware of important factors such as interviewer and interviewee prejudice, and the need to have the necessary interviewing skills. Therefore, the researcher made a deliberate effort to uphold impartiality to the greatest degree feasible.

3.9 Ethical Considerations

Ethical considerations play a crucial role when conducting research interviews. The safeguarding of human participants through the application of appropriate ethical principles is crucial in every research study (Arifin, 2018). The participants were initially provided with an information sheet including the objective of the research study. They were assured of confidentiality and anonymity and were informed that their partaking was entirely voluntary, with the freedom to retract without facing any negative penalties. After fully understanding the provided information, the participants were requested to sign the consent form (refer to **Appendix B**: Consent Form).

3.9.1 Voluntary Participation

Participation in the research study was optional, and no sort of remuneration or compensation was given to participants. Participants were granted the liberty to discontinue their involvement in the study at any point without encountering any personal or occupational repercussions. The researcher made it clear that participants were not forced or intimidated, and their cooperation in the research project was completely free and unbiased.

3.9.2 Confidentiality and Anonymity

The information obtained during the interview sessions was handled with the highest level of secrecy, creating a safe setting for both the researcher and the participants to avoid any possible harm. Data presentation strictly adhered to the principle of maintaining anonymity, ensuring that no personal names were revealed at any stage of the research. To ensure the confidentiality of the banks and participants, no names are revealed, and all individuals involved are named as participants.

3.9.3 Storage of Data

The entirety of the data collected throughout the research project has been securely archived to ensure the protection of the anonymity and confidentiality of all participants. Electronic information is stored on a device that requires a password for access, while paper-based data is kept safely in a locked filing cabinet at the researcher's residence. The collected data in this study will be retained for a period of six months.

3.10 Chapter Summary

This chapter furnishes a comprehensive overview of the research methodology employed in this study. The elucidation encompasses a thorough discussion of the research design, the data collection, and the method of analysis. The explication of the research population and sample is both detailed and justified. Furthermore, the procedures implemented to ensure the validity and reliability of the research findings are meticulously outlined. Additionally, this section meticulously addresses ethical considerations and processes associated with the study.

4. Presentation and Interpretation of Research Results

4.1 Introduction

This chapter presents the results of the analysis of the data collected through in-depth interviews conducted with professionals working in the banking sector and one educational institution in South Africa. An analysis of the qualitative interview data has provided valuable insights into comprehending the specific talents required by the banking business in South Africa. More precisely, the emphasis is on how well the industry understands the use of big data analytics, the effectiveness of the methods used to handle massive amounts of data, and the necessary resources and capabilities required.

This chapter includes an introduction to the study participants and presents the outcomes associated with each of the three key research objectives introduced in Chapter 1.

4.2 Demographics of Sample

Twenty individuals were chosen for interviews using purposeful non-probability sampling. The eligibility criteria required individuals to possess a minimum of five years of experience in the banking sector, namely in professional positions such as Information Technology executive head of a function (Executives), senior manager, and specialist. All individuals who were approached satisfied the specified criteria for inclusion.

4.2.1 Description of the Sample

A total of 14 participants were interviewed to gather the data required to understand the skills in demand by the banking sector and the readiness of graduates for the workplace in Big data and analytics. Table 1 represents the participant identifiers, role, and years of experience.

Table 1: Description of Study Participants

Identifier	**Role in Bank**	**Banking Experience**	**Duration**
Participant 1	Chief Information Officer	18 years	19:09 min

Participant 2	Head: Data Management Services	11 years	50:20 min
Participant 3	Senior Manager: Data, Analytics & AI	23 years	54:30 min
Participant 4	Head: Enterprise Information Management	31 years	52:28 min
Participant 5	Head: Personalisation & Profit Science	20 years	49:00 min
Participant 6	Executive: Personalisation & Profit Science	15 years	55:46 min
Participant 7	Manager: Enterprise Information Management	5 years	42:48 min
Participant 8	Senior Manager: Data Analytics	9 years	37:47 min
Participant 9	Head: Advanced Data Analytics	16 years	41:50 min
Participant 10	Senior Manager: Data Analytics	5 years	29:38 min
Participant 11	Chief Data Officer	5 years	59:21 min
Participant 12	Executive Head: Data Services (South Africa)	15 years	29:27 min
Participant 13	Senior Business Intelligence Manager	28 years	40:14 min

Participant 14	Academic Director	3 years	60:03 min

The participants in the study were individuals from diverse organisational, corporate, and functional levels who are responsible for developing and implementing Big data and analytics initiatives in the banking sector in South Africa. A participant in the study, with extensive experience in the banking industry, currently has an executive position at an educational institution. The bulk of participants, specifically 71%, were male, while the remaining 29% were female. Out of the total participants, twelve individuals (86%) were directly engaged in data and analytics, whereas two people (14%) possessed prior experience in the field of data and analytics.

The mix of the professional roles of the participants were as follows:

- 57% executives
- 29% senior managers
- 14% middle managers

The mean number of years of participant experience in the banking business was 15 years, while the mean participant experience in academics was 3 years. The participant with the highest number of years of experience in the banking industry had 31 years of experience and the participant with the least number of years had 5 years of experience. The participant from the educational institution had 3 years of experience in academia.

The mix of banking industry experience years and academia experience years are as follows:

Banking industry experience:

- More than 5 years – 31%
- More than 10 years – 7%
- More than 15 years – 31%
- More than 20 years – 31%

Academia industry experience

- More than 5 years – 0%

4.3 Presentation of Results

This section provides the analysis of the responses from the participants relating to the research objectives. These objectives encompass identifying the skills demanded by the banking industry. Secondly, to identify the abilities that are currently lacking in the banking industry. Thirdly, to examine the skills that can be obtained through educational institutions and banks. The themes relating to the objectives are also presented and interpreted.

4.3.1 Skills and competencies for Big data analytics

The first objective of this research is to identify the skills demanded by the banking industry. This section presents and interprets the themes that emerged relating to this objective. The themes relating to this objective are tables in Table 2.

Table 2: Themes on skills and competencies for Big Data analytics

	Themes	**Sub Themes**	
		Objective 1	**Objective 2**
1	Academic Qualification	For a career in big data analytics, several university degrees in Engineering, Economics, Statistics, Mathematics, Computer Science, Information Systems, Data Science, and Actuarial Science stand out in Big data analytics.	There was no indication of insufficiency in the academic qualifications of the graduates.
2	In-Demand Technical Skills	Skills such as Programming Languages, Statistical Analysis, Cloud Computing	The technical competencies identified as deficient within the banking sector encompass programming languages,

		Data Analytics, Data Quality Data Modelling, Data Science, and Data Engineering were in demand by the banking industry.	Cloud Computing, Data Science, and Data Engineering.
3	Business Skills	The requisites to augment technical competencies encompass proficiencies in business acumen, strategic acuity, ethical considerations, and domain-specific expertise and knowledge.	The pivotal business proficiencies identified as deficient encompass business acumen and domain expertise.
4	Interpersonal Skills	The identified soft skills are communication, problem-solving, presentations, storytelling, negotiation, continuous learning, critical thinking, teamwork, and customer-centricity.	The essential soft skills identified as deficient include communication, problem-solving, storytelling, and a commitment to continuous learning.

4.3.1.1 Theme 1: Academic Qualification

The analysis of the participants' responses revealed that all the participants expressed the significance of tertiary qualifications in engineering, mathematics, economics, statistics, computer science, information systems, and actuarial science in providing the theoretical foundation for Big data and its analytics. The participants implied that specific tertiary qualifications such as in statistics, economics, and actuarial science provide graduates with enhanced skills for data science. They also discussed the importance of mathematics and engineering tertiary qualifications for data engineering roles which is key for building data pipelines. Academic qualifications in computer science, information systems and engineering were deemed to provide graduates with some skills for data engineering roles.

Participant 7 asserted that whilst tertiary qualifications are the foundation of Big data and analytics there is still a need to develop these graduates for the data scientist and data engineering roles. *"Graduates have a degree which gives them the foundation but there is still a steep learning curve and upskilling for data scientist and data engineering."*

Participant 12 asserted that an actuarial science qualification better equips graduates for data science roles. "Students are more curios, resilient and I find that they have domain some experience. And the same with chemical and mechanical engineering. Informatics resources don't really know how the data pipelines fits together."

Two of the participants perceived that institutions of higher learning are trying to introduce new curricula. They mentioned that they were aware of a tertiary qualification in Data Science, however, this was only offered at master's level.

Participant 4 perceived the academic curriculum to have improved over the last five years; however, it was asserted that more still to be done from institutions of higher learning due to the high demand for data engineering and machine learning skillsets. *"In the varsities now, they are starting to train the students more on data management, how to use data, what a data strategy is… I'm starting to see that come through more in the interviews now as compared to 5 years ago. So, varsities have changed their curriculum somewhat."*

Two participants mentioned that some of the graduates had obtained relevant qualifications from institutions of higher learning in other countries such as the United Kingdom and the United States of America. This was highlighted as a concern due to affordability for local students.

4.3.1.2 Theme 2: In-Demand Technical Skills

In all of the cases, the participants stated that technical skills were mandatory for Big data and analytics within their business environment. The technical skills that were stated by executives and senior managers were programming, statistical analysis, Cloud computing, data analysis, data modelling, machine learning, data governance, data science and data engineering.

Six participants confirmed these skills to be a minimum requirement for the role when recruiting graduates for Big data and analytics roles. For instance, **Participant 1** and **Participant 3** asserted that "*the base programming tools used by most banking institutions are Python and R.*" All participants confirmed that practical experience is severely lacking by graduates when entering the workplace.

Participant 5 and **Participant 10** included additional programming languages such as Statistical Analysis System (SAS) and Java for the use of advanced analytics, statistical analysis, and predictive analytics. **Participant 1** asserted that there is a big demand for Big data and analytics skills *"what we are seeing is that after 18 to 24 months they leave and get increases of more than 30 or 40%. Many of the applicants are self-taught on the latest technologies through these online learning platforms."*

Participant 1 articulated the role of a data scientist as involving the capacity to "manipulate data, extract insights from it, and, in my opinion, the role is complex as it requires someone who comprehends both the data and the tools, possesses the capability to interpret the data, and also has an understanding of the specific business domain. So, they need to understand the problem they're trying to solve so that they derive the correct insights for business."

Participant 4 described some of the attributes that collectively contribute to the effectiveness of the data engineers and data engineering team *"From a technical capability, people with more experience in assessing relevant tools, platforms, etcetera, understanding of warehouse, data lakes and then data quality tools, you know, so, what data quality tool is the best."*

Participant 9 perceived that whilst there was a lack of skills and practical experience there was an overall improvement in the graduates that were recruited. *"We have seen, like, really strong candidates coming through the system over the last few years, so that might be testament to the fact that the universities have improved their game in that respect, so I think it would actually make sense to work together to make the syllabus more practical."*

Participant 11 succinctly explained the core distinction between traditional analytics and big data analytics using the following illustration: *"In the traditional approach, the data structure is established beforehand, and then the data is organised accordingly, with column headings being defined..."Specify the number of rows you have included...Once the necessary structure is established, all the data is imported. However, with big data, the approach is to first load all the data into the database without considering its appearance. Then, patterns are identified,* *and based on these patterns, a structure is created. This requires a shift in mindset. This explanation emphasises a distinct methodology that shifts from a mostly organised approach relying on existing knowledge to a more spontaneous and investigative approach primarily influenced by the data.”*

Expanding on **Participant 11's** claim that *"a completely new set of skills is necessary for analysts,"* **Participant 5** contended that "the essential abilities and capacities we need are fundamentally distinct. *"Traditionally, the role of a data base administrator (DBA) was essential. However, in the present, the demand has shifted towards data engineers and data scientists. The required resources and expertise for shifting from conventional analytics to big data analytics are fundamentally distinct.”*

The statements made by **Participants 5** and **Participant 11** validated the notion that the skillset needed for big data analytics is completely distinct from the skill set of traditional analysts. In the past, database administrators were necessary, however, now there is a demand for data scientists and data engineers. The necessity of a comprehensive cultural transformation and shift in perspective, along with the acquisition of a distinct technical talent, suggests that it may be more feasible to train and introduce new personnel into the system for big data, rather than trying to convert typical analysts into big data analysts.

4.3.1.3 Theme 3: Business Skills

Most of the participants stated that having business skills is essential for deriving meaningful business outcomes. Some of the emphasis provided by the participants are as follows: **Participant 2** “So I mean, even if the institution of higher learning was giving students sort of examples or use cases, they would still struggle when they come into a corporate environment to apply it you know, they actually need the business knowledge”. **Participant 4** “What is ethics around data, that type of stuff is a little bit lacking when they join the organisation". **Participant 5** “graduates wouldn't really have business experience, but you know, business knowledge, like, economics, that kind of thing is important. Business and commercial acumen generally comes with people with experience”. **Participant 6** “You cannot service business unless you understand you know what their business is about". **Participant 7** “from my experience, I do think in some of the banking industry you do get the combination of business and technical skill, but there's definitely a gap when it comes to central banking". **Participant 10** “just having to understand how business operates takes a bit of time. So that need to understand business is there even not just with graduates, but even with, you know, competent data analyst, data engineers, they're not quite in tune with the business and then takes a little bit of time to feel what the business requires". Two of the participants confirmed that understanding business strategy and being able to link this to projects and use cases assigned. **Participant 5** "so if they understand strategy, this enables the data scientists to make informed decisions about which data projects or problems to prioritize". **Participant 10** "So, but you need to understand the business stuff that you need to really know what the business is strategy is and what value you need to get out of it".

Emphasis was placed on linking these use cases to ensure that business problems and opportunities are addressed. Furthermore, business domain knowledge was strongly emphasized as an important contributor to value been derived due to the understanding of the business context of the problem.

4.3.1.4 Theme 4: Interpersonal Skills

In addition to the technical and business skills highlighted above, all participants emphasised the importance of interpersonal skills as vital qualities necessary for dealing with big data and carrying out analytical tasks. Interpersonal skills that were identified include communication, problem-solving, presentation skills, storytelling, customer-centric, negotiation, continuous learning, critical thinking, and teamwork. Some of the emphasis provided by the participants are as follows:

According to **Participants 1** and **3**, continual learning enables individuals to develop their technical skills, which can lead to meaningful business outcomes. **Participant 1** stated: *"The skills are sufficient, however, continuously learning and developing your communication and presentation skills. You need to be able to translate the technical findings into meaningful business outcomes.".* And **participant 3** quotes as flows: **Participant 3** *"theory is there but practical problem solving is an issue. Hypothesis build is in a challenge. Computer science degreed students are not strong problem solvers."*

Participant 5 "The next thing that you need to work with is understanding language. This is the softer skill of understanding what the problem is and what is the expected result."

Participant 4 "So the youngsters get quite flabbergasted between all the complexity around managing and learning the existing on prem solutions and now the new cloud, where seasoned individuals manage better because I've been around."

Participant 8 "You must keep on like it's a continuous evolution. So, if you don't have that culture of continuous learning, you're going to fall flat soon because I think the world is moving faster than we are."

4.3.2 Skill gap in Big data analytics in the banking sector

The second objective of this study is to assess the skills gap in Big data analytics in the banking sector. The themes related to the objectives are presented in Table 2.

4.3.2.1 Theme 2: In-Demand Technical Skills

Under this objective, the participants confirmed that there is a skills gap in Big data and analytics in the financial services industry. The senior managers cited the lack of practical experience as a primary as majority contributing factor to the skills gap. A small number of the participants believed that graduates obtained an adequate skillset from their higher education institution to meet work needs. However, the primary obstacle in facilitating an easy transition into their roles was the absence of practical experience. The skills that were highlighted by the participants were divided into technical and interpersonal skills. The skills lacking were problem-solving, the ability to utilize analytical tools, machine learning, programming capabilities, methodologies, data science, and data engineering.

Participant 4 highlighted that global markets have been successful at adapting to this much quicker "in the South African context where there's a big skills gap in the country, simply the reason is that we are not adopting big data analytics tools and technologies quick enough, technologies have evolved into really high-end data analytics and academic institutions and business have to play catch-up."

Participant 14 also deemed that graduates are not exposed to the various tools and technologies that are used by industry whilst completing their academic qualifications *"... they have formal training, it's not substantive enough and a lack of experience in analytical tools, right, having, you know, use big data in terms of the analysis, analytics and producing insights."*

Participant 2 asserted “Universities don't prep students on technologies and coding, methodologies such as machine learning, feature engineering, data governance frameworks and processes that corporate use, the learning curve takes 2 years, "it becomes so difficult that they start looking for new opportunities and leave."

Participant 4 asserted “They're not fully there yet from a full scope of a data management life cycle on how you use reference data, metadata, ethics around data.

Participant 4 emphasised the current skills challenge within their immediate business unit “firstly, the skillset that we’re looking for data engineering capabilities in the Big data domain, is where the biggest skills gap. We have a lot of CVs of individuals that can do data analysis and data analytical work on BI dashboards but the hardcore work of developing pipelines, machine learning capabilities is where we have the biggest gap. Practical experience is lacking. Data management lifecycle, understanding reference and master data. The tools are rapidly changing and it’s difficult to keep up.”

Participant 2 asserted “we don’t have robust education and training. Understanding of the underlying methodologies of the tools and coding language used. Learning curve in terms of how the technologies work. Resources change jobs frequently to obtain exposure in the different technologies.”

According to **Participant 1**, “graduates do have the necessary degrees; however, the gap needs to be bridged with training on the various technologies and interpersonal skills. Business domain in which they work.”

4.3.2.2 Theme 3: Business Skills

Participant 4 expressed a significant concern regarding the candidates' insufficient business acumen. The participant expressed their frustration with the difficulty of implementing Big Data Analytics in specific business sectors, such as global markets. They observed that while candidates possessed some technical skills in utilising Big data analytics, they lacked a comprehension of the business aspects of financial markets that would allow them to contribute effectively.

Participant 2 asserted that a strategic shift towards recruiting graduates with completed qualifications in economics, statistics, and mathematics yielded superior results compared to hiring graduates with some technical experience but lacking domain knowledge.

4.3.2.3 Theme 4: Interpersonal Skills

Majority of participants reported that the quality of insights derived from Big data is significantly diminished due to several factors: (1) inadequate definition of analytics scope, (2) analysts lacking understanding of business requirements, (3) analysts lacking the understanding of the business domain, and (4) insights not being presented in a manner that effectively addresses business needs.

For instance, **Participant 7** asserted that teamwork and collaboration with the business users can assist with a better understanding of the objectives and the scope of the problem and requirements.

Participant 2 "So I would say, subject matter domain expertise would be very important, it helps the person understand the problem they're solving for and test the results against this."

Participant 8 "The problem is that when you run a POC, as long as you can communicate this clearly, show business value and be sound in showing your train of thought, it's easier for these things to sell to business."

Participant 9 "basic problem solving and communicating the results has been quite a challenge with the newly appointed grads. They don't use the correct terminology as they have not been exposed to the specific business domain It is also important for them to determine upfront which industry, they would like to specialize in obtain exposure in these areas."

Twelve participants stressed that the less experienced individuals lack problem-solving capabilities more so the ability to persevere when experiencing more working with more complex type models.

Participant 4 "critical thinking is also poorly lacking, like technical aptitude, they can do low-code, no-code, but problem solving, and design thinking is not there."

The study participants also highlighted the challenge of continuous learning where there is limited capacity for newly appointed graduates to learn existing technologies and to continuously upskill themselves in the new tools and technologies that are evolving across the industry.

Participant 3 reported that consulting companies are onboarded to assist with meeting the demands of the business. The individuals with more extensive experience demonstrate a heightened comprehension of concepts, enabling them to adapt and acquire proficiency in new tools at an accelerated pace.

Some participants voiced concerns about the lack of compelling narrative and presentation abilities. **Participant 13** highlighted the crucial importance of analysts being able to express their findings in a way that non-technical stakeholders can understand, which helps in making well-informed decisions.

Participant 12 "they need the ability to translate the technical findings into a more business type language so that the end user can use this for decision making."

Participant 5 "confidence is built over time and through experience, it's important that the data science grads be exposed to presenting their results as they are at the forefront with the business stakeholders. Data engineers focus more on building the backend pipelines for business."

4.3.3 Results for Research Objective 3

The third objective of this study explores the roles of academic institutions in closing the Big data skills gap. The themes relate to Table 2.

All the participants emphasised the crucial significance of cultivating partnerships and sustaining continuous collaboration between higher education institutions and the banking sector. The collaborative approach, as revealed by the participants is considered crucial to achieve a harmonious alignment between the capabilities required by the business and those provided by educational institutions. Furthermore, the participants stressed the importance of including industry experts in academia as a strategic approach to improve the collaboration and compatibility between academic programmes and the evolving needs of the banking industry.

Participant 14 asserted "I think what the banking sector needs to do, to put a framework to engage with institutions that probably are offering degrees that are targeted at the banking sector to see, what these programs entail in terms of big data analytics and then assess if there is a gap there and see how that gap can be fulfilled."

Executives have proposed that institutions of higher learning actively engage with the banking industry and vice versa to ascertain the specific skills deemed essential for addressing both current and future needs. It is recommended that educational institutions strategically tailor their curricula to align with these identified skill requirements. Additionally, executives emphasized the utilization of industry forums, such as the Banking Association of South Africa (BASA) and BankSETA, as valuable platforms for fostering collaboration between institutions of higher learning and the banking sector. Such collaborations are seen as instrumental in collectively addressing and mitigating the existing skills gap.

Participant 2 asserted that "I find myself spending more time in the industry at conferences and listening to these other data practitioners as opposed to spending time with some of these higher learning institutions and shaping future data practitioners, and the, that's where I believe personally, we are missing it."

Participant 8 asserted that “If you look at the whole, information management data value chain, Institutions of higher learning are not covering that. So academic institutions need to adapt quickly to the changing environment”.

Participant 2 asserted that “I think that international institutions of higher learning, yes, their curriculum is quite comprehensive, but they are more mature. They have worked well with industry in some countries”.

Participant 4 “I think there are pockets of institutions that are dabbling in Like, Stellenbosch now has a master’s in data science. So, some are starting to explore a little".

4.4 Discussion of Findings

4.4.1 Skill and competencies for Big data analytics

From objective one, which encompasses the skills and competencies for Big data analytics, the technical competencies delineated by all participants encompass programming, statistical analysis, Cloud computing, data analysis, data modelling, machine learning, data governance, data science, and data engineering. While these proficiencies were acknowledged as being in demand within the banking industry, the participants underscored a discernible gap, particularly in the domain of interpersonal skills, as these are notably absent from higher education curricula. Based on the results, the data scientist is highlighted as an essential and dynamic resource who possesses extensive technical expertise in software technologies and data structures. The data scientist is responsible for manipulating both structured and unstructured data and must possess the business acumen necessary to extract and deliver relevant insights to the business in a comprehensible fashion. This enables them to effectively solve business problems.

Furthermore, interpersonal skills, including communication, problem-solving, presentation skills, storytelling, negotiation, continuous learning, and teamwork, were identified as positively correlating with organizational performance. This underscores the imperative for the banking industry to promptly accord precedence to the development of these interpersonal skills, recognizing their critical role in adapting to the evolving industry landscape and fostering human capital development. These findings are in line with a previous study conducted by Dean & East (2019) which suggests that technical skills are no longer sufficient for workers to compete in the work environment.

A study by Samuel & Moagi (2022) suggested that it is crucial for South Africa to begin and speed up skills development plans that would enable its entire workforce to effectively operate in the next industrial revolution. A study by Mabe & Bwalya (2022) determined that although technical skills are crucial for creating and executing Big data applications, the current era of postmodernity requires that interpersonal skills be paired with these technical abilities in order to successfully incorporate technology into different socio-economic value chains.

4.4.2 Skills gap in Big data analytics in the banking sector

Based on research objective 2, which aimed to assess the skills gap in Big data analytics in the banking sector, this study underscores the pressing need for highly specialized technical skills in big data analytics within the banking sector. Mastery in computer languages, data manipulation, machine learning methods, and data visualisation tools are essential for data analysis, trend analysis predictive analysis, and data-driven decision making. The investigation uncovers discrepancies in the levels of technical proficiency among professionals, emphasising a significant disparity that presents obstacles to fully utilising sophisticated analytics for decision-making in financial institutions.

In addition to technical expertise, the assessment highlights the need to have good business acumen as a crucial element of a comprehensive set of skills. Graduates need to have a profound comprehension of banking operations and the capacity to synchronise data analytics activities with broader company goals. The findings underscore the need for training programmes that connect technical and business expertise, empowering analytics experts to convert insights into practical initiatives that enhance the financial prosperity and long-term viability of the banking sector. The research identifies a critical deficit in interpersonal skills among big data analytics professionals within the banking sector. Proficient communication, cooperation, and the capacity to articulate intricate technical knowledge to individuals without technical expertise are of utmost importance. The data analysis demonstrates that the lack of these interpersonal skills obstructs the smooth incorporation of data-driven insights into decision-making processes, thereby restricting the overall effectiveness of analytics projects.

The technical competencies emphasized by interview participants encompass programming, machine learning, data engineering, data analytics, and data science. These skills were identified as pivotal prerequisites for roles in big data and analytics within the banking sector. A study conducted by (Tokarcikova & Malichova, 2020), revealed that employers and students advocate for heightened educational initiatives to enhance these skills to better prepare graduates by providing innovative teaching methods.

The essential soft skills highlighted by the interview participants encompass problem-solving, communication, and proficiency in both oral and written presentations, along with a strong emphasis on teamwork. These identified soft skills align with existing research conducted by Ping Yong and Ling (2023), indicating that proficiency in problem-solving, collaborative abilities, and teamwork complements technological expertise. These skills collectively contribute to a more seamless and effective transition for graduates entering the workplace.

4.4.3 Role of academic institutions in closing the Big data skills gap

Based on research objective 3, which aimed to identify the role of academic institutions in closing the Big data skills gap, tertiary institutions play a crucial role in providing the necessary skills to fulfill industrial needs and promote lifelong learning. However, closing the skills gap requires continuous collaboration and engagement between academics and industry. In order to create this cooperative relationship, it is necessary for both higher education institutions and the industry to transform into organisations that prioritise continuous learning.

The findings revealed that educating individuals in emerging skill sets demands the adoption of novel and innovative pedagogical approaches. Respondents in this study advocated for the augmentation of teaching methodologies, endorsing strategies such as simulation-based or experiential learning, and the incorporation of vocational training. These approaches were highlighted as instrumental in furnishing participants with practical experience and indispensable skills requisite for the contemporary work milieu. It was emphasized that for vocational training to be efficacious within the digital economy, it should be meticulously tailored to align with the current technological landscape. This adaptation should encompass the integration of cognitive and soft skills, deemed crucial in navigating the intricacies of a technologically driven economy.

The findings show that higher education institutions have faced challenges in keeping up with the technical progress witnessed in the banking sector's digital revolution. Especially in higher education, the primary emphasis is on teaching theoretical knowledge and technical proficiency, sometimes overlooking the need for crucial interpersonal skills required by employers (Cunningham & Villasenor, 2016). Participants observed that firms mitigate this disparity by implementing internal learning initiatives, while higher education institutions fail to sufficiently prepare graduates with these competencies. In the present technological milieu, the significance of interpersonal skills is amplified, as they transcend the realm of digitalization or automation (Frey & Osborne, 2017; PwC, 2017b).

Consistent with findings from comparable research studies (Pauw et al., 2006; Yu, 2013), senior managers in this study expressed the obstacle presented by a lack of practical experience during the transfer from academics to the workplace. It has been noticed that higher education institutions largely prioritise theoretical knowledge, however, graduates frequently have difficulties when it comes to applying this academic framework to real-life job situations. This observation suggests that the current output of higher education institutions may not completely match the demands of the modern work environment. As a result, employers may need to provide additional training to address the disparity between theoretical knowledge and practical skills.

4.5 Chapter Summary

This chapter offers a comprehensive analysis, presentation, and discussion of the findings of this study. It is noted that the relevance of a degree program also depends on the specific roles within Big data, as the field encompasses various areas such as data engineering, data science, machine learning, and more. Additionally, practical experience, internships, and proficiency in relevant tools and technologies play a crucial role in preparing graduates for Big data roles. Many universities also offer specialized master's programs in Big data analytics or related fields for those seeking advanced, targeted training in big data technologies.

5. Conclusion and Recommendations

5.1 Introduction

This research endeavoured to assess and ascertain the requisite skills mandated by the banking industry in South Africa. The primary objective is to comprehend the existing gaps and cultivate the necessary competencies essential for meeting industry demands. This concluding chapter provides a concise overview of the primary research discoveries that address the research inquiries and highlight the impact on both business and theory. Additionally, it provides advice to industry and higher education institutions, as well as other stakeholders, on how to address the skills gap in the field of big data and analytics. Additionally, the chapter finishes by emphasizing the limitations of the study and suggesting potential areas for future research.

5.2 Conclusion

5.2.1 Summary

The main objective of this study was to comprehend the necessary skills and competencies, evaluate the deficiency in abilities, and ascertain the role of educational institutions and the banking industry in bridging the gap in skills for Big data analytics. The researcher was able to achieve the study's aims by combining a literature review and empirical inquiry.

Firstly, this study identified the key technical and interpersonal skills of the banking industry requires now and in the future. These include programming, statistical analysis, Cloud computing, data analysis, data modelling, machine learning, data governance, data science, and data engineering. These skills have been described in other studies and have been shown to improve business performance (Mikalef & Krogstie, 2019). There is a scarcity of research that has evaluated the issue of the skills gap in Big data analytics from the viewpoints of both industry and educational institutions. This paper contributes to the lack of research in the literature.

Secondly, this study evaluated and confirmed the presence of a skills gap in the South African banking industry specifically in the area of Big data analytics. The participants believed that educational institutions failed to provide graduates with the necessary vital technical and interpersonal skills to effectively adjust to their evolving environment. Consequently, businesses had to provide additional training to ensure that employment needs were met. The skills gap should be addressed by institutions of higher learning, as they are responsible for developing and instilling the essential skills needed for employability. The industry should engage in collaboration with educational institutions to identify and disseminate the necessary new skills and maintain a consistent alignment between the demand and supply of skills. Employers have utilised internal training, mentorship initiatives, and online learning platforms to enhance the skills of their staff. A noteworthy discovery was that banks also collaborate with consulting firms to address the scarcity of talent, highlighting the significance of symbiotic or strategic collaborations.

Finally, the study offered valuable insights into the strategies that both industry and educational institutions need to employ in order to address the current and future skills gap. Both the industry and the educational institution offered suggestions regarding the respective roles that both parties should assume in bridging the skills gap. The study demonstrates that effectively bridging the skills gap necessitates ongoing collaboration and synchronisation between academia and industry.

5.2.2 Recommendations

As the financial industry continues to evolve, the importance of leveraging data and analytics for informed decision-making becomes increasingly evident. To ensure that banks stay at the forefront of innovation, the recommendations implementing a comprehensive upskilling program for graduates in the areas of big data and analytics.

5.2.3 Recommendations to Banks

Based on the findings of this research, the following suggestions for banks recommended:

The first recommendation is for banks to forge collaborative alliances with universities and educational institutions to formulate curriculum content that aligns intricately with the distinctive requisites and demands of the banking industry. Banks need to share knowledge about the essential skills and competencies that are required now and, in the future, to assist academia in reshaping the curriculum to meet the demands.

The second recommendation is for banks to assist institutions of higher learning with case studies, practical projects, and examples related to the sector in the coursework. Incorporate assessments that simulate real-world scenarios to evaluate their proficiency in applying big data analytics skills in practical contexts. Provide continuous feedback to students concerning their performance and progression relating to the practical projects and assessments.

The third recommendation is to initiate Internship programs in the first year of the degree program, affording students the chance to participate in "winter" or "summer" schools and engage in vacation work. This approach facilitates the acquisition of practical experience and the refinement of interpersonal skills, thereby enhancing their preparedness for the workplace. Employers have been upskilling and reskilling employees to ensure that their workforce is equipped with the relevant skillet. Employers have been building their own academies and using internal training such as internships, graduate programmes, and other learning platforms to upskill and reskill their employees.

The fourth recommendation is that banks support graduates and institutions of higher learning with access to cutting-edge tools and technology utilized in big data analytics within the banking sector. This may involve granting users access to software platforms, cloud services, and analytical tools. Financial assistance is provided to students to obtain industry certifications in the field of big data analytics and Cloud Computing technologies. These credentials authenticate their expertise and enhance their desirability to prospective employers.

The fifth recommendation suggests that banks organize guest lectures and industry speakers to offer valuable insights into the practical implementation of Big data analytics within the banking sector. Engaging with professionals from the industry provides students with invaluable perspectives on emerging trends, new tools, and hands-on expertise.

The sixth recommendation suggests that banks collaborate and partner with secondary schools to introduce or enhance programming modules and include Big data in their educational curricula. This equips high school students with a strong foundation in data science for progressive proficiency. Students will learn problem-solving, critical thinking, collaboration, and communication skills which is a critical requirement for graduates transitioning into the workplace to be effective in their roles.

Lastly, establish mentorship initiatives at secondary and tertiary levels that facilitate the pairing of students with professionals from the financial sector. Mentors can offer direction, impart industry insights, and provide recommendations on cultivating the requisite abilities for excelling in big data analytics positions.

5.2.4 Recommendations to Institutions of Higher Learning

Based on the findings of this research, the following suggestions for institutions of higher education recommended:

The first recommendation is for institutions of higher learning to redesign and continuously enhance the academic curriculum and teaching methodologies to reflect current trends and technologies. Specialised courses and modules to be incorporated into the curriculum for data science, machine learning, and design thinking. Institutions of higher learning to include soft skills as part of the curriculum since there is a strong emphasis on the relevance of these skills.

Secondly, partnerships must be formed with business entities in the financial services industry including consulting companies to obtain access to their technologies and licenses to enable students to obtain exposure and practical applications. Collaboration with industry to enhance the practical relevance of assignments and projects in the academic qualifications suggests a purposeful incorporation of real-world scenarios and technologies obtained directly from the banking sector. By integrating industry-relevant technologies into the curriculum, students will gain practical experience and become familiar with and stay informed about the modern tools used in the banking industry. These tools and technologies should be provided in a research and development (R&D) environment for students to stay ahead of the curve in terms of Big data tools and technologies. Employers have been partnering with FinTechs to reskill and upskill employees. This emphasises the importance of collaboration and strategic partnerships.

The third, recommendation is to develop evaluation standards that not only examine theoretical knowledge but also assess the capacity to apply learnings to practical situations. It is advisable to include project-based examinations that replicate real-world company problems.

All participants offered recommendations regarding the obligations that stakeholders should assume in addressing the skills shortage. The proposals involved promoting collaboration between higher education institutions and industry to align curriculum content and design with industry needs, incorporating industry experts into academic environments, improving teaching methods, and including assessments that involve real-world challenges in the current context. Furthermore, the proposals underscored the significance of providing direction and mentorship to the individuals engaged.

5.3 Research Limitations

The limitations of this study involve a small sample size obtained from a specific geographical area, hence preventing the extrapolation of results to the wider South African banking industry. Furthermore, the study's cross-sectional methodology captures a single moment in time, offering insights into viewpoints at a specific point. This suggests that talents considered essential at the current moment may change and develop over time. Furthermore, given the ever-evolving nature of technology, the skill requirements in the banking sector are likely to undergo alterations in the future. Moreover, the presence of potential interviewer bias and differing levels of interviewing expertise may have impacted the process of gathering data and the subsequent analysis of the findings.

5.4 Recommendation for Further Research

Additional research should involve research scope to encompass a more heterogeneous sample from different geographical locations within South Africa. Examine the influence of regional differences on the perceived proficiency in big data analytics and evaluate the extent to which the findings can be applied to the wider banking industry. Finally, future research should conduct longitudinal investigations to ascertain the progressive development of important abilities in tandem with advancing technologies. By broadening the research endeavours, there is an opportunity to augment the comprehensiveness, scope, and practicality of knowledge pertaining to big data analytics proficiency in the South African banking sector. This will contribute to well-informed decision-making for both academic institutions and industry stakeholders.

REFERENCES

Akaranga, S. I., & Makau, B. K. (n.d.). Ethical Considerations and their Applications to Research: A Case of the University of Nairobi. . . *Pp*, *3*.

Akhtar, P., Frynas, J. G., Mellahi, K., & Ullah, S. (2019). Big Data-Savvy Teams' Skills, Big Data-Driven Actions and Business Performance. *British Journal of Management*, *30*(2), 252–271. https://doi.org/10.1111/1467-8551.12333

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

Asenahabi, B. M. (2019). *Basics of Research Design: A Guide to selecting appropriate research design*. *6*(5).

Aziz, N. A., Long, F., & Wan Hussain, W. M. H. (2023). Examining the Effects of Big Data Analytics Capabilities on Firm Performance in the Malaysian Banking Sector. *International Journal of Financial Studies*, *11*(1), 23. https://doi.org/10.3390/ijfs11010023

Barham, H. (2017b). Achieving Competitive Advantage Through Big Data: A Literature Review. *2017 Portland International Conference on Management of Engineering and Technology (PICMET)*, 1–7. https://doi.org/10.23919/PICMET.2017.8125459

Barney (1991).pdf. (n.d.).

Bryman_and_bell_research_methods.pdf. (n.d.).

Chahal, H., Gupta, M., Bhan, N., & Cheng, T. C. E. (2020). Operations management research grounded in the resource-based view: A meta-analysis. *International Journal of Production Economics*, *230*, 107805. https://doi.org/10.1016/j.ijpe.2020.107805

Dachyar, M., Zagloel, T. Y. M., & Saragih, L. R. (2019). Knowledge growth and development: Internet of things (IoT) research, 2006–2018. *Heliyon*, *5*(8), e02264. https://doi.org/10.1016/j.heliyon.2019.e02264

Dai Thich Phan1, & Lam Quynh Trang Tran. (n.d.). BUILDING A CONCEPTUAL FRAMEWORK FOR USING BIG DATA ANALYTICS IN THE BANKING SECTOR. *Intellectual Economics*.

De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review*, *65*(3), 122–135. https://doi.org/10.1108/LR-06-2015-0061

Dean, S. A., & East, J. I. (2019). Soft Skills Needed for the 21st-Century Workforce. *International Journal of Applied Management and Technology*, *18*(1). https://doi.org/10.5590/IJAMT.2019.18.1.02

Dosso, D., Ferilli, S., Manghi, P., Poggi, A., Serra, G., & Silvello, G. (2021). Information and Research Science connecting to Digital and Library Science: Report on the 17th Italian Research Conference on Digital Libraries, IRCDL2021. *ACM SIGMOD Record*, *50*(2), 44–47. https://doi.org/10.1145/3484622.3484635

Drydakis, N. (2022). Artificial Intelligence and Reduced SMEs' Business Risks. A Dynamic Capabilities Analysis During the COVID-19 Pandemic. *Information Systems Frontiers*, *24*(4), 1223–1247. https://doi.org/10.1007/s10796-022-10249-6

Dubey, R., Gunasekaran, A., Childe, S. J., Blome, C., & Papadopoulos, T. (2019). Big Data and Predictive Analytics and Manufacturing Performance: Integrating Institutional Theory, Resource-Based View and Big Data Culture. *British Journal of Management*, *30*(2), 341–361. https://doi.org/10.1111/1467-8551.12355

Dworkin, S. L. (2012). Sample Size Policy for Qualitative Studies Using In-Depth Interviews. *Archives of Sexual Behavior*, *41*(6), 1319–1320. https://doi.org/10.1007/s10508-012-0016-6

EDSA-OEGlobal17.pdf. (n.d.).

Eriksson, P., & Kovalainen, A. (n.d.). *Qualitative Methods in Business Research*.

Ferreira, C., Robertson, J., & Pitt, L. (2023). Business (un)usual: Critical skills for the next normal. *Thunderbird International Business Review*, *65*(1), 39–47. https://doi.org/10.1002/tie.22276

Fisher, D., DeLine, R., Czerwinski, M., & Drucker, S. (2012). Interactions with big data analytics. *Interactions*, *19*(3), 50–59. https://doi.org/10.1145/2168931.2168943

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

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

Hasan, M., Hoque, A., & Le, T. (2023). Big Data-Driven Banking Operations: Opportunities, Challenges, and Data Security Perspectives. *FinTech*, *2*(3), 484–509. https://doi.org/10.3390/fintech2030028

He, W., Hung, J.-L., & Liu, L. (2022). Impact of big data analytics on banking: A case study. *Journal of Enterprise Information Management*. https://doi.org/10.1108/JEIM-05-2020-0176

Henry, R., & Venkatraman, S. (2015). *BIG DATA ANALYTICS THE NEXT BIG LEARNING OPPORTUNITY*. *18*(2).

Hugh J., W. (2019). Update Tutorial: Big Data Analytics: Concepts, Technology, and Applications. *Communications of the Association for Information Systems*, 364–379. https://doi.org/10.17705/1CAIS.04421

Huttunen, J., Jauhiainen, J., Lehti, L., Nylund, A., Martikainen, M., & Lehner, O. M. (2019). *BIG DATA, CLOUD COMPUTING AND DATA SCIENCE APPLICATIONS IN FINANCE AND ACCOUNTING*.

Hyman, D. M. R. (2016). *Open- versus Close-Ended Survey Questions*. *14*(2).

Jagani, N., Jagani, P., & Shah, S. (2021). BIG DATA IN CLOUD COMPUTING: A LITERATURE REVIEW. *International Journal of Engineering Applied Sciences and Technology*, *5*(11). https://doi.org/10.33564/IJEAST.2021.v05i11.029

Kubina, M., Varmus, M., & Kubinova, I. (2015b). Use of Big Data for Competitive Advantage of Company. *Procedia Economics and Finance*, *26*, 561–565. https://doi.org/10.1016/S2212-5671(15)00955-7

Lee, I. (2017). Big data: Dimensions, evolution, impacts, and challenges. *Business Horizons*, *60*(3), 293–303. https://doi.org/10.1016/j.bushor.2017.01.004

Lindgren, B.-M., Lundman, B., & Graneheim, U. H. (2020). Abstraction and interpretation during the qualitative content analysis process. *International Journal of Nursing Studies*, *108*, 103632. https://doi.org/10.1016/j.ijnurstu.2020.103632

Mabe, K., & Bwalya, K. J. (2022). Critical soft skills for information and knowledge management practitioners in the fourth industrial revolution. *SA Journal of Information Management*, *24*(1). https://doi.org/10.4102/sajim.v24i1.1519

Malaka, I., & Brown, I. (2015). Challenges to the Organisational Adoption of Big Data Analytics: A Case Study in the South African Telecommunications Industry. *Proceedings of the 2015 Annual Research Conference on South African Institute of Computer Scientists and Information Technologists*, 1–9. https://doi.org/10.1145/2815782.2815793

Marshall, M. N. (n.d.). *Sampling for qualitative research*.

Maruyama, H., Kamiya, N., Higuchi, T., & Takemura, A. (2015). *Developing Data Analytics Skills in Japan: Status and Challenge*.

Mathias Kalema, B., & Mokgadi, M. (2017). Developing countries organizations' readiness for Big Data analytics. *Problems and Perspectives in Management*, *15*(1), 260–270. https://doi.org/10.21511/ppm.15(1-1).2017.13

Michalisin, M., Kline, D., & Smith, R. (1970). Intangible Strategic Assets and Firm Performance: A Multi-Industry Study of The Resource-Based View. *Journal of Business Strategies*, *17*(2), 91–117. https://doi.org/10.54155/jbs.17.2.91-117

Mikalef, P., & Krogstie, J. (2019). Investigating the Data Science Skill Gap: An Empirical Analysis. *2019 IEEE Global Engineering Education Conference (EDUCON)*, 1275–1284. https://doi.org/10.1109/EDUCON.2019.8725066

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: A systematic literature review and research agenda. *Information Systems and E-Business Management*, *16*(3), 547–578. https://doi.org/10.1007/s10257-017-0362-y

Mohd Arifin, S. R. (2018). Ethical Considerations in Qualitative Study. *INTERNATIONAL JOURNAL OF CARE SCHOLARS*, *1*(2), 30–33. https://doi.org/10.31436/ijcs.v1i2.82

Neves, P. C., Schmerl, B., Cámara, J., & Bernardino, J. (2016). Big Data in Cloud Computing: Features and Issues: *Proceedings of the International Conference on Internet of Things and Big Data*, 307–314. https://doi.org/10.5220/0005846303070314

Noble, H., & Smith, J. (2015). Issues of validity and reliability in qualitative research. *Evidence Based Nursing*, *18*(2), 34–35. https://doi.org/10.1136/eb-2015-102054

Oluwajodu, F., Blaauw, D., Greyling, L., & Kleynhans, E. P. J. (2015). Graduate unemployment in South Africa: Perspectives from the banking sector. *SA Journal of Human Resource Management*, *13*(1), 9 pages. https://doi.org/10.4102/sajhrm.v13i1.656

Poornima, S., & Pushpalatha, M. (2016). *A JOURNEY FROM BIG DATA TOWARDS PRESCRIPTIVE ANALYTICS*. *11*(19).

Provost, F., & Fawcett, T. (2013a). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, *1*(1), 51–59. https://doi.org/10.1089/big.2013.1508

Provost, F., & Fawcett, T. (2013b). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, *1*(1), 51–59. https://doi.org/10.1089/big.2013.1508

Provost, F., & Fawcett, T. (2013c). Data Science and its Relationship to Big Data and Data-Driven Decision Making. *Big Data*, *1*(1), 51–59. https://doi.org/10.1089/big.2013.1508

Quasim, M. T., Johri, P., Meraj, M., & Haider, S. W. (2019). *5 V'S OF BIG DATA VIA CLOUD COMPUTING: USES AND IMPORTANCE*.

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

RBV_Strategic_Approach_MIT.pdf. (n.d.).

Report on Skills Supply and Demand in South Africa—2022 (1).pdf. (n.d.).

Sadiku, M. N. O., Eze, K. G., & Musa, S. M. (2018). *The Essence of Data Engineering*. *5*.

Samuel, O. M., & Moagi, T. (2022). The emerging work system and strategy for skills transition in South Africa. *Management Research Review*, *45*(11), 1503–1523. https://doi.org/10.1108/MRR-05-2021-0356

Saunders, M. N. K., Lewis, P., & Thornhill, A. (2023). *Research methods for business students* (Ninth edition). Pearson.

Scoping-Study-Skills-Gap-in-the-Banking-Sector-in-South-Africa..pdf. (n.d.).

Shakya, S., & Smys, S. (2021). Big Data Analytics for Improved Risk Management and Customer Segregation in Banking Applications. *Journal of ISMAC*, *3*(3), 235–249. https://doi.org/10.36548/jismac.2021.3.005

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, *22*(2), 63–75. https://doi.org/10.3233/EFI-2004-22201

Smaya, H. (2022a). The Influence of Big Data Analytics in the Industry. *OALib*, *09*(02), 1–12. https://doi.org/10.4236/oalib.1108383

Smaya, H. (2022b). The Influence of Big Data Analytics in the Industry. *OALib*, *09*(02), 1–12. https://doi.org/10.4236/oalib.1108383

Smith, K. A., Vasudevan, S. P., & Tanniru, M. R. (1996). Organizational learning and resource-based theory: An integrative model. *Journal of Organizational Change Management*, *9*(6), 41–53. https://doi.org/10.1108/09534819610150512

Sofaer, S. (n.d.). *Qualitative research methods*.

Soltani Delgosha, M., Hajiheydari, N., & Fahimi, S. M. (2021). Elucidation of big data analytics in banking: A four-stage Delphi study. *Journal of Enterprise Information Management*, *34*(6), 1577–1596. https://doi.org/10.1108/JEIM-03-2019-0097

Talwar, S., Kaur, P., Fosso Wamba, S., & Dhir, A. (2021). Big Data in operations and supply chain management: A systematic literature review and future research agenda. *International Journal of Production Research*, *59*(11), 3509–3534. https://doi.org/10.1080/00207543.2020.1868599

The Role of Big Data ^0 Predictive Analytics in the Employee Retention A Resource Based View (1).pdf. (n.d.).

Torres, R., Sidorova, A., & Jones, M. C. (2018a). Enabling firm performance through business intelligence and analytics: A dynamic capabilities perspective. *Information & Management*, *55*(7), 822–839. https://doi.org/10.1016/j.im.2018.03.010

Torres, R., Sidorova, A., & Jones, M. C. (2018b). Enabling firm performance through business intelligence and analytics: A dynamic capabilities perspective. *Information & Management*, *55*(7), 822–839. https://doi.org/10.1016/j.im.2018.03.010

University of Zilina, Slovak Republic, Tokarcikova, E., Malichova, E., University of Zilina, Slovak Republic, Kucharcǐkova, A., University of Zilina, Slovak Republic, Durisova, M., & University of Zilina, Slovak Republic. (2020). Importance of Technical and Business Skills for Future IT Professionals. *Www.Amfiteatrueconomic.Ro*, *22*(54), 567. https://doi.org/10.24818/EA/2020/54/567

Walker, R. S., & Brown, I. (2019). Big data analytics adoption: A case study in a large South African telecommunications organisation. *SA Journal of Information Management*, *21*(1). https://doi.org/10.4102/sajim.v21i1.1079

Whitman, M., & Woszczynski, A. (Eds.). (2004). *The Handbook of Information Systems Research:* IGI Global. https://doi.org/10.4018/978-1-59140-144-5

Widyaningrum, D. T. (2016). Using Big Data in Learning Organizations. *Proceedings of the 3rd International Seminar and Conference on Learning Organization (Isclo-15)*. 3rd International Seminar and Conference on Learning Organization, Yoyakarta, Indonesia. https://doi.org/10.2991/isclo-15.2016.53

Appendix A: Participation Information Sheet

Dear Sir / Madam

My name is Reece Johnson I am a Masters student in Business Administration at the University of the Witwatersrand, Johannesburg. My supervisor is Dr Euphemia Godspower Akpomiemie. I am conducting a research study about Big Data Analytics in the financial sector. The study title is **Bridging Big Data Analytics Skills Gap in the Financial Services Sector**.

I am inviting you to take part in a once-off interview which will be recorded to respond to questions relating to the topic of Big Data. If you decide to take part, your participation in this research study will last about one hour The interview will take place at your office at a mutually agreed time.

With your permission, I would like to audio record the interview to ensure the accuracy of the report. This data will be stored on my computer for 1 year and deleted after 1 year. Only the researcher will have access to the data.

During the research activity, I will need to ask for some personal information about you, including organization, area of specialization, role in organisation and years of experience.

The interview will be confidential and anonymous. When I share the results of the research study, I will not include your name or anything else that could identify you.

If you decide to take part in the research study, it should be because you want to volunteer. You do not have to take part. You can stop being in the study at any time. You do not have to answer any questions if you do not want to. You will not get any direct benefits if you choose to join the research study. You will not lose any services, benefits or rights you would normally have if you decide not to join. Taking part in the research study will not cost you anything. You will not be paid for being in this research study.

The risks for this research study are minimal with unintended consequences, and conflicts of interest have been adequately mitigated.

This research study will be written up as a research report. If you would like to receive a summary of this report, I will be happy to send it to you.

If you have any questions during or afterwards about this research study, feel free to contact me at the details listed below. If you have any concerns or complaints about the ethical procedures of this research study, you are welcome to contact the University Human Research Ethics Committee (Non-Medical), telephone +27(0) 11 717 1408, email hrecnon-medical@wits.ac.za.

Yours sincerely,
Reece Johnson

Researcher: Reece Johnson, 2414115@students.wits.ac.za, +27 82 8786394
Supervisor: Dr Euphemia Godspower Akpomiemie, euphemia.godspowerakpomiemie@wits.ac.za

Appendix B: Consent Form

Bridging Big Data Analytics skills gap in the financial services sector

Reece Johnson

I, …………………………………….., agree to participate in this research project.

I agree to the following:

(Please circle the relevant options below)

The research study was explained to me. I understand what this study is about and what is required of me	YES	NO
I understand that I can volunteer to take part in the study	YES	NO
I agree that the interview activity may be audio recorded	YES	NO
I agree that direct quotations from my interview may be used by the researcher in their research paper	YES	NO
I agree that my participation will remain anonymous (my name or other identifying data will not be used by the researcher in their research paper	YES	NO
I agree that other researchers may use the information I provide in my interview (depending on their own ethics clearance being obtained) but my name and any personal information will not be used or passed on	YES	NO
I confirm that I am authorized and take responsibility for the proprietary information shared during this interview.	YES	NO

(Signature)……………………………………
(Name of participant) ……………………………………
(Date)……………………………………

R.Johnson

Reece Johnson …………………
(Date

Appendix C: Interview Guide

Organization:	
Area of specialization	
Years of experience	
Role in current place of employment	
Date of interview	
Time of interview	
Q1: What are some of the challenges you are experiencing when recruiting graduates for Big data roles?	
Q2: What skills and competencies are required for a Big data analytics?	
Q3: Which of the skills and competencies are minimum requirements for the role?	
Q4: How are the required skills obtained by employees in your business?	
Q5: Are these skills sufficient to enable employees to adapt to the banks' changing needs in this big data analytics era.?	
Q6: How is Big data analytics used in the banking industry and how do you see it evolving in the next few years?	

Q7: To what extent are the methodologies employed for the processing of Big data considered adequate to leverage BDA in the banking industry? Please motivate your answer.
Q8: Is your current skillsets able to support the technological advances in your sector now and in the future? Please elaborate?
Q9: Do you think there is a skills gap in Big data analytics? Please elaborate?
Q10: Do you think institutions of higher learning offer a comprehensive curriculum on Big data analytics?
Q11: Do you think institutions of higher learning produce graduates that are ready for the bank' current environment or industry needs?Please elaborate?
Q12: How can institutions of higher learning assist the banking industry with closing the skills gap?
Q13: What can banks do to close this skills gap?

Appendix D: Ethics Clearance Letter

Graduate School of Business Administration
University of the Witwatersrand, Johannesburg

Wits Business School Ethics Committee
Constituted under the University Human Research Ethics Committee (Non-Medical)

Ethics Clearance Certificate

Ethics protocol number: WBS/BA2414115/989

This certificate is only valid with a legitimate ethics protocol number and signed by the Researcher (below).
This certificate is only valid if permission has been granted by the Registrar's Office of Wits University.

Project title	Bridging big data analytics skills gap in the financial services sector
Investigator / Researcher	Mr Reece Johnson
Nature of Project	MBA (Research Article)
Decision of the Committee	Approved, provided stakeholders and participants are guaranteed confidentialty.
Issue Date of Certificate	9/11/2023
Expiry date	Date of submission of the project / research report
Chairperson	Dr Pius Oba ☎ +27 11 717 3976 +27 82 733 6587 pius.oba@wits.ac.za

Declaration by Researcher

One copy must be signed by the Researcher and returned to the Chairperson of the Wits Business School Ethics Committee.

I fully understand the conditions under which I am authorized to carry out the abovementioned research and I guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I undertake to resubmit the protocol to the Committee.

R Johnson
Signature

13th September 2023
Date: