Investigating how SMEs can benefit from Big Data Analytics

SL Makhele

cid.org 0000-0003-4798-5364

Mini-dissertation submitted in partial fulfilment of the requirements for the degree *Master of Business Administration* at the North-West University

Supervisor: Mr. JC Coetzee

Graduation May 2018

Student number: 24710350

ABSTRACT

Globally the Small and medium enterprises (SMEs) are the main drivers of the economic growth and the big data analytics has been seen as a support solution to help the SMEs to grow and become competitive. The primary goal of this study is to focus on growth, efficiency and effectiveness of SMEs using big data analytics. It also stresses how the big data analytics can add value to the SMEs by using available information to do the meaningful and well-informed business decisions (Ogbuokiri et al., 2015).

Coleman et al. (2015) claim that the big data analytics is currently the buzzword in many organizations globally; it can add value in products and services development and improvements, in customer relations, in profitability and in the creation of competitive advantage using available data and information. They also claim that according to the research, many SMEs are not yet using big data analytics and they are therefore behind regarding the big data analytics benefits (Coleman et al., 2015).

In this study, the challenges that SMEs face in the adoption of the big data analytics and how to overcome them will be discussed. There will also be some recommendations in order to assist SMEs in how they can manage their available resources to be able to adapt the big data analytics.

Keywords: SMEs, Big data analytics, IT, Data Warehousing, Business Intelligence, Cloud, MI, Open Source, Internet, Computing.

ACKNOWLEDGEMENTS

I would not be able to achieve this research without the various assistance and encouragements that I received from friends and family, especially my wife Ditlhare Makhele and my daughter Bokamoso Makhele; they have been a pillar of strength for me. They continuously understood when I was spending the weekends away from them. I am very appreciative of them.

I want to express the great gratitude that I owe to Mr. JC Coetzee, my supervisor for his knowledge sharing. He assisted me wholeheartedly and tirelessly. His constructive criticisms and suggestions made this research a successful completion.

Finally, I wish to show special appreciation to Mr. John Mukomberanwa. He is a statistician who assisted me with the data analysis. His hard work, friendliness and warm welcomes that I always got when visiting his place are not gone unnoticed.

Everyone who assisted in various ways to make this project a success, I thank you as well.

TABLE OF CONTENTS

ABSTRACT ii

ACKNOWLEDGEMENTS iii

LIST OF FIGURES vii

LIST OF TABLES vii

LIST OF ABBREVIATIONS viii

CHAPTER ONE 1

1.1 ORIENTATION AND PROBLEM STATEMENT 1

1.2 INTRODUCTION 1

1.2.1 Small and Medium Enterprises 4

1.3 CONTEXT 5

1.4 CAUSAL FACTORS 6

1.5 IMPORTANCE OF THIS STUDY 7

1.6 PROBLEM STATEMENT 8

1.7 RESEARCH OBJECTIVES 9

1.7.1 Primary objectives 9

1.7.2 Secondary objectives 9

1.8 THE HYPOTHESIS TO BE TESTED 10

1.9 RESEARCH METHODOLOGY 10

1.9.1 Literature and theoretical review 10

1.9.2 Empirical research 10

1.9.3 Limitations 11

1.10 LAYOUT OF THE STUDY 11

CHAPTER 1: Orientation and problem statement 11

CHAPTER 2: Literature review 11

CHAPTER 3: Empirical study 11

CHAPTER 4: Conclusions and Recommendations 12

1.11 CONCLUSION 12

1.12 CHAPTER SUMMARY 12

CHAPTER TWO 13

2.1 INTRODUCTION 13

2.2 DEFINITION OF BIG DATA 13
2.2.1 Big data benefits 14
2.2.2 Big data challenges 16
2.2.3 Data security 16
2.2.4 Data ethics and governance 17
2.2.5 Open source 17
2.2.6 Digitalization and e-commerce 19
2.2.7 Cloud computing 20
2.2.8 Cloud computing benefits and the challenges 22
2.2.9 SMEs of the future 23
2.2.10 Analytics and digital marketing 24
2.2.11 B2B big data analytics benefits for SMEs 25
2.2.12 Managerial implication 25
2.3 CONCLUSION 26
2.4 SUMMARY 27
CHAPTER THREE 28
3.1 RESEARCH METHODOLOGY AND DESIGN 28
3.2 INTRODUCTION 28
3.3 PURPOSE OF THE STUDY 28
3.4 RESEARCH DESIGN 29
3.5 POPULATION AND SAMPLE 29
3.5.1 Population 29
3.5.2 Sample size and sampling method 30
3.5.3 Non-probability sampling 30
3.5.4 Assumptions 30
3.6 DATA-COLLECTION INSTRUMENT 31
3.7 STATISTICAL ANALYSIS 31
3.7.1 Simple Statistics 32
3.7.2 Key highlights 35
3.8 DESCRIPTIVES 35
3.8.1 A socio-demographic profile of repondents: 36
3.9 SURVEY RELIABILITY 39

3.9.1 What are the perceived barriers to adopting Big Data technologies: Univariate Analysis. 43
3.9.2 Hypothesis testing 44
3.10 CONCLUSION 55
3.11 SUMMARY 56
CHAPTER FOUR 57
4.1 INTRODUCTION 57
4.1.1 Research finings 57
4.2 RECOMMENDATIONS 58
4.2.1 How to overcome big data adoption challenges 58
4.2.2 How to deploy big data technologies 58
4.3 MANAGERIAL IMPLICATION 60
4.4 REFERENCES 61

LIST OF FIGURES

Figure 2.1: Benefits of big data analytics. 14

Figure 2.2: Why big data analytics is important. 15

Figure 2.3: Cloud computing network graphical presentation 21

Figure 3.1: Type of SMEs. 36

Figure 3.2: Age of business. 36

Figure 3.3: Business model. 37

Figure 3.3 shows the type of business model adopted by the SMEs. The physical stores are on 44% and online stores shows only 17%. 37

Figure 3.4: Primary business location 37

Figure 3.5: Annual turnover 38

Figure 3.5 depicts the annual turn over of the SMEs, 40% of which is below R100 000. .38

Figure 3.6: Number of employees 38

Figure 3.6 shows that 70% of the SMEs employ less than five people. 39

LIST OF TABLES

Table 1.1 Big Data components 1

Table 3.1: Research questions 32

Table 3.2: Cronbach's alpha test on questions 39

Table 3.3: Factors/Theme variances 45

Table 3.4: Factor 1 - level of awareness 46

Table 3.5: Factor 2 - Leveraging of existing technology 47

Table 3.6: Factor 3 - Belief in big data 47

Table 3.7: Factor 4 - Perceptions of adopting open-source technologies 48

Table 3.8: Factor 5 - Implementing and adopting 49

Table 3.9: Factor 6 - Perception to adoption 49

Table 3.10: Factor 7 - Future growth perception 49

Table 3.11: Factor 8 - Leverage of data 50

Table 3.12: Factor 9 - Accessibility and Storage 50

Table 3.13: Factor 10 - Challenges 51

Table 3.14: Summary Table 51

Table 3.15: Factors/Themes 54

LIST OF ABBREVIATIONS

(BDA): Big Data Analytics

(**SMEs):** Small and Medium Enterprises

(IT): Information Technology

(DW): Data Warehouse

(BI): Business Intelligence

(MI): Management of Information

(GDP): Growth Domestic Products

(R&D): Research and Development

(VAT): Value Added Tax

(**ICT):** Information and Communication Technology

(B2B): Business to Business

(B2C): Business to Customer

(IoT): Internet of Things

(GPS): Global Positioning System

(SaaS): Software as a service

(SA): South Africa

(ZB): Zettabyte

(MIS): Information Management Systems

CHAPTER ONE

1.1 ORIENTATION AND PROBLEM STATEMENT

1.2 INTRODUCTION

The 'Big Data' is a combination of vast and complicated amounts of data (sometimes called the volume, variety and velocity problem). The Big Data information increases exponentially and therefore becomes difficult or impossible to work with using standard database management systems or analytical solutions (Davenport, 2013:4).

Scholz (2016:14) unpacked the Big Data in the following eight components:

Table 1.1 Big Data components

Components	Definitions
Volume	Various digital channels such as internet, social media, online shopping sites and Google analytics increase the amount of information available every second.
Variety	The available data is in different kinds of structures. It is unstructured, semi structured and structured.
Velocity	This is a rate of speed at which that the data can be produced, processed and consumed.
Veracity	The level of data trustworthiness. The level of trust that one can rely on when using the data.
Variability	The data comes in different forms and from different sources and is produced in different periods.
Complexity	The more the amount of data increases, the more difficult it is to manage and to use. The amount of information increases every second and, therefore, it ends up being a

	challenge to control and manage correctly.
Value	This refers to the benefit that one receives from collecting and using the data. Proper usage of the information must add a value one way or the other.
Viability	This refers to cost effectiveness of using and investing in the data management. The befit should exceed the cost of data collection and management.

Source: Scholz (2016:14) unpacked the Big Data in the following eight components.

Zicari (2015) maintains that Big Data is not a clear-cut term; but rather is a classification of the ever-increasing cumulative amount of unstructured, semi-structured and structured data. It consists of massive datasets that are increasing exponentially every second and that are too enormous and raw for processing using traditional data processing tools such as MIS. The datasets have grown beyond the capability of normal database software tools to capture, process, analyze, store, control, manage and turn them into useful business decisions.

In simple terms, the Big Data is the process whereby organizations accumulated so much information from both internal and external sources, and now it is difficult for them to process and make right business decisions from that data, especially using traditional data processing applications (Wedgwood, 2014:2). SAS (2016) also, explains Big Data as an exponentially growing collection of both structured and unstructured data. It increases on a daily basis in zettabytes (ZB) which is the highest number of the unit byte for digital information. The Big Data assists on insight analysis and helps in improving companies' strategic decision-making process.

The considerable increase in the use of the social media, web and other disruptive technologies, has created a massive growth in data recently. These enormous amounts of information are now captured, stored, processed, analyzed and shared on a daily basis by many organizations and this results in what is known as the Big Data. Nowadays data is regarded as one of the most important assets for any agency and should be fully utilized by organizations in order to fully benefit from it (Abdulla, 2014:7).

In many organizations, the problem is that new data can take two to three days before it is analyzed and by the time it is ready to be used, it is already out-of-date. A real-time analytical tool in needed when interpreting Big Data; because the current Information Management Systems (MIS) are not as efficient as they should be (Oracle, 2014:26).

Alsooj (2013:38) claims that the utilization of analytics that are entirely driven by the collected data within an organization, mainly helps the businesses to evaluate their performances; and for making well-informed business decisions that are useful and beneficial for all stakeholders. The collection and application of such data and related analytics have led to the development of the concept of Big Data analytics.

Many researchers, such Benda (2016:311-312); Zang & Ye (2015:42) have a universal agreement that Big Data has four components which are named the V features, these are volume, variety, velocity, and value. The V features comprise an extensive collection of data, exponential information growth and more convincing results.

Volume: The rapidly increasing information on internet popularity, social media, cloud computing and wireless technologies have resulted in massive amounts of data.

Variety: Data complexity is another feature of Big Data, nowadays data is not only structured, but it is also unstructured and semi-structured.

Velocity: Due to the fast-growing data volumes, the extremely fast processing tools are also needed to process and make sense out of this enormous amount of information in the shortest space of time, or in real-time if possible. If not, the data will lose its value. Real-time response is another feature of the Big Data.

Value: Most of the unstructured data does not have valuable information that is needed to make business decisions; first it has to be cleaned and adequately manipulated before it can be used. Therefore, raw value data is one of the leading features of the Big Data analytics.

The four primary dimensions mentioned above are used to better understand Big Data (Suwaidi, 2014:2, Bateman, 2015:2 and Davenport, 2013:3).

The purpose of the strategic drivers for Big Data applications is to provide the organizations with the quality of the data, to enhance the operations of the businesses, to make better business decisions, to avoid disruptions and also to achieve a competitive edge on other competitors. Below are some of the benefits of adopting Big Data:

- Risks detection and management

- Failure or incident forecast
- Threat analysis
- Non-compliance detection
- Diagnostic and forensic analysis
- Customers' Behavior analysis
- Suppliers, sales, and customer profiling
- Perception and feeling monitoring
- Geographic, demographic and location tracking
- Product and service improvement
- Data mining and aggregation
- Network monitoring and continuity analytics (Henke, 2016:45)

The creation of new business opportunities that can uplift the consumer can be easily achieved using Big Data analytics. One example is the Uber business model; it has developed innovative and improved services in the taxi industry that benefit the consumers, while creating a competitive advantage for the company at the same time (Wallsten, 2015:2-3).

1.2.1 Small and Medium Enterprises

The Small Business Development Act (102 of 1996) explains a small business as a separate and different business body in any sector; it may include nonprofit making organizations, privately owned by one or more owners and formal or informal.

Mahembe (2011:30) interpreted that Small and Medium Enterprises can be any enterprise that employs fewer than 200 employees and/or has an annual turnover of less than R64 million in South Africa.

The description of the SMEs sector covers a wide range of enterprises, formally registered, informal and non-VAT registered organizations. The SME industry varies from medium-sized companies, such as casual micro-enterprises, self-employed vendors, to family businesses employing from one to over a hundred employees. The higher end of the range is what is known as the small- and medium-sized enterprises (SME) segment (SEDA, 2016:6).

Many SME organizations are still using Information Management (IM) reporting tools, which is a system through which businesses are trying to increase their efficiency in the usage of their historical data in order to make better business decisions. A problem with the IM system is the fact that it uses outdated information, which is not in real-time and,

therefore, leads to obsolete decision making. It also leads the organizations to lose their competitive advantage and to eventually close down. Thus, Big Data is just as relevant in SMEs as it is to more prominent organizations (Oracle, 2014:26).

Oracle (2014:26) stresses that nowadays advanced tools such as Big Data analytics are making it possible for any type and size of organization to create its competitive advantages such as real-time reporting, customer Behavior predictions and innovation.

Big Data analytics enables SMEs to have advanced marketing analytical capabilities and empowers them to analyze and compare suppliers, customer trends, material costs, product quality and prices with those of competitors (Heyningen, 2012:28).

1.3 CONTEXT

Many SMEs have failed to take advantage of the Big Data analytics and to benefit from its capabilities because it used to be expensive. With the development of open-sources applications such as MongoDB, Hadoop, and Spark, servers and also data centres accessible through the cloud vendors, the cost of using Big Data analytics is now affordable. Many SMEs can now afford to adopt the Big Data analytics technologies, especially the cost effective open-source techniques. The well-advanced analytics and data integrity can provide any organization in any sector with many business opportunities (Ogbuokiri, Udanor & Agu, 2015:1-2).

Big Data analytics benefits include competitive advantages: data-driven business decisions, innovation: unique services and product development: cross-selling; customer relations and retention: organization's performance evaluation: fraud detection and prevention. All these benefits assist enterprises to become sustainable and to grow beyond their territories (Ogbuokiri *et al.,* 2015:1-2).

The SMEs are the backbone of a country's economy, and it is essential to assist them in how they can grow and create efficient, effective and sustainable enterprises. Adoption and implementation of advanced technologies such as the Big Data analytics is one of the critical points that can assist the SMEs develop (Ogbuokiri *et al.,* 2015:1-2).

The value of SMEs is acknowledged globally, regardless of the country's economic size and stage. The contribution that this sector makes towards economic growth, job creation and social advancement is highly valued and are, therefore, their adoption of Big Data analytics is considered the most critical tool that can assist in achieving economic growth and a solution to the unemployment challenges (Mahembe, 2011:13).

Regardless of the importance of SMEs and the contribution they make to economic growth; they face many challenges that hamper entrepreneurial growth. According to the SBP (2014:1), inadequate funding and access to finance, poor management skills, lack of adequate training and education are the main challenges that the SMEs are facing.

The SBP (2014:1) also suggests that all the governments should prioritize research and development. The primary goal should be to promote creativity and innovative ways of conducting business.

SMEs frequently expend most of their efforts on core business activities, instead of researching and investing in new technologies to improve their operations. As a result, many end up not being able to keep up with the technological changes and innovations (SBP, 2014:2).

1.4 CAUSAL FACTORS

According to Davenport (2013:12), the adoption rate of Big Data analytics by SMEs globally was only 5% in 2014 and thus SMEs will continue to lag behind the evolution of large companies (Coleman, Göb, Manco, Pievatolo, Martorelle & Reis, 2016:2).

The causal factors for this study were as follows:

- An Increasing awareness of the opportunities and benefits of adopting Big Data analytics. Previous researchers have argued that by efficiently taking and implementing the Big Data technologies, organizations will undoubtedly reap the reward of creating a competitive advantage and, therefore, sustainable enterprises (Obisesan, 2016:140 & Naru, 2016:5).
- Issues and possible solutions around Big Data analytics' implementation. The recommendations and potential solutions will be discussed throughout this study.
- The value of investing in the Big Data analytics. There needs to be a clear comparison of the costs of adopting Big Data versus the return on investment. This issue will also be discussed in this study.
- Resources management and alignment for the SMEs to enable this sector to adopt Big Data analytics.

Another crucial causal factor for this study is the fact that businesses who use customers' personal information should use this information ethically. Thus, implementation of the Big Data should always adhere to ethical procedures, such as storing and/or sharing personal information only with prior consent, allowing a customer to have the option of opting out of

such information usage, simplifying terms and conditions on documents and always being honest with customers' information (Tiell & Metcalf, 2016:4-14).

1.5 IMPORTANCE OF THIS STUDY

An efficient SME sector is the backbone of the economy of any nation (SEDA, 2016:6) by creating more job opportunities for the majority of the people, providing toward the GDP growth, encouraging entrepreneurial innovation skill and helping to increase the country's exports.

The SMEs are experiencing huge barriers when trying to adapt to the utilization of the Big Data analytics, challenges such as lack of skilled developers, data scientists and analysts, infrastructure setup and cost management. It is, therefore, complicated to forecast the cost implementation of introducing Big Data analytics into SME projects.

Should Big Data software be adequately implemented, it can give any organization (regardless of its size) a competitive advantage such as innovation, growth and profit margins, and it can, therefore, eventually contribute to economic growth (Rising, Kristensen & Hansen 2014:18-22).

Bughin, Livingstone & Marwaha (2012:48-49) stress that the SMEs that try to enhance their value and reap the strategic advantage of their participation in the digital marketplace, face the challenges of integrating various internal and external functional areas. Most SMEs with limited resources find it very challenging to adopt new advanced technologies, due to lack of both in-house skills and high outsourcing costs.

The improvements in technology, risk assessment and business models can progressively improve the SMEs' effort in becoming sustainable and competitive businesses. Nowadays digitalized markets are the primary enablers when it comes to the organizations' growth and can offer the SMEs tools to expand and enjoy the potential benefits linked to them. The entire value chain can be digitalized from the suppliers, retailers, customers and other business partners and, therefore, increase efficiency and profitability (Rausas, Manyika, Hazan, Bughin, Chui & Said, 2012:21-22).

According to Brown, Manyika & Chui (2012:30), the next-generation businesses that adopt Big Data analytics will be able to trail the Behavior of individual customers from internet click streams, store their purchase preferences and model their likely action in real time. They will be able to recognize when a customer is about to make a purchase decision and move the transaction to completion by pushing preferred products and making offers with reward program savings. These are some of the competitive advantages that businesses

that adopt Big Data analytics can enjoy, by being proactive and staying ahead of their competitors most of the time.

For all the reasons mentioned above the researcher feels that it is essential to conduct this study to find simplified processes for the adoption and implementation of the Big Data analytics by the SMEs.

1.6 PROBLEM STATEMENT

Many SMEs struggle to adopt and implement the Big Data analytics for their benefit, because of the wrong assumption that it is expensive and complicated. Challenging factors include lack of resources, tight budgets, data acquisition, storage, integration, cleaning, sharing, visualization, analytics and real-time reporting costs. The shortage of skills shortage is another obstacle, and it is also not easy to develop such skills overnight in-house (Rising *et al.,* 2014:18-22).

The assumption that these problems are insolvable is common but incorrect and there are some solutions such as open-source, cloud computing, social media and Google analytics that are not expensive and can be used to overcome challenges such as the cost factor.

Should Big Data analytics be adequately implemented, it comes with many opportunities and advantages such as more real-time reporting, competitive advantage, better business decision making, accurate performance data and creative ways of operating a business, all in real time (Heyningen, 2012:6).

There are new trends and other disruptive technologies such open source and social media, that allow companies of any size to participate in implementing the Big Data analytics for a small access fee. The SMEs can gain access to numerous open-source technology tools such as web analytics and cloud-based options to analyze data, for example Google Analytics. The SMEs also need to use the appropriate infrastructure to efficiently manage increasingly growing volumes of data and to ensure that their data security and privacy implementations are in order. The support system must satisfy the real-time information usage to realize its true value. Failure to implement it correctly will result in unnecessary costs, time wastage and other potential risks (Ogbuokiri, *et al.,* 2015:2).

According to Simms (2015), an effort is being made by companies to assist the SMEs to tackle the issues mentioned above. Topics such as skills shortage and costs versus benefits around Big Data analytics implementation will be discussed in this study together

with recommendations as to how the existing resources can be managed and aligned in this sector to adapt to application of Big Data analytics.

1.7 RESEARCH OBJECTIVES

The research objectives of the study comprise both primary and secondary objectives:

1.7.1 Primary objectives

The primary purpose of this research is to raise the SMEs' awareness and understanding of Big Data technologies, to investigate the influence of the Big Data on SMEs, as well its capability to transform them. The employment of simple methods to access, adopt and implement this system will also be examined. Lastly an investigation into the challenges and entry barriers, as well as the benefits of Big Data analytics will be conducted.

- The perceived barriers to SMEs adopting and implementing Big Data technologies.
- The factors associated with the adoption and implementation of Big Data technologies.
- The influence of the various factors on the level of turnover within the SMEs.

1.7.2 Secondary objectives

To achieve the primary aim of this study, the secondary goals to be realized are:

- Introducing methods to improve or leverage the existing resources.
- Introducing methods to assist the SMEs to set up their technological infrastructure in order to derive value from Big Data's capabilities.
- Investigating the challenges associated with Big Data analytics, such as skills shortage and costs versus benefits.
- Making recommendations for solving the problems related to the adoption and implementation of Big Data analytics.
- Creating rewards for formulating a managerial framework.
- Investigating awareness with regard to Open Source technologies
- Investigating how the SMEs can leverage Big Data ecosystems for innovation and creativity as well as costs versus benefits.
- Raising awareness on how the SMEs can create sustainable enterprises through Big Data analytics.
- Raising awareness around the ethics of data collection and handling.
- Discussing security risks and possible solutions

This research study was conducted to encourage and to make it easy for the SME sector to adapt to the Big Data analytics. According to many researchers, there are countless benefits, such as better business decision-making and increased revenue, from using this technology (Ogbuokiri *et al.,* 2015:1–2) & (Rising *et al.,* 2014:18-22)

1.8 THE HYPOTHESIS TO BE TESTED

The purpose of this analysis is to test the following hypotheses in order to realise the objectives of the study:

Hypothesis 1:

The perceived barriers to SME's adopting and implementing BIG data technologies.

Hypothesis 2:

The factors are associated with adoption and implementing BIG Data technologies.

Hypothesis 3:

The influence of the various factors on the level of turnover within the SMEs'.

1.9 RESEARCH METHODOLOGY

1.9.1 Literature and theoretical review

A literature and theoretical survey of the areas of data analytics, information management and disruptive technologies (such as cloud computing, wireless and social media) and opportunities and challenges in the SME sector will be conducted in Chapter 2. Specific focus will be given to the Big Data analytics for SMEs together with the benefits and problems of its implementation as well as recommendations for resolving these problems.

1.9.2 Empirical research

In order accomplish the objective of this study, empirical research will be conducted, the sample size will be selected from the SMEs in different businesses and industries that engage in business with one of the major financial institutions in South Africa. This study aims to reach a sample of 50 respondents, which will be formed by a group of very knowledgeable SME owners. They will, therefore, be the best sample to provide the most relevant and useful information for this study. Their level of working experience will range from junior to executive level.

They were asked to answer a 8-minute questionnaire, comprising questions related to their businesses demographics, understanding of data, their experience of Big Data analytics,

processes, technological terminologies and IT infrastructures, data management and the benefits and challenges.

The questions will include closed-ended questions to acquire short answers that can be used to measure the results of the study. The quantitative research methodology will be used in this study to gain an understanding of the respondents underlying reasons, opinions and motivations with regard to Big Data analytics.

1.9.3 Limitations

1.9.3.1 Sources

The publications currently available in South Africa do not have sufficient literature on the topic on Big Data analytics. The majority of available sources reflect the American or European point of view, and do not consider the data management challenges that African SMEs might be dealing with.

1.9.3.2 Research

This study will attempt to explain how the SMEs can benefit from Big Data analytics by managing and aligning their existing resources, and will be limited to information provided by:

- The SMEs and their Information Technology departments;
- The Information Technology professionals who will participate in this study.

1.10 LAYOUT OF THE STUDY

This mini-dissertation is divided into four chapters, which will be presented as follows:

CHAPTER 1: Orientation and problem statement

This chapter discusses the background, context of and causal factors of the study as well as the problem statement. It also presents an overview of the research design and layout of the remaining chapters.

CHAPTER 2: Literature review

This chapter investigates the nature of both SMEs and Big Data analytics. It also examines how the SMEs can manage and align their available resources to adopt Big Data analytics so that it can create a sustainable competitive advantage. This investigation will be conducted through a literature review.

CHAPTER 3: Empirical study

This chapter presents the research methodology and discusses the sampling data and methods used in the study, how the survey instrument (a questionnaire) will be compiled, as well as the study participants and the data collection. The data analysis and the results of the investigation will also be presented and discussed.

CHAPTER 4: Conclusions and Recommendations

The conclusions of the study, based on the literature review and the empirical investigation, as well as recommendations for further research, are presented in this final chapter.

1.11 CONCLUSION

According to the literature, Big Data analytics can bring so many opportunities for SMEs. If implemented correctly, the Big Data can increase the enterprise's competitive advantages and, therefore, result in sustainable business. With right data strategies, the SMEs can grow from strength to strength. The SMEs managers need to fully understand the threats and opportunities of the Big Data analytics, as well as the solutions to the challenges in order to succeed in implementing it fully. Knowledge of the data governance and the standard ethics is also essential.

The SMEs need to invest in research and skills development which are the only ways of finding the creative and innovative initiatives that can assist the businesses to grow and become sustainable.

1.12 CHAPTER SUMMARY

The basic purpose of this study is to investigate how the SMEs can manage and align their existing resources to adopt and implement the Big Data analytics so that they can create a competitive advantage and grow their business as a result.

Based on the literature review, the relevance, importance and challenges of the Big Data analytics for SMEs were discussed and assessed in this study and also the proposals on issues related to the Big Data implementations will be put forward.

In the problem statement, it was mentioned that Big Data analytics tools are available on the open-source platforms. The implementation rate will be measured by a questionnaire in Chapter 3 and the responses will be interpreted. The literature review, will be used to determine best practices and what forms the basis for an effective Big Data analytics implementation which will gain value for the SMEs.

CHAPTER TWO

2.1 INTRODUCTION

Rising *et al.,* (2014:3) stresses that the SMEs are finding it difficult to adapt and take advantage of data-driven business decision-making; they are bewildered by the intricacies of Big Data analytics. Recently global economies have not performed well and, as a result, it is imperative for SMEs to fully understand their customers' needs and preferences in order to stay competitive. Customer loyalty is more critical than ever before due to the fierce competition in most business sectors.

Some of the main Big Data analytics challenges facing SMEs are: accessing the right data, managing it, making value out of it and meeting or exceeding customers' needs. The technological requirements are another challenge when trying to adopt Big Data analytics (Wegener & Sinha, 2013:1-2).

Big Data analytics is a tool available to assist companies to access useful insights such as the type of the customers, their Behavior, the competitors, the best suppliers and partners. Analyzing this information from many different sources can allow businesses to take advantage of various patterns and correlations in the data to advance their business operations (EY & Nimbus Ninety, 2015:9).

The various tools, object and methods that can be used by SMEs to adopt Big Data technologies will be investigated and discussed in this chapter. Security is always a primary concern in data management and this will also be addressed. Recommendations as to how SMEs can adopt Big Data without encountering security concerns will also be examined.

2.2 DEFINITION OF BIG DATA

In many organizations processing analytics takes more than a day due to a fact that data increases in zettabytes every second and by the time it is ready to be used, it is outdated. The real-time object and analytical tools such as Big Data analytics are needed to overcome this problem (Oracle, 2014:26).

Big Data is high-volume, high-velocity and high-variety information assets that demand low-cost, creative methods of data processing that enable improved insight, decision making and process automation (Gartner, 2015:1). In simple words, volume means the quantity or size of data available, velocity is the speed of data generation and variety implies that the data comes from different sources and is structured, unstructured & semi-structured (Scholz, 2016:14).

According to Khan, Yahoo, Baker, Hashem, Inayah, Ali, Alma, Shiraz & Gain, (2014:5), Big Data comes with different forms of applications, both open-source and commercial ones. Some of the open source applications are MongodbDB, Hadoop, and Spark Open-source which were developed in a collaborative public manner. Examples of the Commercial applications are Aster, Datameer and SAP HANA.

2.2.1 Big data benefits

Kessel (2014:2) states that the data is growing exponentially currently, it is estimated that in 2020 it will be more than 100 0000 Exabyte's. So any business that uses advanced technologies will be able to achieve greater insights on markets, customers and other needs.

Figure 2.1 shows how Big Data can assist in business decision making, especially with the amount of data that is available; it will require very little effort to understand the customers' needs, expectations and preferences better. Once their needs, expectations and choices are realized, it will be easy to customize and, therefore, improve their services (Kessel, 2014:1).

Figure 2.1: Benefits of big data analytics.

Source: Vision graphics, (2017)

New business products and service development are also two of the solutions that result from Big Data. Given the amount of information available based on customer transactions,

complaints and compliments, other lines of services or products that can add value to the business can be developed and, therefore, increase the SMEs revenue (Henke, 2016:45).

According to Davenport (2013:16-17), optimization of costs, returns on investment and improved revenues are some of the benefits from Big Data analytics. By implementing integrated financial values and revenue systems, the profit margins can improve considerably because businesses can predict possible future transactions and cancellations of sales and, therefore, save on costs related to such events upfront, as shown in Figure 2.2.

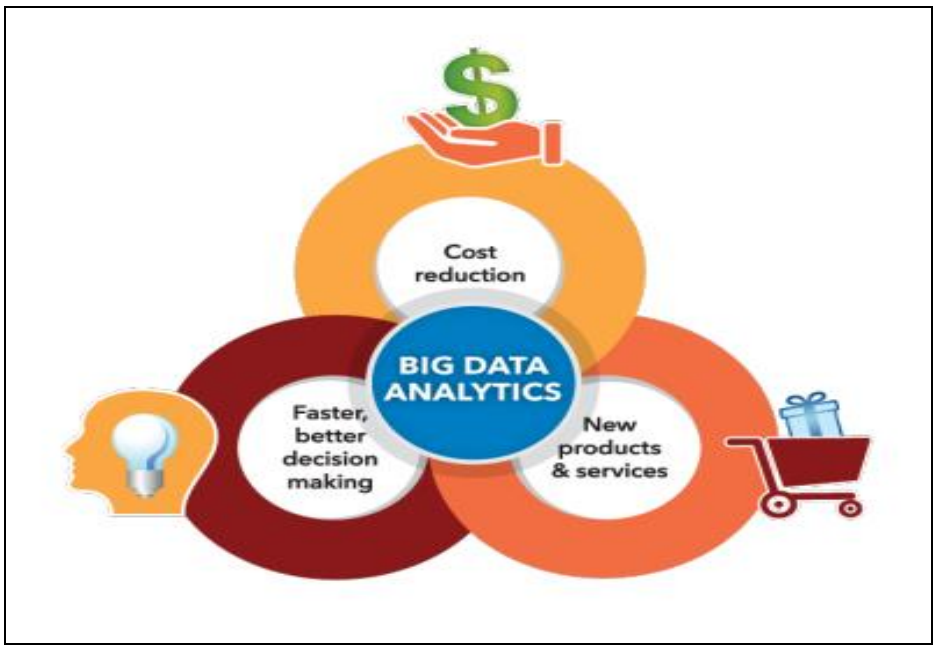


Figure 2.2: Why big data analytics is important.

Source: Adapted from SAS, (2017)

Big Data analytics provides real-time analysis opportunities and, therefore, can offer businesses efficient real-time data analysis and, thus, enhance their business decision making processes. The time required to produce, promote, market and sell a product can be reduced significantly with the help of Big Data (Kessel, 2014:1).

There are countless sources of the data such as social media, cellphones, email, web and internet and many different types of analytics can be conducted from these sources which will assist businesses to grow, based on the results achieved from the wide variety of data available (SAS, 2016).

2.2.2 Big data challenges

A wide range of problems exist with Big Data in general and involve issues such as costs, data security and privacy, storage and processing, analytical, technical and data value challenges.

According to Zicari, Rosselli, Ivanov, Korfiatis, Tolle, Niemann & Reichenbach (2016:18), it is complicated to set up a new Big Data infrastructure from scratch and to manage it. The scalability, which means adding more resources to the whole system as it grows, is also not easy. Another major problem is the shortage of IT experts, such as developers, data scientists and analysts with sufficient knowledge and skills to successfully implement a Big Data project.

Many organizations find it difficult to source suitable talent with capable skills to work with new technologies and to interpret the data in order to extract meaningful business insights. Data scientists are the IT professionals who work with data models, algorithms and visualizations to assist with data manipulation from the enormous amounts of data available. These organizations are also looking for professionals with the requisite business acumen so that they can use the data insights to satisfy their customers' needs (Heeley, 2014:80).

2.2.3 Data security

Kessel (2014:20) states that the new technologies have certain security threats, therefore, organizations should always safeguard their privacy by applying information governance and ensuring the protection of the data and information. Security is one of the significant challenges in data management, and especially for Big Data, given its complexity.

In the Big Data environment, information is stored on many different storage sites such as cloud-based storage, web systems, social media and other physical servers, therefore, making it very difficult to secure the data completely, thus bringing a considerable security risk to the whole environment. In one month only in 2013, cybercriminals hacked over 400 million clients' transactional information, this figure shows how important security is in the Big Data environment, consequently more time and budgets should be spent in security management than currently. A 2014 research study found that the cost of cybercrime was between US$375 billion and $US575 billion, or about 0.6 percent of GDP worldwide (Kim, 2016:223).

Disaster recovery is another factor that needs to be taken seriously. Should a server or the whole information storage be destroyed, SMEs should be able to use their backed-up data and continue with their business with as little inconvenience as possible (Lyon, 2015).

2.2.4 Data ethics and governance

Ethics are a crucial part of IT usage globally, especially when dealing with customers' data. Confidential customer information should not be used without the customer's prior consent, not misused in any way. Policies governing the ethical use of the customer information should be in place and always adhered to (Mittelstadt & Floridi, 2015).

(Kessel, 2014:10) stresses that good governance plays an important role in the success of Big Data initiatives in any organization; it incorporates reliable guidance processes and clear management decision-making. Therefore, SMEs must ensure they adhere to standard and comprehensive data capture procedures without compromising the security measures, they need to protect the data all the time throughout all the levels and functions within the organization.

2.2.5 Open source

Almeida & Bernardino (2016) define that open source Big Data platforms as a combination of IT hardware and software tools that are implemented to extract, store, process and analyze data. These devices are combined within a shared architectural and virtualized cloud environment. The data can be stored, processed and analyzed from this virtual cloud environment at a minimal cost.

Many organizations that do not have excellent financial muscle can then adopt open source tools to take advantage of the benefits of Big Data analytics. Cloud computing and open source software provides these companies with an affordable way to embrace Big Data analytics. Another way of accessing Big Data analytics is to use free public sources such as the internet, 'blogs' and social media platforms (Vătuiu, Udrică & Tarcă, 2013:404).

According to Liu, Man, Chong & Chan (2016:1), free web analytics such as Google analytics and social media are available to any organization at any given point and time, the only downside will be the prior implementation of skills development and training, which can be done easily. Web analytics can be used to collect certain information, for example tracking users' locations, what websites were used and when were they accessed. This information can be collected via cookies and sessions that are stored on the users' systems, and then used to discover trends in popular online items which can be

analyzed to establish customer Behavior. Antivirus and other data security measures will also need to be implemented in this process to minimize the risks of cyber-attacks.

SMEs can use these free web analytics to better understand and customize their customers' needs and preferences. These web analytics are typically performed using web interface using one machine and no specific hardware or software is needed. The only costs would be the training of staff, which should be regarded as an investment. In this way the benefits of Big Data analytics can be achieved at a minimal cost (Liu, *et al.,* 2016:1-2).

Open source data is provided by most search engine organizations such as Yahoo, Google and Wiki. These and other social media applications can be used to improve the business activities at a very low cost. Open source data can be used to analyze customer trends and, therefore, assist in developing appropriate products and services. It can also be used for predicting future needs and demands which can result in greater efficiency and increased revenues for the business, while simultaneously reducing costs (Vătuiu, *et al.,* 2013:404).

According to Bughin, Chui & Manyika (2013:8), the search and trends analysis data is available for free on many online platforms and can be used by companies to improve their businesses. This date assists businesses to track which items are searched and liked the most on social media so that they can start producing those products and services. This information can be used to analyze the company's competitive situation, as well as how it can improve its products and services. It can also be used establish which products or services are currently in demand and then promote and market them accordingly at market related prices. SMEs can likewise use this kind of information to enhance their business activities through the development, production, segmentation, marketing and promotion of new products and services.

SMEs can monitor the social media input to establish the public's perception of their products and services in order, to improve the value they offer and better satisfy customers. The negative comments from social media can be used as constructive criticisms. The number of 'likes' recorded on social media can be used to meet demands for specific products and services. If used correctly, this kind of information is beneficial to the businesses (Carter, 2014:9-12).

Carter (2014:9-12) also states that the cost of using social media is cheaper and more efficient than traditional technologies and, thus, social media can be a suitable and useful marketing tool for SMEs.

2.2.6 Digitalization and e-commerce

According to Dall'Omo (2017:4) digitalization means leveraging technologies and insights from information and data. The whole idea behind this practice is to ensure that the customer's needs are met efficiently. Examples of digital technologies such as cellular telephones, landline telephones, video teleconferencing, cloud computing, sensors, analytics and the Internet of Things (IoT) are fast-tracking the way business should be transacted and are used to collect, analyze, store and share information globally.

There are different low-cost digital marketing solutions that can be used by SMEs to expand and become sustainable, such as social media. More than 93% of the marketers use social media and about 75 % of social media users are more likely to buy something that they see promoted on social media than through traditional forms of advertising. Social media users are growing exponential on a daily basis (Dall'Omo, 2017:4).

The World Economic Forum (2016:10-11) states that the digital innovation is changing all the industries by disrupting current operational models. It has brought about entirely new business opportunities and challenges. It is fast-tracking the globalization concept, it is also assisting many businesses to become more competitive, not only in their territories but on a global level by increasing efficiency, reducing costs and growing revenue. Digitalization assists businesses to utilize customers' information more creatively and enhance their customer's experiences.

Big Data overpowers traditional limitations in a cost-effective manner and opens opportunities to consume, store and process data from new sources, such as external social media data, market data, communications, interaction with customers via digital channels (World Economic Forum, 2016:10-11).

Digitalization brings new business models that enable the business to attract an ecosystem of partnerships which help to ensure customer loyalty. It allows the company to partner with other stakeholders internally and externally, such as suppliers, intermediaries and offering ecosystems. The old legacy systems can be integrated into the whole digital setup as well, and this can result in creation of new products and services for the customers (World Economic Forum, 2017:13-20).

The World Economic Forum (2016:21) stress that digitalization provides customers with more efficient shopping options because it offers access to the global online markets and a greater variety of products and services to choose from, at lower costs. The e-commerce platforms are integrated and automated and, therefore, result in only a small fee being associated with transacting because the advertising fees on business websites are used to finance the operations.

The Big Data capabilities and techniques that have been used in physical stores can be applied even more in the e-commerce environment. Digitalization allows businesses to deliver a tailored-made experience to their customers; Nike is one example that offers customized products. Companies can access all the necessary information regarding their customers' buying habits which can enable them to create tailor-made products and services for individual customers resulting in both increased revenues for the business and satisfied customers. Biz-2-Me is a good example of business digitalization (Deloitte, 2017:12-14).

According to Brown, Schuler & Sikes (2012:10), the many highly digitalized organizations have implemented the concept of 'single customer view'; an approach that incorporates multi-channel sales and support components. This approach offers customers the option to conduct business via a variety of preferred channels, at a convenient time and place. It also allows companies to take advantage of each customer interaction. Most of these companies have moved away from the traditional business methods, such as a one-way website channel and have adopted social media and web 3.0, Enterprise 2 and Business 2 as tools to use in engaging and interacting with the customers. These companies, therefore, benefit from gaining insight into customers' trends, needs and Behavior, which can be used to improve their market share and revenues while satisfying their customer needs.

Brown *et al.,* (2012:11) also mentions that these companies are also creating more efficient and effective internal processes to improve their decision-making processes, which then benefit their customers through a shortened delivery period. They also apply these technologies to their core businesses in order to find the new business opportunities that create additional profits from new revenue lines.

2.2.7 Cloud computing

Cloud Computing technology is a graphic network representation of a cloud network. It allows the delivery of computing as a service instead of physical hardware. It also uses the

internet to access the virtual resources services that the user has rented. The user can manage and control these resources through the internet connection. This process also refers to the hiring of the virtual hardware and software resources from third party companies, thus cloud computing technology software will be installed on the server that is not owned by a client. The processing and storage capabilities are accessed online (Vătuiu *et al.,* 2013:398-399).

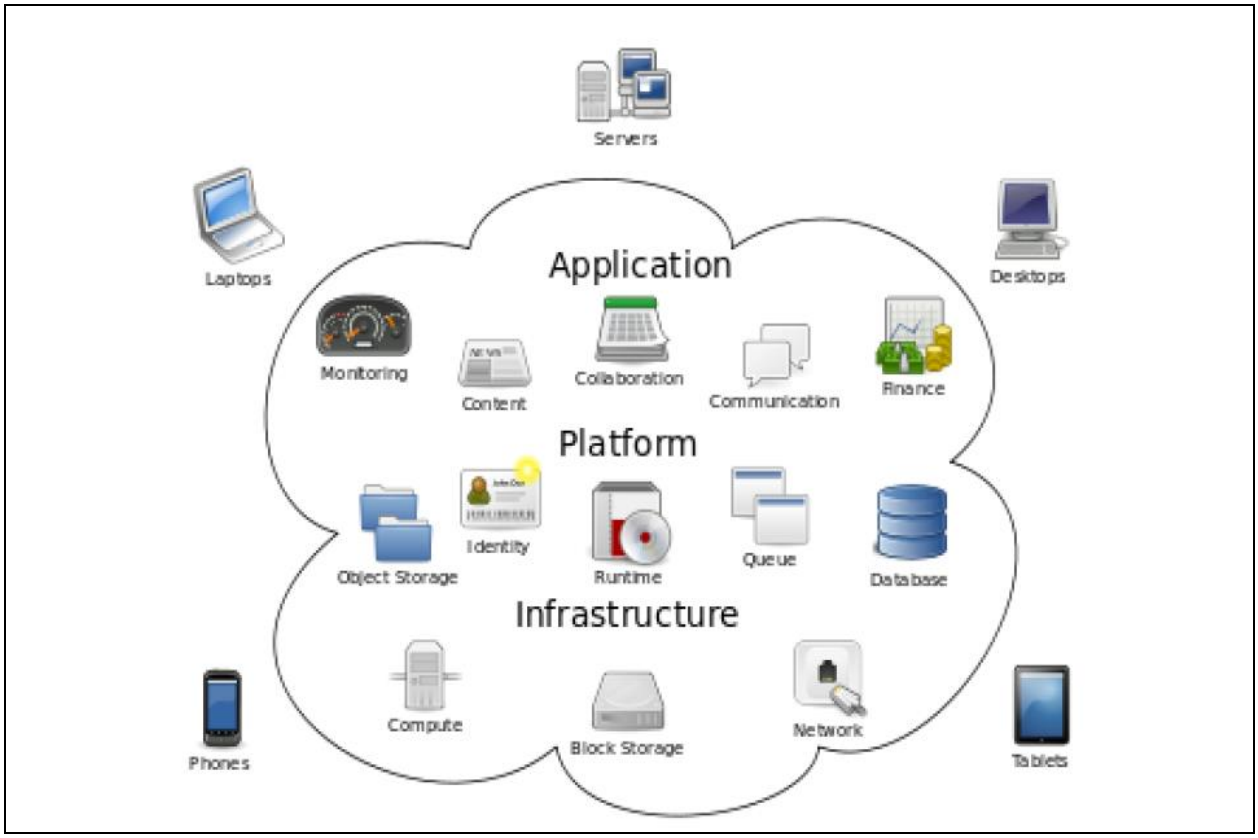


Figure 2.3: Cloud computing network graphical presentation

Source: Adapted from Vătuiu *et al.,* (2013:399).

Tole (2014:49-50) stresses that cloud computing solution implementation can result in a significant cost reduction and, thus, its use is recommended for SMEs. It also allows SMEs to implement Big Data solutions without really spending large sums on hardware.

According Vajjhala & Ramollari (2016:131), there are two primary cloud computing service models which are:

- Software-as-a-Service (SaaS), on this service, users can access and execute applications on the cloud infrastructure managed by the cloud supplier.
- Platform-as-a-Service (PaaS), PaaS permits users to deploy applications on the infrastructure provided and maintained by the cloud provider.

(Vajjhala & Ramollari, 2016:131) mention that private, community, public and hybrid clouds are the four main cloud computing deployment models. Private clouds allow the cloud infrastructure to be deployed entirely for use by a single organization. The primary purpose of private clouds is to maximize the level of control over data security and privacy. SMEs would typically not require a private cloud. Community clouds allow the cloud infrastructure to be used by a specific community of organizations or consumers sharing common concerns.

Public clouds are accessible to the general public and the infrastructure is also available to the public. In a public cloud environment, resources are organized on a self-service basis and service is provided through web services or other alternative forms.

Public clouds can be deployed off-site over the Internet and are available to the general public. The benefit of a public cloud is that it offers excellent efficiency and resources at low cost. This type of cloud infrastructure could be quite useful for SMEs which often have limitations with both budget and skills training. Public clouds could be beneficial for SMEs because the analytics and data management services are offered by the cloud service provider which is also responsible for the quality of service. Hybrid clouds are a combination of the other three cloud deployment models bound together by branded technology, enabling application convenience (Vajjhala & Ramollari, 2016:131).

2.2.8 Cloud computing benefits and the challenges

Cloud computing resources and procedures can be implemented to solve the problems and challenges associated with Big Data analytics. SMEs can take advantage of cloud computing methods to source the opportunities and benefits of Big Data, without making significant investment in technology and skills training (Vajjhala & Ramollari, 2016:130-131).

Cloud computing provides some benefits for businesses with low budgets, the organizations can run reports and dashboards at very little cost. It is not necessary to employee an IT professional or to own IT resources. The software on clouds is updated automatically and can be scaled quickly, free of charge. The only cost to users is an affordable monthly subscription usage fee, instead of an upfront licensing cost. Cloud computing fast tracks service delivery, decreases costs, improves the flexibility, reusability and scalability of data (Kaur, Azad & Singh, 2013:25).

Also, according to Vătuiu *et al.* (2013:399), cloud computing reduces capital and operating costs. There is no need to invest in physical data centres, hardware costs are limited and

customers only pay for the resources they require and use. Application deployment and management are simplified in cloud computing which offers access to the vital ecosystems as well. Cloud computing is an efficient way of managing information because it is supported by the continuous increase in broadband availability.

Agostini (2013:9-11) also stresses that cloud computing is valuable for SMEs because the technological obstacles have been removed and both IT beginners and experts can manage this method efficiently.

Security is wrongly assumed as a significant challenge when it comes to cloud computing. There is a wrong perception that cloud computing is less secure than physical data centres, but the reality is the opposite. Data is safer in clouds than in physical data centres, although it will always be important for SMEs to back up their data (Kaur *et. al.,* 2013:25).

2.2.9 SMEs of the future

Big Data can turn today's small business into tomorrow's big business. In today's world, regardless of whether a company is small or significant, it is crucial for it to keep up with the needs and preferences of its customers, so that these requirements can be satisfied within a reasonable space of time. In SMEs it is easy to be creative and innovative, given the level of flexibility in this sector, compared to the big companies. This situation also applies to the Big Data process which can be used to implement a new way of doing business (Wall, 2014).

Wallsten (2015:3-5) claims that two typical examples of successful businesses that began as SMEs are Uber and Mr. Delivery. Uber's business model entirely depends on the Big Data crowdsourcing principle, which has contributed enormously to the success of this business. Any person owning a car who is willing to join this market is allowed to do so. The smartphone application is used to assist passengers to book the taxis. It has a comprehensive database of drivers in different locations, and the locations, passengers and the drivers are matched within a few minutes using Big Data analytics to offer a service. The fares are calculated based on the distance travelled using the GPS with a Big Data informed pricing model; it also applies real-time reporting to check on traffic situations. To build a good relationship with passengers and to make well-informed business decisions, Uber uses rating systems and social media to get some feedback from the passengers regarding the service (Wallsten, 2015:4).

Wallsten (2015:4) also stresses that the success of SMEs depends on the best practices of information management and the use of data available to improve and grow their businesses.

2.2.10 Analytics and digital marketing

Digital marketing is the implementation of digital tools to make a comprehensive, targeted and measurable communication channels which can assist in customer acquisition and retention and create the lasting and long-term relationships with the customer base (Kottier, 2017:8).

Although digital marketing is a new trend globally there are many advantages to using it. Only about 25% of the marketing management teams used data-driven marketing methodologies in 2015. The business decision making can be improved greatly by adopting Big Data analytics which improves accuracy, time management, segmentation and provides real-time marketing. Big Data analytics assist companies to achieve about 75% of the marketing information and data; and SMEs should not lag behind in opportunities such as data-driven decision-making and data analytics for marketing purposes (Kottier, 2017:7).

In Big Data analytics, real-time analysis and reporting is a priority offering. For enterprises to access the real-time reporting and analysis, they need to connect to the mobile devices such as smartphones, tablets, and other mobile devices. With small investment on a mobile application, the companies can then reap huge rewards and become market leaders, as well as being ahead of the competition all the time (Intuit, 2016:6-7).

By gaining access to their customers' social media, SMEs will have a direct relationship with their customers and will be able to target, promote and market what their customers need and prefer at the right time and price. Benefits such as targeting potential customers, gaining access to competitors' information and profiles; tracking customers' Behaviors; and competitiveness are easily accessed with digital marketing. In addition, the fact that their competitors will also be using the social media will make SMEs work smarter than the competition. Access to social media is very affordable and can, therefore, be the best marketing tool for SMEs (Kottier, 2017:10).

Digital marketing analytics can assist SMEs to track new types of customer Behaviors. Many customers spend time 'surfing' the internet before buying products or services in order to compare the prices and quality of goods and services from different competitors, in order to buy the best quality goods at the lowest cost. This kind of available information

can assist marketers in targeting customers. Digital marketing enables SMEs to try a prototype campaign with little risk and limited financial budgets and, if all goes according to plan, the company can make a well-informed business decision on whether or not to implement these campaigns. It is also possible to measure the return on investment when using digital marketing (Europages, 2016:4).

2.2.11 B2B big data analytics benefits for SMEs

Big Data applications allow business to automate and integrate the whole supply chain. It enables enterprises to attain values such as being able to source the best suppliers and distributors based on better costs, with the quality of goods and services they offer and to track customers' Behavior, per industry. It also provides the opportunity to personalise customer needs through trends manipulation and to come up with the new products and services that never existed before. SMEs could take advantage of such data and information that is already available in B2B chain, apply and analysis it to realize value from it. The access to digital creativity can also assist the SMEs in improving their B2B payments methods which can also result in enhanced revenues, growth, costs reduction, improved customer experience and create a competitive advantage (Deloitte, 2017:19).

2.2.12 Managerial implication

Implementing Big Data analytics within organizations, especially SMEs, is a big challenge, but it is crucial for them to deal with Big Data. The research has identified a number of managerial implications for SME leaders. The adoption of Big Data analytics is a complicated exercise. The SMEs should be upskilled in their understanding of and dealing with Big Data analytics.

The SMEs leaders need to consider factors such as:

- Is data regarded as an asset that can assist in business decision making by their enterprises?
- Is there a belief that data can be used to improve and grow the business?
- Is there awareness around Big Data analytics and its benefits?
- How can barriers to adopting and implementing Big Data analytics be overcome?

If the answer is no to any of these questions, then it is suggested that the leaders of SMEs should consider engaging in a change management process. This practice will encourage the adoption and implementation of the Big Data technologies supported by every member of the SME structure.

Once this change has been achieved, then the next step for the SME leaders is to consider if they can use the data and the resources that they have already to adopt Big Data analytics. It has already been proven in the literature study that Big Data analytics adds so much value to a business enterprise that, consequently, it can add as great value to SMEs.

If management sees an opportunity for the use of Big Data analytics, then they need to up-skill or leverage the owned existing resources. They can start by up-skilling their internal staff through an open source course, Google analytics, cloud computing, social media, as well as the Big Data analytics and also raise awareness of how all these technologies can benefit SMEs. It is not affordable to employ Big Data experts because they are expensive due to their scarcity. Therefore, SME employees need to acquire Big Data literacy together with business skills.

Everyone within the SME has to support Big Data implementation project and the value derived from Big Data analytics.

2.3 CONCLUSION

Big Data analytics can help in offering businesses the necessary tools to grow and become competitive by transforming data into useful information that can assist enterprises to make well-informed decisions. Data is an asset in itself. Big Data analytics is relevant in SMEs as well, and that is where it is needed to assist this sector to grow and to become efficient and sustainable.

One way or the other, Big Data analytics is impacting many different types of organizations, in various sectors and diverse economies. There are some wrong perceptions about Big Data challenges and most of the assumed obstacles are solvable. There are many different approaches that can be followed to adopt Big Data analytics successfully. It has been mentioned that the cost-effective procedures such as open source; web analytics; cloud computing and social media are some of the approaches that can be used by SMEs to benefit from the Big Data analytics. SMEs should also invest in in-house skills training, to avoid paying high fees to Big Data experts and there are the different online courses available at very little cost. There are many different readings and tutorials which proves that there is guidance available for anyone wishing to make use of Big Data analytics. SMEs are more flexible than large businesses and can, therefore, adopt the Big Data approach quickly, because they do not have significant data available,

nor do they have the old legacy systems to integrate and so can effortlessly build up their database.

2.4 SUMMARY

Big Data is the process whereby organizations have accumulated so much information from both internal and external sources, and now find it difficult to process and make appropriate business decisions from it, especially using traditional data processing applications.

Many different kinds of problems exist with Big Data in general, issues such as high costs, lack of data security and privacy, storage and processing, analytical, technical and data value challenges.

All the benefits that have been mentioned above regarding the Big Data analytics can still apply to SMEs. The literature stresses that regardless of the size of a business, the amount of data that is available to the public, combined with other methodologies such as open source, social media, and cloud computing can be used to set up the Big Data analytics for SMEs. The adoption of Big Data analytics is not expensive.

Most SMEs do not understand the importance of Big Data analytics due to its complications and the fact that it is a combination of multiple technologies. By using the right combination and techniques, Big Data analytics can bring numerous benefits for SMEs. It is vital that SMEs ae not left behind regarding Big Data adoption, in today's data society, no company regardless of its size and nature, can ignore it (Macinnes, 2017).

SMEs can also leverage Big Data to create a strategic competitive advantage, but the lack of resources limits them. There are various cost-effective technologies and tools that can be used to assist SMEs in adopting Big Data analytics such as social media, open-sources, Google analytics and cloud computing.

CHAPTER THREE

3.1 RESEARCH METHODOLOGY AND DESIGN

3.2 INTRODUCTION

In Chapter 2, the literature study showed how vital the SME sector is. It also stresses how the how the new methods of conducting business should be adopted to ensure growth and sustainability for SMEs. Big Data analytics is recommended as a means of assisting this sector to improve and expand their businesses. Big Data and other free related accessible technologies are defined; their benefits, challenges and the threats around the adoption, management and implementation factors were also discussed in Chapter two.

Chapter 3 focuses on the research methodology that is utilized to assist in achieving the goals and objectives of this research study. The principal focus will be on the research design, sampling methodology, instruments used as well as data analysis. The questionnaire is divided into different themes, which are awareness around Big Data analytics, perception, challenges, benefits and what resources are needed to adopt Big Data analytics.

Questions are divided into themes that include the demographics, available resources in SMEs, perception regarding data and information management, awareness of technology and familiarity of analytical technologies. A questionnaire is attached as one of the Annexures.

3.3 PURPOSE OF THE STUDY

The main reasons for this research are to examine different ways of making SME owners aware of Big Data analytics, its challenges and benefits and how it can be accessed. It further investigates the way the adoption, implementation and management processes can be simplified. The existing resources leverage, skills shortage, costs, open source tools, data security as well as perceptions will also form a main part of the investigation and discussion.

This research is conducted to encourage the SMEs sector to adopt Big Data analytics. Lastly, suggestions will be made on how to overcome the challenges and destructive perception myths surrounding Big Data analytics in order to improve and expand the SME sector.

Different strategies will be suggested with regard to the various available advanced technologies that can be used to improve the service of SMEs.

3.4 RESEARCH DESIGN

Kotzé (2007:44) outlined a research design as the map and the process for research. It includes the detailed methods of data collection and analysis. It should address the type of research problem being examined.

In other words, the research design addresses critical factors such as the underlying assumptions, how the research is going to be conducted, specifically the data collection and analysis method that is going to be used. Lastly, it explains why a certain chosen specific plan will be used for this study (Kotzé, 2007:44).

Research Methodology outlines the strategies that will be applied to data collection, data processing, results presentation and interpretation (Kotzé, 2007:44).

3.4.1 Quantitative research

Creswell (2014:32) suggested that a quantitative research approach is appropriate to test theory objectives in order to understand the relationship between variables. This study adopted an explanatory research approach that sought to understand the influence of procurement rules and legislation on achieving sustainable development objectives.

According to Fraser Health Authority (2011), quantitative research is based on traditional scientific methods, which generate numerical data and usually seek to establish causal relationships (or association) between two or more variables, using statistical methods to test the strength and methods and significance of the relationships. It is a deductive process that tests ideas or hypotheses on numeric data to enable statistical analysis.

To undergo this quantitative research project, the researcher designed a questionnaire using SurveyMonkey application, which was distributed through emails and WhatsApp application to those respondents who did not have email addresses.

SurveyMonkey is a digital survey development cloud-based software service (SaaS) company. It offers free surveys; it also includes data analysis, sample selection, bias elimination and data representation.

3.5 POPULATION AND SAMPLE

3.5.1 Population

Alvi (2016:10) refers to population as all the members who meet a certain specified criterion intended for a research investigation.

The targeted population of this study consisted of any person who owns an SME in South Africa. The unit of study ranges from one individual owner to more than one owner.

3.5.2 Sample size and sampling method

Alvi (2016:11) also explains sampling as the process through which a sample is selected from a population because it is impossible to assess every single element of a population. Therefore, a sample is smaller in number than the population chosen for the assessment.

According to Levine *et al*., (2008:303), the sample size equation is depicted as below:

Equation 3.1: Sample size evaluation

$$n = \frac{Z^2 \pi\ (1 - \pi)}{e^2}$$

Where:

n = the sample size required for the given parameters

Z = the number of standard deviations for the given accuracy

π = the proportion of sample of interest (a value of 0.5 maximises the sample size, therefore minimising the error)

e = the error allowable, for instance, 10%

In this study SMEs owners throughout South Africa, who are banking with one of the major banks in South Africa, were invited, using an email invitation, to participate in the research study by filling in a questionnaire.

3.5.3 Non-probability sampling

Aoto *et al*. (2014:31) explain non-probability sampling as a recognized and powerful tool that affords significant time-saving in the acquisition of data by sampling only a percentage of the overall sample size.

In this study non-probability sampling was used, becuase this form of research does not require an excess of either time or resources.

3.5.4 Assumptions

The assumptions that were made for the research sample were as follows:

An understanding of English language and technology,

An understanding of IT terms.

The sample collected is a true representation of the population

3.6 DATA-COLLECTION INSTRUMENT

The questionnaire (see Annexure A) was developed in SurveyMonkey application targeting a convenient sample of the SME owners banking with one of the major banks in SA. The questionnaires were distributed to each responded through emails, and WhatsApp application was used to distribute the questionnaires to those who did not have email addresses.

The survey consisted of sixty questions, and of the 80 questionnaires sent out, only fifty-five participants responded which is a 69% response rate. The data collected was handed over to the Statistical Consultation Services (? From which University) for further analysis.

SurveyMonkey also undertakes sample selection, bias elimination; necessary data analysis and data representation.

On average it took the research participants eight minutes to respond to all sixty questions.

The questionnaire was split into two broad themes, which are the demographics and the general understanding of Big Data, it benefits and challenges. In this survey, the researcher did not include any pre-existing questions becauseand this topic is a relatively new one for SMEs.

With regards to the second section, the questions are based on a Likert scale of 1-5 and the rating scale was based on 1- 5, where 1 was “strongly disagree” and 5 was “strongly agree”. Researchers such as Rao, Solka & Wegman (2005) are supportive of the Likert scales due to their simplicity of interpreting contrast views. However, it is worth noting that the Likert scale has its limitations such as the inability to manipulate further detailed views on the subject rated by the respondents. The drop-down listed questions and the restrictions on the close-ended questions are also included in a questionnaire to obtain short answers that can be used to measure the results of the study.

3.7 STATISTICAL ANALYSIS

The purpose of this analysis was to test the following hypotheses:

Hypothesis 1:

The perceived barriers to SMEs adopting and implementing Big Data technologies.

Hypothesis 2:

The factors that are associated with adopting and implementing Big Data technologies.

Hypothesis 3:

The influence of the various factors on the level of turnover within the SMEs'

3.7.1 Simple Statistics

Table 3.1: Research questions

Question	Per cent Positive responses	Category
Q45. I believe my business can be competitive in future.	96%	
Q25. I understand the basics of data collection and management.	94%	
Q33. I believe I can use the data that I collect to create better client value propositions.	92%	
Q44. I believe that there are ecosystems that we can share and leverage of data in our industry.	92%	
Q34. I believe data analytics can create a value in my business.	90%	
Q46. I believe in digitalization which is taking place in my industry.	90%	
Q37. I often read about the topics on how technology can improve the business.	88%	
Q48. I believe I understand the needs and the preferences of my customers.	88%	
Q51. I believe information management and business intelligence can improve my business.	85%	
Q13. Types of IT resources you have in your business.	84%	Laptop
Q42. I believe that data usage can provide my	82%	

business with the competitive advantage.		
Q43. I believe that data usage can assist in growth of my business.	82%	
Q49. I believe I understand the future needs and preferences of customers .	80%	
Q23. Data is a key driver in my business decisions	74%	
Q11. Do you use any social media tool for your business?	72%	
Q30. I am aware of how I can leverage my data in my business decision making.	72%	
Q32. My competitors are actively adopting technologies.	72%	
Q20. Any information back up storage?	71%	
Q36. I am fully aware how big data analytics is driving the ecommerce agenda.	71%	
Q22. I rate data as an asset for my business.	70%	
Q26. Data security is major issue with SMEs.	68%	
Q35. I am currently using the publicly available data to drive my business.	68%	
Q47. I believe I am rapidly adopting to meet the ecommerce demand of my industry.	67%	
Q18. Average number of customers.	66%	
Q54. I understand what open source technologies are.	65%	
Q55. I believe that open source technologies and applications can add value to any business.	65%	
Q12. How easy is it to get the resources you need to run your business?	62%	
Q28. I believe my business has enough resources	62%	

to manage data and information.		
Q27. I believe my business has enough skills to manage data and information.	60%	
Q56. I see open source technologies as the solutions to the SMEs IT needs.	59%	
Q60. I believe open source technologies have the ability to provide my business with the competitive advantage.	59%	
Q24. I believe that I do not need big budget to collect and store data in my business.	56%	
Q52. I believe in multiple cloud based solutions.	55%	
Q53. I foresee cloud based solutions as a critical partner for my business.	53%	
Q41. I understand the usage, maintenance and management processes of big data.	51%	
Q10. Do you have any website for your business?	50%	
Q19. Records keeping method?	50%	
Q17. Communication channels	48%	Email
Q57. I believe that open source technologies are safe and secure.	45%	
Q31. My business is geared to adopt big data technologies.	44%	
Q50. I believe my business has technologies and resources for the customer retention and the acquisition.	43%	
Q16. Advertising & Marketing channels	42%	W.O.M
Q38. I have resources that can assist me in implementing big data technologies.	38%	
Q14. Any IT intellectual property?	34%	

Q58. I believe that there are skills available in my business to leverage on open source technologies.	33%	
Q15. Is IT a core part of your business?	32%	
Q59. My business has the necessary support and structure to deploy open source technologies.	31%	
Q40. It is easy to adopt big data technologies.	26%	
Q39. I have the necessary budget that can assist me in implementing big data technologies.	22%	

3.7.2 Key highlights

Access to budget (Q39: 22%) and the perception that Big Data technologies are complex and expensive (Q40: 26%) appear to be the biggest obstacle with regard to adopting Big Data technologies.

It is important to note that almost all SMEs interviewed understood the basics of data collection (Q25: 94%) and also believe that data does provide the competitive advantage needed for them to remain relevant.

Support and skills to deploy open source technologies (Q59) appear to be significant stumbling blocks to SMEs adopting Big Data technologies. According to theory, Q55 and Q56 from the questionnaire showed that the SMEs are not aware of open-source technologies but perception on skill and support are obstacles.

3.8 DESCRIPTIVES

The figures listed show the different demographics from the research participants and also interpret these responses in relation to this study.

3.8.1 A socio-demographic profile of repondents:

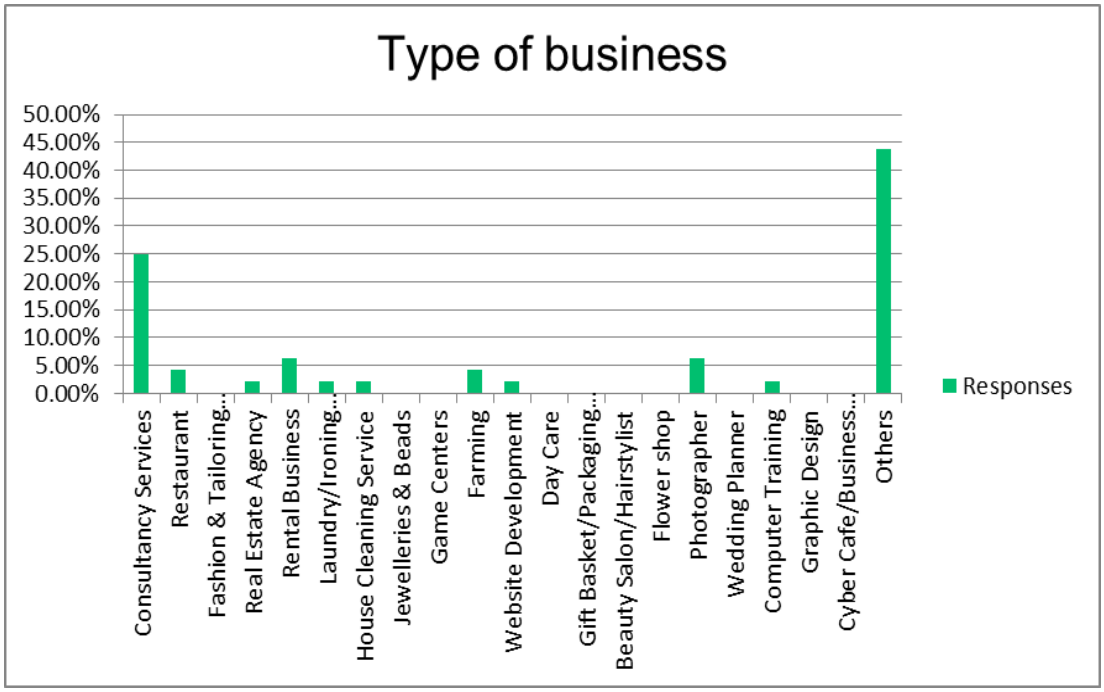


Figure 3.1: Type of SMEs.

Figure 3.1 depicts the types of SMEs that participated in the survey. Other types of business provided the largest response of 43% followed by Consulting Services with 25%.

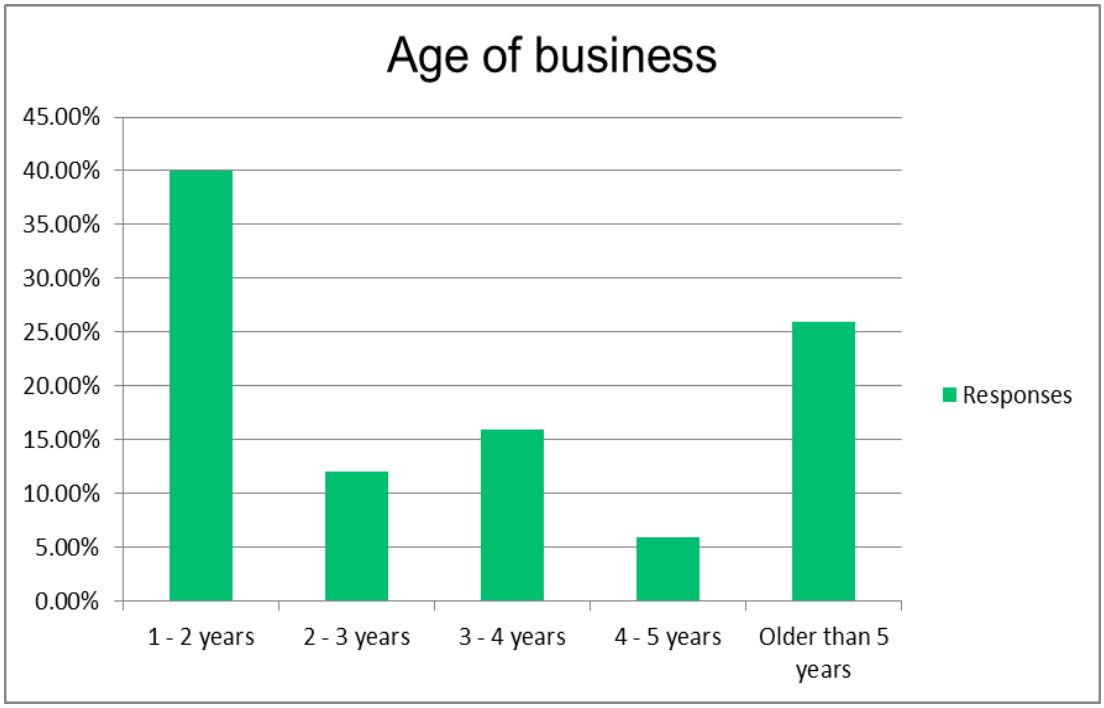


Figure 3.2: Age of business.

Figure 3.2 depicts how long the SMEs have existed. 40% of participants are between 1-2 years old, followed by those that are older than 5 years, while 5% are between 4 -5 years.

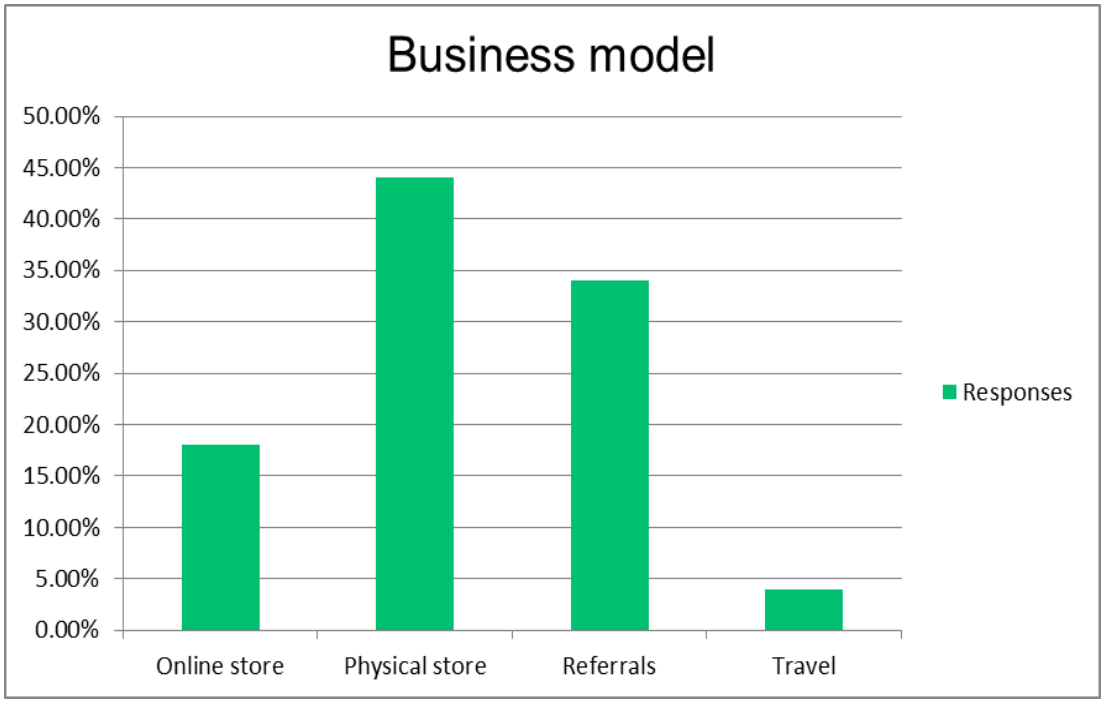


Figure 3.3: Business model.

Figure 3.3 shows the type of business model adopted by the SMEs. The physical stores are on 44% and online stores shows only 17%.

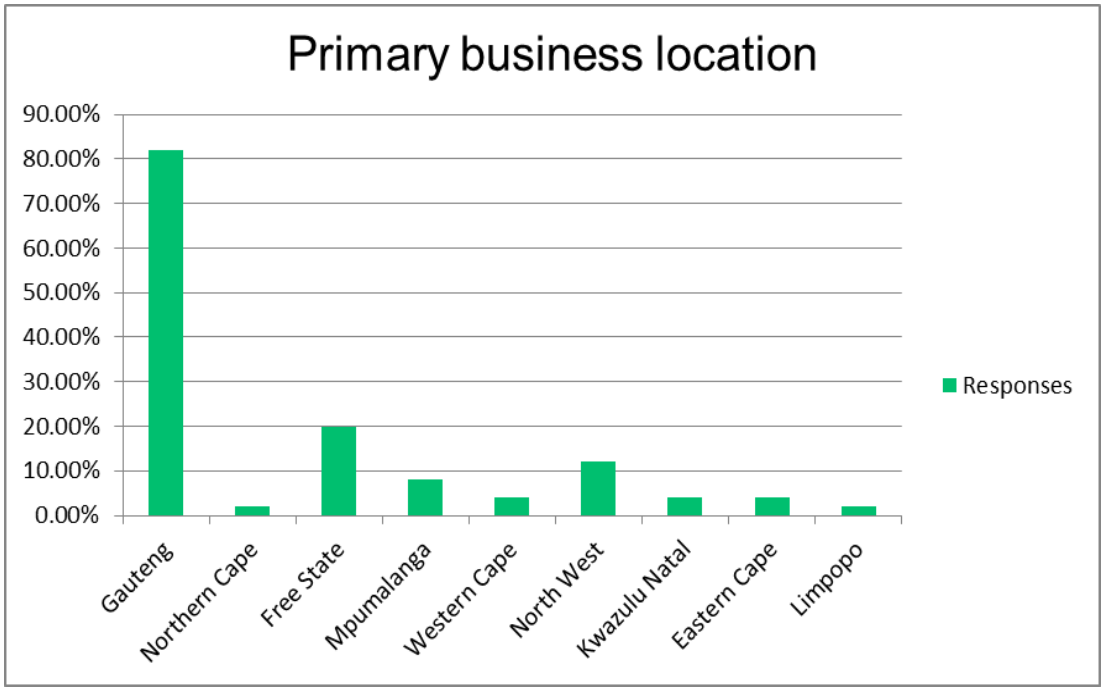


Figure 3.4: Primary business location

Figure 3.4 shows the geographical representation of the SMEs, 80% of the SMEs are located in the Gauteng Province.

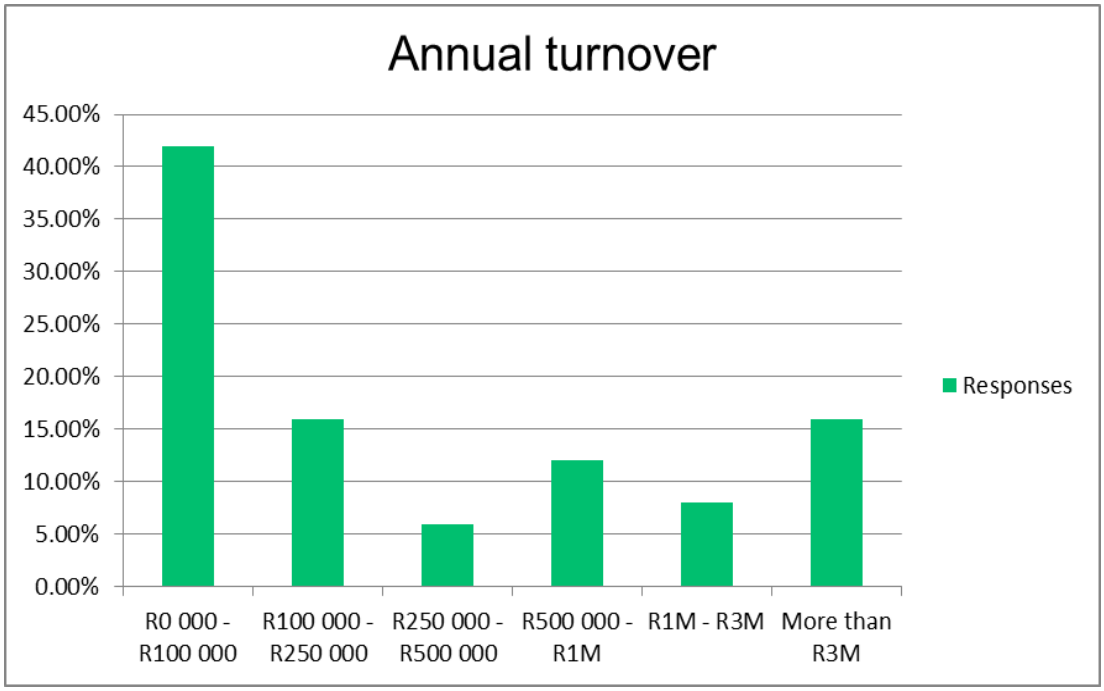


Figure 3.5: Annual turnover

Figure 3.5 depicts the annual turn over of the SMEs, 40% of which is below R100 000.

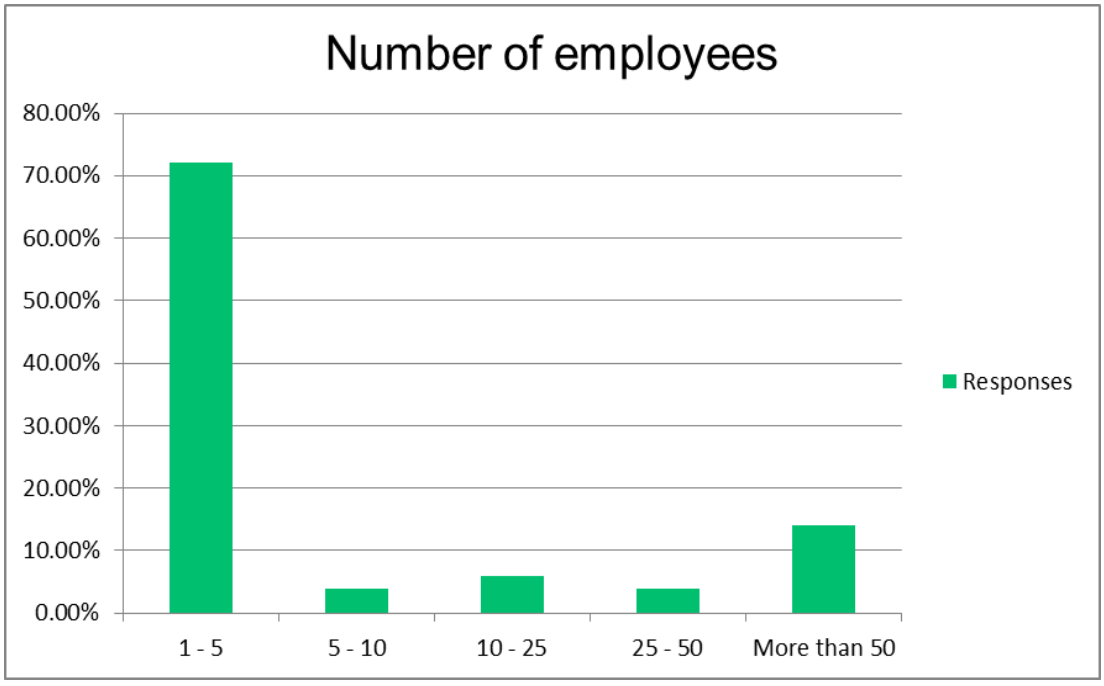


Figure 3.6: Number of employees

Figure 3.6 shows that 70% of the SMEs employ less than five people.

3.9 SURVEY RELIABILITY

This section analyses the response data received from the survey questionnaires. The data was then tested for survey response reliability and internal consistency using the Cronbach's alpha. If the survey is not reliable then the research conclusions and recommendations would be incorrect. Cronbach's alpha is bounded between zero and one, and if closer to one then the reliability is very strong and if closer to zero then the responses are unreliable. Our overall Cronbach's alpha for the entire survey was 0.66 and for individual questions greater than 0.6 which means that the survey responses are reliable. Anything response less than 0.5 is deemed unreliable according to the industry. The table below show the reliability figure for each question (Cho & Kim, 2014).

Our overall Cronbach's alpha for the entire survey was 0.66 and for individual questions grater that 0.6 which means that our survey was reliable. Anything that is less than 0.5 is deemed not reliable according to the industry. The table below show the reliability for each question.

Table 3.2: Cronbach's alpha test on questions

Question	**Alpha**
Type of business	0.81
Data Storage and Security as challenge in terms of storing your company data	0.67
Do you have any website for your business?	0.66
Any IT intellectual property?	0.66
Is ITÂ core part of your business?	0.66
I believe I am rapidly adopting to meet the ecommerce demand of my industry.	0.66
Any mobile app used in your business?	0.66
Technologies as challenge in terms of storing your company data	0.66
I understand what open source tech2logies are.	0.66
Advertising & Marketing channels	0.66

Data security is major issue with SMEs.	0.66
I understand the basics of data collection and management.	0.66
Business model	0.66
I believe I understand the needs and the preferences of my customers.	0.66
Do you use any social media tool for your business?	0.66
How easy is it to get the resources you need to run your business?	0.66
I believe I understand the future needs and preferences of customers.	0.66
I believe that there are ecosystems that we can share and leverage of data in our industry.	0.66
Any information back up storage?	0.66
Accessibility as challenge in terms of storing your company data	0.66
I believe my business can be competitive in future.	0.66
Budget as challenge in terms of storing your company data	0.66
Primary business location	0.66
I believe in digitalization which is taking place in my industry.	0.65
Records keeping method?	0.65
Skill as challenge in terms of storing your company data	0.65
Communication channels	0.65
I believe that open source technologies are safe and secure.	0.65
I believe my business has technologies and resources for the customer retention and the acquisition.	0.65
Type of entity	0.65
I believe that I don't need a big budget to collect and store data in my business.	0.65
I believe in multiple cloud based solutions.	0.65
I believe open source technologies have the ability to provide my business with the competitive advantage.	0.65

I believe information management and business intelligence can improve my business.	0.65
I see open source technologies as the solutions to the SMEs IT needs.	0.65
I believe that data usage can provide my business with the competitive advantage.	0.65
I believe that open source technologies and applications can add value to any business.	0.65
I believe data analytics can create a value in my business.	0.65
It is easy to adopt Big Data technologies.	0.65
I believe that data usage can assist in growth of my business.	0.65
I believe that there are skills available in my business to leverage on open source technologies.	0.65
My competitors are actively adopting technologies.	0.65
My business has the necessary support and structure to deploy open source technologies.	0.65
I am currently using the publicly available data to drive my business.	0.65
I understand the usage, maintenance and management processes of Big Data.	0.65
I believe my business has enough skills to manage data and information.	0.65
I foresee cloud based solutions as a critical partner for my business.	0.65
I often read about the topics on how technology can improve the business.	0.65
I believe I can use the data that I collect to create better client value propositions.	0.65
I believe my business has enough resources to manage data and information.	0.64
I am aware of how I can leverage my data in my business decision making.	0.64
I have the necessary budget that can assist me in implementing Big Data technologies.	0.64

I have resources that can assist me in implementing Big Data technologies.	0.64
I rate data as an asset for my business.	0.64
I am fully aware how Big Data analytics is driving the ecommerce agenda.	0.64
My business is geared to adopt Big Data technologies.	0.64
Data is a key driver in my business decisions	0.64
Number of IT resources you have in your business.	0.64
Number of businesses	0.64
Age of business	0.64
Average number of customers.	0.63
Number of employees	0.63
Annual turnover	0.63

Pearson Correlation tests on continous variables

		1	2	3	4	5	6	7	8
1	Do you have any website for your business	1							
2	Do you use any social media tool for your business	0.53	1						
3	Skill as challenge in terms of storing your company data	-0.34	-0.17	1					
4	Data Storage and Security as challenge in terms of storing your company data	-0.08	-0.19	0.17	1				
5	I believe that data usage can provide my business with the competitive advantage.	-0.02	-0.06	-0.06	0.13	1			
6	I believe that data usage can assist in growth of my business.	-0.13	-0.12	-0.08	0.11	0.89	1		
7	I believe that open source technologies are safe and secure.	-0.08	-0.06	0.03	-0.17	0.19	0.19	1	
8	I understand what open source technologies are.	0.06	0.2	-0.04	-0.03	-0.02	-0.03	-0.46	1

Figure 3.7: Pearson Correlation tests on continous variables

For Pearson correlation tests, the general rule is that if the correlation between 2 variables is above +/- 0.7 then the variables are highly correlated, if +/- 0.4 --- +/- 0.7 correlated, if +/-0.2 -- +/- 0.4 weakly correlated and if 0 -- +/- 0.2 then there is no correlation. If the value is positive then it is called positive (+) correlation, if it is negative then it is called negative (-) correlation. As an example, variable 1 (using website for business) is positively correlated to variable 2, (using social media as a tool for business) (Freeman & Young, 2016).

This study explores the correlation in the data set by selecting some of the variables that are related to the theme that this research is trying to address. A total of eight questions were used and the Pearson correlation was evaluated. The following deductions were drawn from the correlations:

- There is a weak positive correlation response between having a website for the business and using social media tools for business (0.53).
- There is a very strong positive correlation between the perception of data usage to gain competitive advantage and using data to assist in growing the business (0.89)
- A weak negative correlation was observed between understanding open source technology and the perception of it being secured, indicating that although respondents understand open source technology they do not think it is secured (-0.46).
- There is also a weak negative correlation between skill as a data challenge and having a company website (-0.34).

SMEs that have websites also use the social media platforms. SMEs believe that usage of data can actually assist their enterprises to grow and also gain competitive advantage. They do not trust open-source security and lastly, they believe that no special skill is required to use a website.

3.9.1 What are the perceived barriers to adopting Big Data technologies: Univariate Analysis.

Hypothesis 1: Respondents are neutral with regard to the barriers to adopting Big Data technologies i.e. respondent generally chose a score of 3 on the Likert scale for this item.

Assumptions:

1. Assume that if respondents are neutral, they would choose a score of 3 on the Likert scale
2. Assume that a score above 3 represents a positive score and a score of below 3 is a negative score.

In order to understand what respondents generally perceive as the major barriers, a univariate analysis was carried out. The results are listed below.

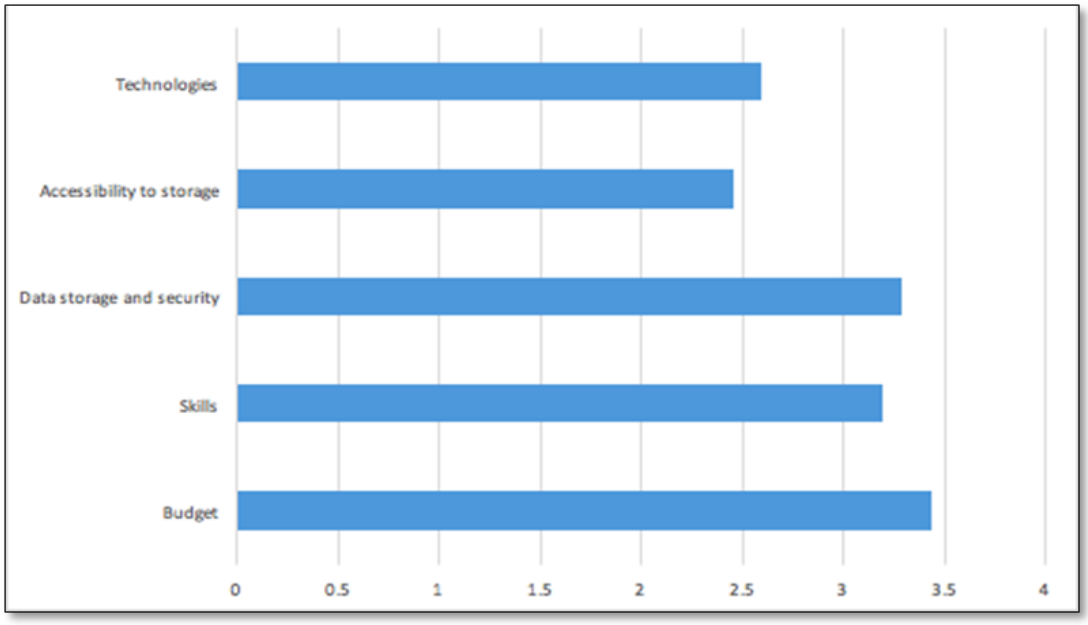


Figure 3.8: Univariate analysis on perception barriers

From the Figure 3.8 above, it is clear that budget, skills and data security were highlighted as the largest impediments. The mean score for these responses was roughly 3.3 and the median of 4 which is higher than the neutral score of 3.

The mean score for this response was 3.99 and the median of 4 which is higher than the neutral score of 3.

Budget, data security as well and skill are regarded as the main barrier when it comes to the adoption of big data technologies by the SMEs.

3.9.2 Hypothesis testing

Methods

This section indicates factors or themes that were particularly strong in the survey response. These themes are usually unobserved and hidden in the data and can only be derived by using dimension reduction technique such as principal component analysis, singular value decomposition and factor analysis.

The general rule used for analysis is that the number of chosen factors or themes must be less than the total number of questionnaires.

Hypothesis 2

Generally, the null hypothesis is that there are no themes that can be derived from the survey response versus the alternative hypothesis that there exists at least one theme in the responses received.

The tests were carried out at 5% significance level.

The null hypothesis was rejected, and it was concluded that there exist at least one or more themes in the responses received.

The results are shown below:

The chart below shows the total number of factors or themes on the x-axis, the eigen value on the primary axis (in green) and the orange line shows the cumulative variance explained by the factors. One can decide the number of themes of factors required by visually creating a cut-off level where it is seen that adding more themes does not significantly increase the cumulative variance.

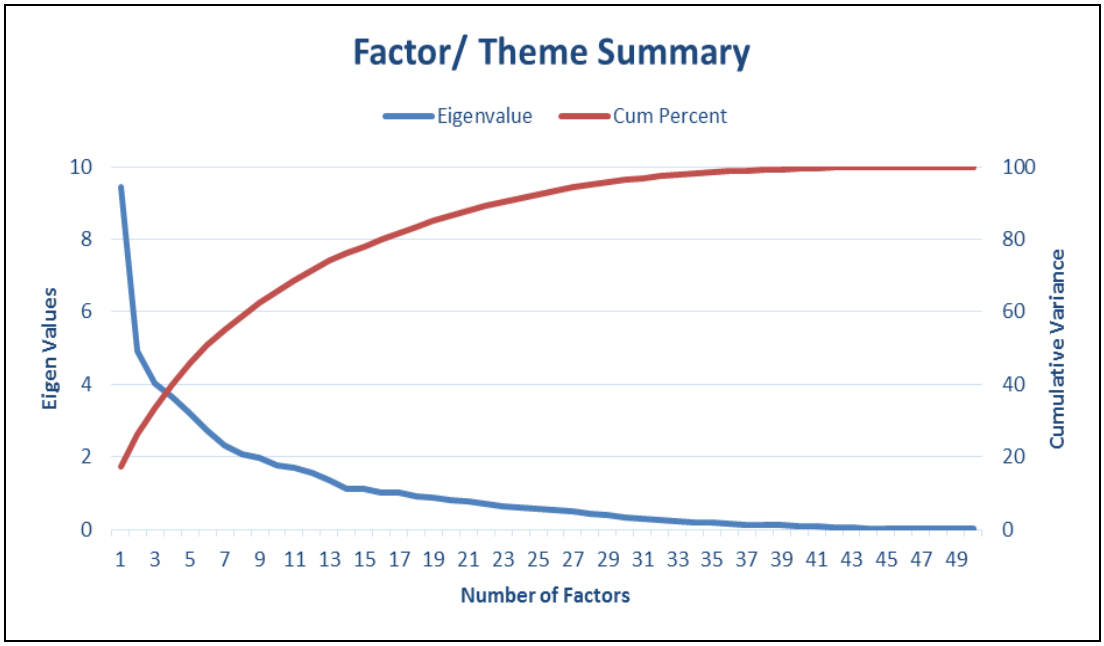


Figure 3.9: Factors/Theme summary

The table below shows the total number of factors or themes that were used in the final model and the cut-off significant level was based on a chi-square test which confirmed that 10 themes were sufficient and explained 65.6% of the variance in the data.

Table 3.3: Factors/Theme variances

Variance Explained			
Factor/Themes	Variance	Percent	Cum Percent
Factor 1: Level of awareness	5.1231	9.315	9.315
Factor 2: Leveraging of	4.777	8.685	18

existing technology			
Factor 3: Belief in big data	4.2754	7.773	25.774
Factor 4: Perceptions of adopting open-source technologies	4.2353	7.701	33.474
Factor 5: Implementing and adopting	4.2124	7.659	41.133
Factor 6: Perception to adoption	3.6586	6.652	47.785
Factor 7: Future growth perception	3.0189	5.489	53.274
Factor 8: Leverage of data	2.4023	4.368	57.642
Factor 9: Accessibility and Storage	2.2549	4.1	61.741
Factor 10: Challenges	2.1472	3.904	65.645

In this section each theme or factor is given a name that is strongly related to it, based on the questionnaire. This practice is commonly referred to as theme loadings. For each theme the questions that are associated with it and the average response is listed.

Table 3.4: Factor 1 - level of awareness

Questions		**Factor 1**
Q37.	I often read about the topics on how technology can improve the business.	0.782962
Q33.	I believe I can use the data that I collect to create better client value propositions.	0.734967
Q32.	My competitors are actively adopting technologies.	0.693466
Q34.	I believe data analytics can create a value in my business.	0.669553

Q36	I am fully aware how big data analytics is driving the ecommerce agenda.	0.663124
Q30.	I am aware of how I can leverage my data in my business decision making.	0.650956
Q35.	I am currently using the publicly available data to drive my business.	0.611661
Q44.	I believe that there are skills available in my business to leverage on open source technologies.	0.538211

This theme groups the respondents who strongly believe that data analytics can create value in their business, the data collected can be used to create better client value proposition and they are currently using publicly available data to drive business.

Table 3.5: Factor 2 - Leveraging of existing technology

Questions		**Factor 2**
Q11.	Do you use any social media tool for your business?	0.755163
Q9.	Any mobile app used in your business?	0.742765
Q10.	Do you have any website for your business?	0.735631
Q15.	Is IT core part of your business?	0.715246
Q14.	Any IT intellectual property?	0.694985
Q16.	Advertising & Marketing channels	0.583675

This theme groups the respondents who use social media, IT is not their core business and there is no IT intellectual property.

Table 3.6: Factor 3 - Belief in big data

Questions		**Factor 3**

Q42.	I believe that data usage can provide my business with the competitive advantage.	0.814783
Q43.	I believe that data usage can assist in growth of my business.	0.785906
Q51.	I believe information management and business intelligence can improve my business.	0.750824
Q23.	Data is a key driver in my business decisions	0.520619
Q53.	I foresee cloud based solutions as a critical partner for my business.	0.495573

This theme groups the respondents who strongly believe that data usage can create competitive advantage, that it is the key driver of their business and that information management and business intelligence can improve their business. This also agrees with the theory in Chapter 2.

Table 3.7: Factor 4 - Perceptions of adopting open-source technologies

Questions		**Factor 4**
Q55.	I believe that open source technologies and applications can add value to any business.	0.816779
Q56.	I see open source technologies as the solutions to the SMEs IT needs.	0.787437
Q57.	I believe that open source technologies are safe and secure.	0.7465
Q60.	I believe open source technologies have the ability to provide my business with the competitive advantage.	0.654547

This theme groups the respondents who had fairly neutral perceptions about the ability of open source to give them competitive advantage and neutral perceptions on open source security.

Table 3.8: Factor 5 - Implementing and adopting		
Questions		**Factor 5**
Q41.	It is easy to adopt big data technologies.	0.724011
Q38.	I have resources that can assist me in implementing big data technologies.	0.702903
Q39.	I have the necessary budget that can assist me in implementing big data technologies.	0.670434
Q41.	I understand the usage, maintenance and management processes of big data.	0.538646
Q45.	I believe my business has technologies and resources for the customer retention and the acquisition.	0.516721

This theme groups the respondents who do not believe that it is easy to adopt Big Data technology or that they have the necessary budget to assist in the implementation of such technology.

Table 3.9: Factor 6 - Perception to adoption		
Questions		**Factor 6**
Q28.	I believe my business has enough resources to manage data and information.	0.8441
Q27.	I believe my business has enough skills to manage data and information.	0.821731

This theme groups the majority of respondents agreed that they have sufficient skills and resources to manage data and information.

Table 3.10: Factor 7 - Future growth perception		
Questions		**Factor 7**

Q46.	I believe in digitalization which is taking place in my industry.	0.663424
Q48.	I believe I understand the needs and the preferences of my customers.	0.650542
Q45.	I believe my business can be competitive in future.	0.564562
Q47.	I believe I am rapidly adopting to meet the ecommerce demand of my industry.	0.470028

This theme groups the respondents who respondents believe in digitalization taking place in the industry, rapidly adopting to meet e-commerce demands and that they understand their customers' needs and preferences.

Table 3.11: Factor 8 - Leverage of data

Questions		**Factor 8**
Q44.	I believe that there are ecosystems that we can share and leverage of data in our industry.	0.63474
Q24.	I believe that I don't need big budget to collect and store data in my business.	0.496579
Q22.	I rate data as an asset for my business.	0.439416

This theme groups the respondents who believe that there is an existing ecosystem where they can share and leverage data and that they do not need large budgets for collecting and storing data. 70% respondents rate data as being an asset for their business.

Table 3.12: Factor 9 - Accessibility and Storage

Questions		**Factor 9**
Q20.	Any information back up storage?	0.629014
Q26.	Data security is major issue with SMEs.	0.409633

This theme groups the respondents who have back-up storage for data and strongly agree that data security is a major issue.

Table 3.13: Factor 10 - Challenges		
Questions		**Factor 10**
Q21.	Accessibility as challenge in terms of storing your company data	0.760701
Q21.	Data Storage and Security as challenge in terms of storing your company data	0.619487
Q21.	Skill as challenge in terms of storing your company data	0.366563
Q21.	Budget as challenge in terms of storing your company data	0.320917

This theme group the respondents who disagree that accessibility, skills and budget are a challenge in storing company data.

Table 3.14: Summary Table

Factor	**Mean Response Score**
Level of Awereness	3.8
Leveraging of exsiting technology	3.75
Belief in BIG data	4.05
Perceptions of adopting open-source technologies	3.56
Implementing and adopting	3.07
Perception of adoption	3.52
Future growth perceptions	2.85
Leverage of existing data	2.48
Accessibility and Data security challenges	2.39

In conclusion, the level of awareness concerning Big Data technology within SMEs is very high and there is generally a highly positive belief that it will provide their business with competitive advantage. The current challenges faced include the lack of trust in open

source technology due to security and data storage fears. There is also a general perception that Big Data technology would require highly skilled personnel and unaffordable resources to implement and maintain.

This agrees with the theory once again; it states that adoption of big data technologies can in fact provide the SMEs with competitive advantage; it mentioned that the open-source security is not trusted.

Hypothesis 3

Model Specification:

The research design model was built to assist in identifying factors that are associated with increased business turnover.

This section employs a generalised linear model using company annual turnover and the ten themes that were created. The whole model test below tests the following hypothesis:

There is no predictor of annual revenue from the created themes versus the null hypothesis that at least one theme is the predictor of annual revenue.

At 1% level of significance the null hypothesis was rejected and it was concluded that at least one theme is a predictor of annual turnover (p-value< 0.01) (Levine *et al.,* 2008:324)

Target Definition

The following transformation was carried out on the variable “annual turnover”:

$$Target = \begin{cases} 1 \; if \; annual \; turnover > 500k \\ 0 \; if \; annual \; turnover =< 500K \end{cases}$$

Sampling

For the purpose of the modelling exercise, the full model data set was used for training the model and a simple random selection of 20% of the data was used to validate the model.

Methodological Approach

Holmes, (2011:13) explained logistic regression as models relationship between set of variables or covariates. This regression was chosen because the model has a dichotomous dependent variable. The logit model predicts the log of the odds of a categorical event or outcome. When the log of the odds from the resulting model is

converted back to a probability the possible range of values is restricted to be between 0 and 1.

For k explanatory variables and i= 1 ,…, n individuals, the model is

Equation 1

$$log\left[\frac{p_i}{1-p_i}\right] = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik}$$

Where p_i is the probability that the dependent variable is 1.

By the very nature of the linear summation, the logit model assumes that the predictors are linear in log-odds space with the outcome. Unfortunately, few relationships are exactly linear. To mitigate the impact of the linearity assumption and for other beneficial reasons, some predictors have been transformed.

Analysis and Fit Statistics

At 1% level of significance the null hypothesis was rejected and it was concluded that at least one theme is a predictor of annual turnover (p-value< 0.01).

Whole Model Test

Model	-LogLikelihood	L-R ChiSquare	DF	Prob>ChiSq
Difference	22.245632	44.4913	10	<.0001*
Full	81.1542614			
Reduced	103.399893			

Goodness Of Fit Statistic	ChiSquare	DF	Prob>ChiSq	Overdispersion
Pearson	75.2124	39	0.0004*	1.5042
Deviance	75.2124	39	0.0004*	

AICc
194.7410

Figure 3.10: Whole factor test

Turner (2008:7-9) explained the Generalized Linear Models (GLM) as a general linear regression that permits for response variables that have error distribution models other than a normal distribution. The table below shows the GLM fit and each of the derived themes. At 5% level of significance, five themes were significant while at 1% level of significance only 3 themes were significant.

The results indicated that Leveraging of existing technology, leveraging of data and Accessibility and storage are the three most important themes in predicting the annual turnover of a company.

Table 3.15: Factors/Themes

Factors/Themes	DF	L-R ChiSquare	Prob>ChiSq
Level of awareness	1	0.6945009	0.4046
Leveraging of existing technology	1	22.12109	<.0001*
Belief in big data	1	5.4034567	0.0201*
Perceptions of adopting open-source technologies	1	2.2675888	0.1321
Implementing and adoption	1	1.9875091	0.1586
Perception of adoption	1	6.1779629	0.0129*
Future growth perception	1	3.1788503	0.0746
Leverage of data	1	10.309514	0.0013*
Accessibility and storage	1	13.61905	0.0002*
Challenges	1	0.0402868	0.8409

The table below shows the strength of each theme in a GLM in terms of being a predictor of annual turnover. The estimates are the model betas and the higher the chi-square value the higher its influence on predicting annual turnover. It can be seen that the highest contributing theme to the prediction of annual turnover is Leveraging of existing technology.

Table 3.16: Strength of each theme

Term	Estimate	Std Error	L-R ChiSquare	Prob>ChiSq	Lower CL	Upper CL
Intercept	2.7602784	0.1734502	90.127358	<.0001	2.4136868	3.10687

Level of awareness	0.1451498	0.1735683	0.6945009	0.4046	-0.201678	0.4919772
Leveraging of existing technology	-0.903636	0.1713117	22.12109	<.0001	-1.245955	-0.561318
Belief in BIG data	0.3697857	0.1548015	5.4034567	0.0201	0.0604584	0.6791131
Perceptions of adopting open-source technologies	-0.264241	0.1734906	2.2675888	0.1321	-0.610914	0.0824309
Implementing and adoption	0.2487275	0.1746784	1.9875091	0.1586	-0.100318	0.5977731
Perception of adoption	0.4458094	0.1738494	6.1779629	0.0129	0.0984202	0.7931986
Future growth perception	0.3170699	0.1750172	3.1788503	0.0746	-0.032653	0.6667926
Leverage of data	-0.479988	0.1418533	10.309514	0.0013	-0.763442	-0.196534
Accessibility and storage	-0.687444	0.1737479	13.61905	0.0002	-1.03463	-0.340257

3.10 CONCLUSION

In conclusion, the SMEs that are highly aware and are leveraging off existing data are associated with a higher annual turnover. These SMEs do not necessarily believe in Big Data technology but they have embraced data technology and are currently facing minimal challenges with regard to data management processes.

Approach

Data was collected in the form of a survey questionnaire. A total of 55 respondents were collected. The following approach was used to analyse the data:

The survey had two broad sections, namely the demographic profile of the business and the descriptive questionnaire, mainly based on a 5 point Likert scale, was provided. The internal consistency of the survey was calculated using Cronbach's alpha.

The mean score of each variable was also calculated. The data profiling and descriptive statistics were first drawn up in order to understand the data in terms of frequency distributions, variable correlations, number of useable variables, levels of 'missingness' in the data and the need for variable transformations where necessary.

Pearson correlation tests were carried out for continuous variables in order to ascertain variable correlation.

A test for association on categorical variables was carried out using Fisher's exact test since the data set was small.

Wilcoxon's sign ranked tests were used to assess whether the responses deviated from the median.

An exploratory factor analysis model was developed in order to identify the key themes that were emanating from the questionnaire.

The median scores of the questions relating to each factor were also calculated in order to understand the general views of the respondents.

A logistic regression model was developed to understand the interaction of the factors which are associated with the SME annual turnover.

3.11 SUMMARY

In conclusion, SME's which are highly aware and are leveraging off existing data are associated with higher annual turnover. The SMEs do not necessarily believe in big data technologies but they have embraced data technologies and are currently facing minimal challenges with regards to data management processes.

CHAPTER FOUR

4.1 INTRODUCTION

In this chapter the theory in Chapter two, the in results and the findings from Chapter three will be discussed. The interesting facts, as well as the managerial implication of the researcher's recommendation will also be discussed. The proposals on issues related to the Big Data implementations will be put forward.

4.1.1 Research finings

According to the research findings, the access to budget and the perception that big data technologies are complex and expensive appear to be the biggest obstacle with regard to adopting Big Data technologies.

Support and skills to deploy open source technologies are also significant stumbling blocks to SMEs. The SMEs believe that usage of data can actually assist their enterprises to grow and also gain competitive advantage. But they do not trust open-source security and lastly, they believe that no special skill is required to use a website.

The level of awareness concerning big data technology within SMEs is very high and there is generally a highly positive belief that it will provide their business with competitive advantage. The current challenges faced include the lack of trust in open source technology due to security and data storage fears. There is also a general perception that big data technology would require highly skilled personnel and unaffordable resources to implement and maintain.

It is important to note that almost all SMEs interviewed understood the basics of data collection and also believe that data does provide the competitive advantage needed for them to remain relevant and grow.

The SMEs that are highly aware and are leveraging off existing data are associated with a higher annual turnover. These SMEs do not necessarily believe in Big Data technology but they have embraced data technology and are currently facing minimal challenges with regard to data management processes.

4.1.2 Literature contribution

The budget, data security as well and skill are regarded as the main barrier when it comes to the adoption of big data technologies by the SMEs according to theory. It also states that adoption of big data technologies can in fact provide the SMEs with competitive advantage.

4.2 RECOMMENDATIONS

4.2.1 How to overcome big data adoption challenges

Many different kinds of problems exist with big data technology adoption in general and they range from issues such as costs, data security and privacy, storage and processing, analytical, technical and data value challenges, especially for the SMEs. Another big challenge is skill shortage. In most cases the SMEs find it difficult to find suitable talent with capable skills to work with new technologies and to interpret the data in order to extract meaningful business insights.

Another big problem is a security in the Big Data environment. Information is sourced from many different storage sites such as cloud-based storage, web systems, social media and other physical servers, therefore, making it very difficult to secure the data completely, thus bringing a considerable security risk to the whole environment. But all these challenge can be defeated.

4.2.2 How to deploy big data technologies

Big Data can assist in business decision making, especially with the amount of data that is available; it will require very little effort to understand the customers' needs, expectations and preferences better. Optimization of costs, returns on investment and improved revenues are some of the benefits from Big Data analytics, but unfortunately to get around deployment appear to be significant stumbling blocks to the SME's adopting Big Data technologies.

There are countless sources of the data such as social media, cellphones, email, web and internet and many different types of analytics can be conducted from these sources which will assist the SMEs to grow, based on the results achieved from the wide variety of data available.

Open source data is provided by most search engine organizations such as Yahoo, Google and Wiki, these and other social media applications can be used to improve the business activities at a very low cost.

The SMES that do not have excellent financial muscle can then adopt open source tools to take advantage of the benefits of big data analytics. Cloud computing and open source software provides these companies with an affordable way to embrace Big Data analytics.

Free web analytics such as Google analytics and social media are available to any organization at any given point and time. The skills development and training should be done prior implementation.

SMEs must ensure they adhere to standard and comprehensive data capture procedures without compromising the security measures, they need to protect the data all the time throughout all the levels and functions within the organization.

SMEs can monitor the social media input to establish the public's perception of their products and services in order, to improve the value they offer and better satisfy customers. The negative comments from social media can be used as constructive criticisms.

There are different low-cost digital marketing solutions that can be used by SMEs to expand and become sustainable, such as social media. Digitalization can enable the SMEs to attract an ecosystem of partnerships which help to ensure customer loyalty. It allows the partnership with other stakeholders internally and externally, such as suppliers, intermediaries and offering ecosystems.

Public clouds can be deployed off-site over the Internet and are available to the general public. The benefit of a public cloud is that it offers excellent efficiency and resources at low cost. This type of cloud infrastructure could be quite useful for SMEs which often have limitations with both budget and skills training. Public clouds could be beneficial for SMEs because the analytics and data management services are offered by the cloud service provider which is also responsible for the quality of service.

SMEs can take advantage of cloud computing methods to source the opportunities and benefits of Big Data, without making significant investment in technology and skills training

Disaster recovery is another factor that needs to be taken seriously. Should a server or the whole information storage be destroyed, SMEs should be able to use their backed-up data and continue with their business with as little inconvenience as possible

Confidential customer information should not be used without the customer's prior consent, not misused in any way. Policies governing the ethical use of the customer information should be in place and always adhered to

SMEs should also invest in in-house skills training, to avoid paying high fees to Big Data experts and there are the different online courses available at very little cost. There are

many different readings and tutorials which prove that there is guidance available for anyone wishing to make use of Big Data analytics.

4.3 MANAGERIAL IMPLICATION

Implementing Big Data analytics within organizations, especially SMEs, is a big challenge, but it is crucial for them to deal with Big Data. The research has identified a number of managerial implications for SME leaders. The adoption of Big Data analytics is a complicated exercise. The SMEs should be up skilled in their understanding of and dealing with Big Data analytics.

The SMEs leaders need to consider factors such as:

- Is data regarded as an asset that can assist in business decision making by their enterprises?
- Is there a belief that data can be used to improve and grow the business?
- Is there awareness around Big Data analytics and its benefits?
- How can barriers to adopting and implementing Big Data analytics be overcome?

It is suggested that the leaders of SMEs should consider engaging in a change management process. This practice will encourage the adoption and implementation of the Big Data technologies supported by every member of the SME structure.

Once this change has been achieved, then the next step for the SME leaders is to consider if they can use the data and the resources that they have already to adopt Big Data analytics. It has already been proven in the literature study that Big Data analytics adds so much value to a business enterprise that, consequently, it can add as great value to SMEs.

If management sees an opportunity for the use of Big Data analytics, then they need to up-skill or leverage the owned existing resources. They can start by up-skilling their internal staff through an open source course, Google analytics, cloud computing, social media, as well as the Big Data analytics and also raise awareness of how all these technologies can benefit SMEs. It is not affordable to employ Big Data experts because they are expensive due to their scarcity. Therefore, SME employees need to acquire Big Data literacy together with business skills.

Everyone within the SME has to support Big Data implementation project and the value derived from Big Data analytics.

4.4 REFERENCES

Abdulla, S. H. 2014. Exploring the benefits of using Big Data when creating mobile applications. (*In Richard, L. M., ed. 2013.* IT practices for SMEs success series book 3: Big Data for SME: Questions of opportunities, challenges, benefits and operations. University of Derby. p. 7). Date of access: 20 Jun. 2017.

Acts ***see*** South Africa.

Agostini, A. 2013. Winning Customers in the Era of Cloud Business Intelligence: Key Adoption Factors from a Small and Medium Enterprise perspective. *Final Dissertation in Technical Project 15 ECTS, Halmstad*.(Dessirtation - Masters). https://ojs.hh.se/index.php/JISIB/article/download/72/76 . Date of access: 3 Sep. 2017.

Almeida, P. & Bernardino, J. 2016. A comprehensive overview of open source Big Data platforms and frameworks. (Abstract). https://www.researchgate.net/publication/310621654_A_COMPREHENSIVE_OVERVIEW_OF_OPEN_SOURCE_BIG_DATA_PLATFORMS_AND_FRAMEWORKS. Date of access: 28 Jul. 2017.

Alsooj, A. 2013. Opportunists for supermarkets: Big Data application in supermarkets. (*In Richard, L. M., ed. 2013.* IT practices for SME success series book 3: Big Data for SMEs: Questions of opportunities, challenges, benefits and operations. University of Derby. p. 38). Date of access: 24 Jun. 2017.

Alvi, M. 2016. A Manual for Selecting Sampling. *Munich personal repec archive: University of Karachi, Iqra University.* Techniques in Research. https://mpra.ub.unimuenchen.de/70218/1/MPRA_paper_70218.pdf . Date access: 31 Oct. 2017.

Bateman, M. 2015. Implementation and Data Latency Issues Regarding Big Data. (*In Richard, L. M., ed. 2013.* IT practices for SME success series book 3: Big Data for SMEs: Questions of opportunities, challenges, benefits and operations. University of Derby. p. 47). Date of access: 30 Jun. 2017.

Benda, R. 2016. Science of Big Data: Background and Requirements, *Advances in Economics, Risk Management, Political and Law Science journal, Tomas Bata University in Zlín.*(Thesis – Masters). http://www.wseas.us/elibrary/conferences/2012/Zlin/EPRI/EPRI-51.pdf. Date of access: 22 Jul. 2017.

Brown, B., Manyika. J & Chui, M. 2012. McKinsey center for business technology. *Are you ready for the era of Big Data?.* https://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/BTO/PDF/MCBT_Compendium_Perspectives_on_Digital_Business.ashx. Date of access: 23 Aug. 2017.

Brown, B., Schuler, S. & Sikes, J. 2012. McKinsey center for business technology. *Capturing the upside of technology-driven threats.* https://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/BTO/PDF/MCBT_Compendium_Perspectives_on_Digital_Business.ashx. Date of access: 27 Sep. 2017.

Bughin, J., Chui, M & Manyika, J. 2013. Ten IT-enabled business trends for the decade ahead. *McKinsey Quarterly May 2013*. https://www.mckinsey.com/industries/high-tech/our-insights/ten-it-enabled-business-trends-for-the-decade-ahead . Date of access: 19 Sep. 2017.

Bughin, J., Livingstone, J. & Marwaha, S. 2012. McKinsey center for business technology. *Seizing the potential of Big Data.* https://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/BTO/PDF/MCBT_Compendium_Perspectives_on_Digital_Business.ashx. Date of access: 23 Aug. 2017.

Carter, J. 2014. Social Media Strategies in Small Businesses Final Report. *Centre for Enterprise, MMU.* http://www.nemode.ac.uk/wp-content/uploads/2014/04/Carter-Social-media-in-SMEs.pdf . Date of access: 10 Oct. 2017.

Cho, E. & Kim, S. 2014. Cronbach's Coefficient Alpha Well Known but Poorly Understood. Organizational Research Methods (Abstract). http://journals.sagepub.com/doi/abs/10.1177/1094428114555994?journalCode=orma . Date of access: 16 Nov. 2017.

Coleman, S., Göb, R., Manco, G., Pievatolo, A., Martorelle, X. T. & Reis, M.S. 2016. How Can SMEs Benefit from Big Data? Challenges and a Path Forward. *Quality and reliability engineering international,* 32(17). (Abstract). http://onlinelibrary.wiley.com/doi/10.1002/qre.2008/full . Date of access: 17 Aug. 2017.

Creswell, J.W. 2014. Research design: Qualitative, quantitative and mixed methods approach. 4th Ed. London. Sage Publications.

Dall'Omo, S. 2017. African digitalization maturity report 2017. *Siemens.* http://www.siemens.co.za/pool/about_us/Digitalization_Maturity_Report_2017.pdf. Date of access: 23 Jul. 2017.

Davenport, T. D. 2013. Big Data Executive Survey 2013: The State of Big Data in the Large Corporate World. *NewVantage Partners LLC.* http://newvantage.com/wpcontent/uploads/2013/09/Big-Data-Survey-2013-Executive-Summary.pdf . Date of access: 4 Sep. 2017.

Deloitte. 2017. SME digital payments new opportunities to optimise. *The Paytech Revolution Series | 2017*. https://www2.deloitte.com/content/dam/Deloitte/au/Documents/financial-services/deloitte-au-fs-zenith-payments-landscape-report-final-220517.pdf. Date of access:21 Sep. 2017.

Europages, 2016. B2B DIGITAL MARKETING FOR SMEs. *Using digital marketing techniques to develop your company's activity.* http://corporate.europages.fr/wpcontent/themes/wordpress-bootstrap-master/pdf/White-Paper-Europages_B2B%20Digital%20Marketing-for-SMEs.pdf?a4972c&a4972c . Date of access: 5 Aug. 2017.

EY & Nimbus Ninety, N. 2015. Becoming an analytics driven organization to create value. *EY building a better world.* http://www.ey.com/Publication/vwLUAssets/EY-global-becoming-an-analytics-driven-organization/$FILE/ey-global-becoming-an-analytics-driven-organization.pdf. Date of access: 20 Oct. 2017.

Fraser Health Authority. 2011. Quantitative Research Methods and Tools. *Fraser health: research and evaluation.* http://www.fraserhealth.ca/media/2011_11_14_Quantitative-Research-Methods-and-Tools.pdf . Date of access: 23 Oct. 2017.

Freeman, J. & Young, T. 2016. Correlation coefficient: association between two continuous variables. *Tutorial|Scope.* https://www.sheffield.ac.uk/polopoly_fs/1.43991!/file/Tutorial-14-correlation.pdf . Date of access: 17 Nov. 2017.

Heeley, R. 2014. Overcoming the talent shortage in Big Data Analytics, a guide for SMEs. (*In Richard, L. M., ed. 2013.* IT practices for SME success series book 3: Big Data for SMEs: Questions of opportunities, challenges, benefits and operations. University of Derby. p. 80). Date of access: 29 Jun. 2017.

Henke, N. 2016. The age of analytics: Competing in a data-driven world. *Mckinsey & company 2016 in collaboration with Mckinsey analytics.* https://www.mckinsey.com/search?q=The%20age%20of%20analytics%20Competing%20in%20a%20data-driven%20world%202016. Date of access: 08 No. 2017.

Heyningen, M. K. 2015. Big Data Analytics: Turning Big Data into Big Money. KVH Industries, Inc. https://thedigitalship.com/conferences/presentations/2015oslo/2.pdf . Date of access: 27 Jul. 2017.

Holmes, L. 2011. Logistic Regression: Basics Prediction Model: Binary Outcomes. *Nemours Stats 101.* https://www.nemoursresearch.org/open/StatClass/January2011/Class8.pdf . Date of access: 18 Nov. 2017.

Intuit, 2010. Twenty trends that will shape the next decade. *Intuit 2020 report.* https://http-download.intuit.com/http.intuit/CMO/intuit/futureofsmallbusiness/intuit_2020_report.pdfDate of access: 25 Sep. 2017.

Kaur, K., Azad, N. & Singh, P. 2013. Cost effective Cloud based Business Intelligence Model for Small Scale Organizations. *International Journal of Computer Applications (0975 – 8887), 63(8).* https://pdfs.semanticscholar.org/9d63/b354d47871b26f899122a50ad853080693b0.pdf. Date of access: 11 Jun. 2017.

Kessel, P. 2014. Big data Changing the way businesses compete and operate. *Insights on governance, risk and compliance, EY.* http://www.ey.com/Publication/vwLUAssets/EY_-_Big_data:_changing_the_way_businesses_operate/%24FILE/EY-Insights-on-GRC-Big-data.pdf. Date of access: 29 Oct. 2017.

Khan, N., Yaqoob, I., Abaker, I., Hashem, T., Inayat, Z., Ali, W. K. M., Alam, M., Shiraz, M. & Gani, A. 2014. Big Data: Survey, Technologies, Opportunities, and Challenges. *The Scientific World Journal Volume 2014 (2014).* https://www.hindawi.com/journals/tswj/2014/712826/. Date of access: 28 Oct. 2017.

Kim, J. Y. 2016. Digital dividends. *Worldevelopment report, world bank group.* http://documents.worldbank.org/curated/en/896971468194972881/pdf/102725-PUB-Replacement-PUBLIC.pdf. Date of access: 18 Oct. 2017.

Kottier, W. 2017. Marketing strategy for online SMEs. *University of Twente Master,* (Thesis – MBA). http://essay.utwente.nl/71976/1/Kottier_MA_BMS.pdf. Date of access: 1 Sep. 2017.

Kotzé, T. 2007. Guidelines on writing a first quantitative academic article. *Department of Marketing and Communication management, Universtiy of Pretoria.* https://www.researchgate.net/publication/228376642_Guidelines_on_writing_a_first_quantitative_academic_article . Date of access: 21 May. 2017.

Levine, D.M., Stephan, D.F., Krehbiel, T.C. & Berenson, M.L. 2008. Statistics for managers: Using Microsoft Excel. 5th edition. *Upper Saddle River, NJ: Pearson*.

Liu, O., Man, K. L., Chong, W. & Chan, C.O. 2016. Social Network Analysis Using Big Data. *Proceedings of the International MultiConference of Engineers and Computer Scientists 2(1).* http://www.iaeng.org/publication/IMECS2016/IMECS2016_pp639-640.pdf . Date of access: 29 Sep. 2017.

Lyon, L. 2015. Disaster Recovery and Big Data. *Featured Database Articles: DB2.* http://www.databasejournal.com/features/db2/disaster-recovery-and-big-data.html. Date of access: 29 Sep. 2017.

Macinnes, B. 2017. *Help SMEs Benefit From Big Data on a Small Scale.* http://www.microscope.co.uk/feature/Help-SMEs-benefit-from-big-data-on-a-small-scale . Date of accessed: 31 Jul. 2017.

Mahembe, E. 2011. Literature Review on Small and Medium Enterprises' Access to Credit and Support in South Africa. *National Credit Regulator (NCR).* http://www.ncr.org.za/pdfs/Literature%20Review%20on%20SME%20Access%20to%20Credit%20in%20South%20Africa_Final%20Report_NCR_Dec%202011.pdf . Date of access: 28 jun. 2017.

Mittelstadt, B. & Floridi, L. 2015. The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Science and Engineering Ethics (Abstract).* https://www.researchgate.net/publication/277079050_The_Ethics_of_Big_Data_Current_and_Foreseeable_Issues_in_Biomedical_Contexts. Date of access: 01 Oct. 2017.

Naru, U.F. 2016. Multimedia Big Data. How to store and manage Big Data for multimedia. (*In Richard, L. M., ed. 2013.* IT practices for SME success series book 3: Big Data for SMEs: Questions of opportunities, challenges, benefits and operations. University of Derby. p. 135). Date of access: 29 Apr. 2017

Obisesan, F. 2016. Big Data Security Challenges In SMEs: How The Big Data Challenges Can Be Mitigated. (*In Richard, L. M., ed. 2013.* IT practices for SME success series book 3: Big Data for SMEs: Questions of opportunities, challenges, benefits and operations. University of Derby. p. 140). Date of access: 24 May. 2017.

Ogbuokiri, B.O., Udanor, C.N. & Agu, M.N. 2015. Implementing Big Data analytics for small and medium enterprise (SME) regional growth. *IOSR Journal of Computer Engineering* 17(6). p. 35.

https://www.researchgate.net/publication/289175303_Implementing_bigdata_analytics_for_small_and_medium_enterprise_SME_regional_growth. Date of access: 23 Sep. 2017.

Oracle. 2014. Information Management and Big Data: *Oracle white paper:A reference architecture.* http://www.oracle.com/technetwork/database/bigdataappliance/overview/bigdatarefarchitecture-2297765.pdf . Date of access: 23 Apr. 2017.

Rao, C.R., Solka, J.L., Wegman, E.J. 2005. Handbook of Statistics, *Vol. 24: Data Mining and Data Visualisation. Elsevier Books*.

Rausas, M.P., Manyika. J, Hazan. E., Bughin, J., Chui, M. & Said, R. 2012. McKinsey center for business technology. *Internet matters: The Net's sweeping impact on growth, job, and prosperity.* https://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/BTO/PDF/MCBT_Compendium_Perspectives_on_Digital_Business.ashx . Date of access: 23 Aug. 2017.

Rising, C. J., Kristensen, M. & Hansen, S. T. 2014. Is Big Data too Big for SMEs? *Leading Trends in Information Technology Stanford University Summer 2014.* https://web.stanford.edu/class/msande238/projects/2014/GainIT.pdf . Date of access: 29 Apr. 2017.

SAS. 2017. Big Data What it is and why it matters. *Big Data Insights.* https://www.sas.com/en_za/insights/big-data/what-is-big-data.html. Date of access: 20 Aug. 2017.

SBP. 2014. Examining the challenges facing small businesses in South Africa. *Business enviroment specialist.* http://www.sbp.org.za/uploads/media/SBP_Alert_-_Examining_the_challenges_facing_small_businesses_in_SA_01.pdf . Date access: 23 Jun. 2017.

Scholz, T. M. 2016. Big Data in Organizations and the Role of Human Resource Management: A Complex Systems Theory-Based Conceptualization. *PERSONALMANAGEMENT UND ORGANISATION Herausgegeben von Volker Stein, band 5.* https://oapen.org/download?type=document&docid=625772 . Date of access: 07 Nov. 2017.

SEDA. 2016. The small, medium and micro enterprise sector of South Africa. *Bureau for economic research.* http://www.seda.org.za/Publications/Publications/The%20Small,%20Medium%20and%20

Micro%20Enterprise%20Sector%20of%20South%20Africa%20Commissioned%20by%20Seda.pdf . Date of access: 23 Jul. 2017.

Simms, W. 2014. Are You Really Ready For Big Data? An Explanation Of The Skill Shortages For Big Data In SMEs (*In Richard, L. M., ed. 2013.* IT practices for SMEs success series book 3: Big Data for SME: Questions of opportunities, challenges, benefits and operations. University of Derby. p. 198).

Small Business Development Act 1996. Small Business Development. https://www.thedti.gov.za/sme_development/docs/act.pdf . Date of access: 7 Oct 2017.

South Africa. 1996. National Small Business Act 102 of 1996

Surveymonkey, 2017. Make smarter decisions. https://www.surveymonkey.com/. Date of access: 31 Aug. 2017.

Suwaidi, A. A. 2014. Big Data for insurance. (*In Richard, L. M., ed. 2013.* IT practices for SME success series book 3: Big Data for SMEs: Questions of opportunities, challenges, benefits and operations. University of Derby. p. 42). Date of access: 23 Jun. 2017.

Tiell, S. C. & Metcalf, J. 2016. The ethics of data sharing: A guide to best practices and governance. *Accenture technology*. https://www.accenture.com/t20161110T001618__w__/us-en/_acnmedia/PDF-35/Accenture-The-Ethics-of-Data-Sharing.pdf . Date of access: 08 Nov. 2017.

Tole, A. A. 2014. Cloud Computing and Business Intelligence. *Database Systems Journal, 4(2014).* http://www.dbjournal.ro/archive/18/18_5.pdf . Date of access: 2 Jul. 2017.

Turner, H. 2008. Introduction to Generalized Linear Models. *ESRC National Centre for Research Methods, UK.* http://statmath.wu-wien.ac.at/courses/heather_turner/glmCourse_001.pdf . Date of access: 18 Nov. 2017.

Vajjhala, N. R. & Ramollari, E. 2016. Big Data using Cloud Computing - Opportunities for Small and Medium-sized Enterprises. *European Journal of Economics and Business Studies, 4(1) 125-136.* https://www.researchgate.net/publication/305210543_Big_Data_using_Cloud_Computing_-_Opportunities_for_Small_and_Medium-sized_Enterprises . Date of access: 09 Sep. 2017.

Vătuiu, T., Udrică, M. & Tarcă, N. 2013. Cloud computing technology - Optimal solution for efficient use of business intelligence and enterprise resource planning applications. *Journal of knowledge management, economics and information technology, special issue*

December 2013. http://www.scientificpapers.org/wp-content/files/28_Vatuiu-CLOUD_COMPUTING_TECHNOLOGY.pdf . Date of access: 1 Aug. 2017.

Vision graphics, 2017. Data analytics http://www.visiongraphics-inc.com/digital-print-services/customer-data-analytics/ . Date of acess: 30 Oct 2017.

Wall, M. 2014. Can Small Firms Benefit From The Big Data Dividend? *Business reporter, BBC News.* http://www.bbc.co.uk/news/business-27065762. Date of accessed: 26 Apr. 2017.

Wallsten, S. 2015. The competitive effects of the sharing economy: How is uber changing taxis?. *Technology policy institute, studying the global information economy.* https://www.ftc.gov/system/files/documents/public_comments/2015/06/01912-96334.pdf. Date of access: 7 Oct. 2017.

Wedgwood, K. 2014. Big Data and analytics in travel and transportation, beyond the hype: Solutions that deliver big value. *IBM Big Data and Analytics White Paper.* http://www.katakenya.org/incs/uploads/pub/6575626643.pdf . Date of access: 15 Aug. 2017.

Wegener, R. & Sinha, V. 2013. The value of Big Data: How analytics differentiates winners. *Bain & Company.* http://aabdp.org/BAIN%20_BRIEF_The_value_of_Big_Data.pdf . Date of access: 10 Oct. 2017.

World Economic Forum. 2016. World Economic Forum, white paper digital transformation of Industries, in collaboration with Accenture. *Digital Enterprise.* http://reports.weforum.org/digital-transformation/wp-content/blogs.dir/94/mp/files/pages/files/digital-enterprise-narrative-final-january-2016.pdf. Date of access: 12 Sep. 2017.

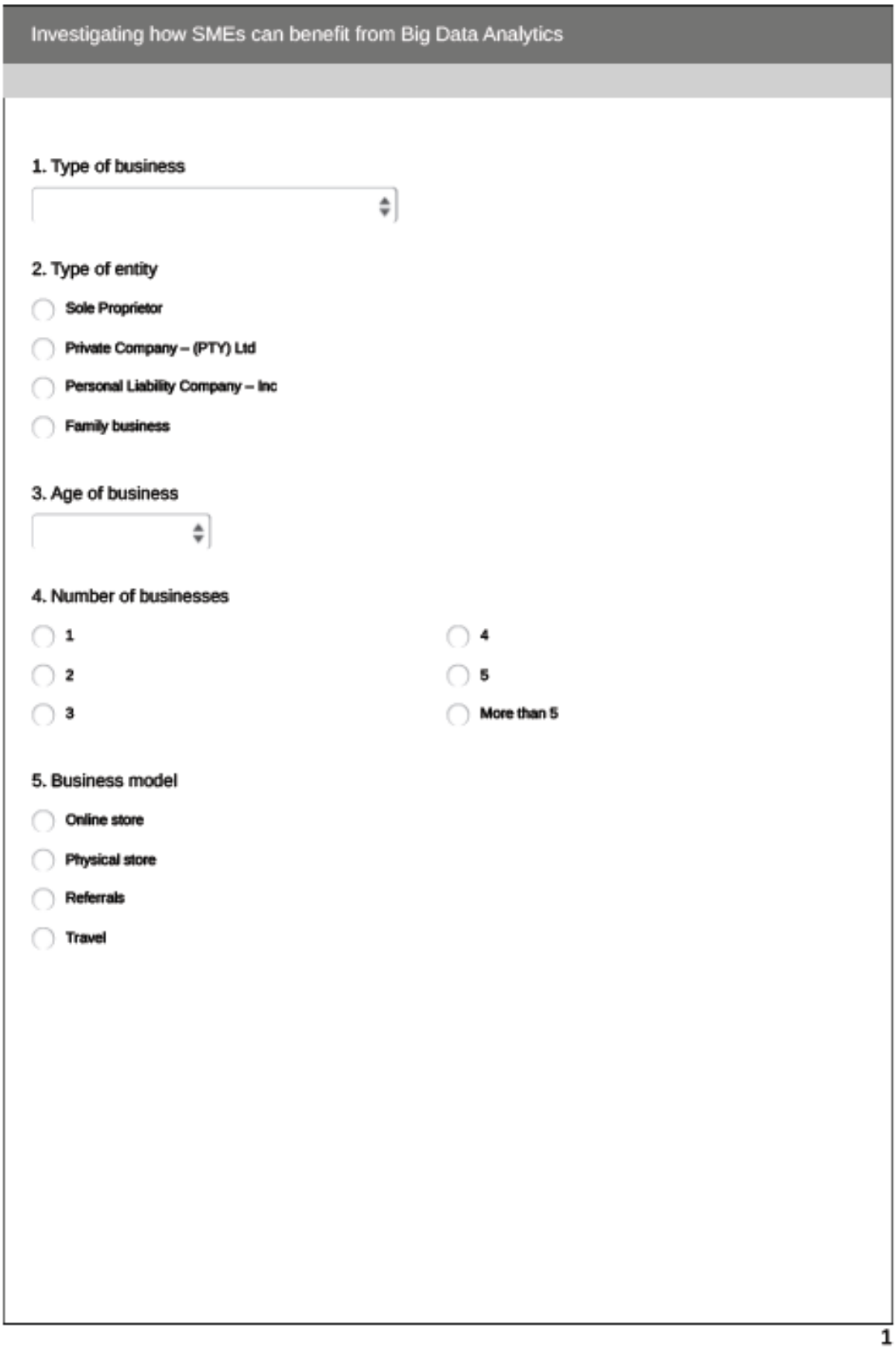
World Economic Forum. 2017. Digital Transformation Initiative In collaboration with Accenture. *Unlocking $100 Trillion for Business and Society from Digital Transformation.* https://www.accenture.com/t20170411T120304Z__w__/usen/_acnmedia/Accenture/Conversion-Assets/WEF/PDF/Accenture-DTI-executive-summary.pdf. Date of access: 14 Oct. 2017.

Zang, S. & Ye, M. 2015. Human Resource Management in the Era of Big Data. *Journal of Human Resource and Sustainability Studies*. (2015) 3, p.41-45. https://file.scirp.org/pdf/JHRSS_2015031714271747.pdf . Date of access: 29 Sep. 2017.

Zicari, R. V. 2015 Big Data: Challenges and Opportunities -ODBMS.org. *Big Data Computing.* http://www.odbms.org/wp-content/uploads/2013/07/Big-Data.Zicari.pdf . Date of access: 08 Nov. 2017.

Zicari, R. V., Rosselli, M., Ivanov, T., Korfiatis, N., Tolle, K., Niemann, R. & Reichenbach, C. 2016. Setting Up a Big Data Project: Challenges, Opportunities, Technologies and Optimization. *Springer International Publishing Switzerland 2016*. http://www.springer.com/cda/content/document/cda_downloaddocument/9783319302638-c2.pdf%3FSGWID%3D0-0-45-1567079-p179864884+&cd=1&hl=en&ct=clnk&gl=za. Date of access: 28 Oct. 2017.

APPENDIX 1: QUESTIONNAIRE



Investigating how SMEs can benefit from Big Data Analytics

1. Type of business

2. Type of entity

- Sole Proprietor
- Private Company – (PTY) Ltd
- Personal Liability Company – Inc
- Family business

3. Age of business

4. Number of businesses

- 1
- 2
- 3
- 4
- 5
- More than 5

5. Business model

- Online store
- Physical store
- Referrals
- Travel

1

6. Primary business location

- [ ] Gauteng
- [ ] Northern Cape
- [ ] Free State
- [ ] Mpumalanga
- [ ] Western Cape
- [ ] North West
- [ ] Kwazulu Natal
- [ ] Eastern Cape
- [ ] Limpopo

7. Annual turnover

- ◯ R0 000 - R100 000
- ◯ R100 000 - R250 000
- ◯ R250 000 - R500 000
- ◯ R500 000 - R1M
- ◯ R1M - R3M
- ◯ More than R3M

8. Number of employees

- ◯ 1 - 5
- ◯ 5 - 10
- ◯ 10 - 25
- ◯ 25 - 50
- ◯ More than 50

9. Any mobile app used in your business?

- ◯ Yes
- ◯ No

10. Do you have any website for your business?

- ◯ Yes
- ◯ No

11. Do you use any social media tool for your business?

- ◯ Yes
- ◯ No

2

12. How easy is it to get the resources you need to run your business?

- ◯ Extremely easy
- ◯ Very easy
- ◯ Somewhat easy
- ◯ Not very easy
- ◯ Not at all easy

13. Types of IT resources you have in your business.

- ☐ Cash machine
- ☐ PC
- ☐ Laptop
- ☐ Scanner
- ☐ Other (please specify)
- ☐ Pont of Sale machine
- ☐ Payment software
- ☐ Server
- ☐ Cameras

14. Any IT intellectual property?

- ◯ Yes
- ◯ No

15. Is IT core part of your business?

- ◯ Yes
- ◯ No

16. Advertising & Marketing channels

- ◯ Social media
- ◯ TV & Radio
- ◯ Newspapers
- ◯ Word of mouth
- ◯ Pamphlets

17. Communication channels

- ◯ Telephone/Cell phone
- ◯ Email
- ◯ Social media

3

18. Average number of customers.

- ◯ 10 - 20
- ◯ 20 - 50
- ◯ 50 - 100
- ◯ 100 - 500
- ◯ More than 500

19. Records keeping method?

- ☐ Manual
- ☐ Electronically

20. Any information back up storage?

- ◯ Yes
- ◯ No

21. In terms of storing your company data/information, may you please list the following challenges in order of importance? (With 5 being a biggest challenge and 1 being a least challenge).

- Budget
- Skills
- Data storage and security
- Accessibility to storage
- Technologies

22. I rate data as an asset for my business.

- ◯ Strongly disagree
- ◯ Disagree
- ◯ Neutral
- ◯ Agree
- ◯ Strongly Agree

23. Data is a key driver in my business decisions

- ◯ Strongly disagree
- ◯ Disagree
- ◯ Neutral
- ◯ Agree
- ◯ Strongly Agree

4

24. I believe that I don't need big budget to collect and store data in my business.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

25. I understand the basics of data collection and management.

- Yes
- No

26. Data security is major issue with SMEs.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

27. I believe my business has enough skills to manage data and information.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

28. I believe my business has enough resources to manage data and information.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

5

29. Please tick the terms that you are familiar with.

- [ ] ICT
- [ ] Big data
- [ ] Internet of things
- [ ] Social media
- [ ] Digitalization
- [ ] Digital marketing
- [ ] Data mining
- [ ] Data analysis
- [ ] Open Source data technologies
- [ ] Cloud based computing
- [ ] Information Management
- [ ] Business intelligence

30. I am aware of how I can leverage my data in my business decision making.

- ( ) Strongly disagree
- ( ) Disagree
- ( ) Neutral
- ( ) Agree
- ( ) Strongly Agree

31. My business is geared to adopt big data technologies.

- ( ) Strongly disagree
- ( ) Disagree
- ( ) Neutral
- ( ) Agree
- ( ) Strongly Agree

32. My competitors are actively adopting technologies.

- ( ) Strongly disagree
- ( ) Disagree
- ( ) Neutral
- ( ) Agree
- ( ) Strongly Agree

33. I believe I can use the data that I collect to create better client value propositions.

- ( ) Strongly disagree
- ( ) Disagree
- ( ) Neutral
- ( ) Agree
- ( ) Strongly Agree

6

34. I believe data analytics can create a value in my business.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

35. I am currently using the publicly available data to drive my business.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

36. I am fully aware how big data analytics is driving the ecommerce agenda.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

37. I often read about the topics on how technology can improve the business.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

38. I have resources that can assist me in implementing big data technologies.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

7

39. I have the necessary budget that can assist me in implementing big data technologies.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

40. It is easy to adopt big data technologies.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

41. I understand the usage, maintenance and management processes of big data.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

42. I believe that data usage can provide my business with the competitive advantage.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

43. I believe that data usage can assist in growth of my business.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

44. I believe that there are ecosystems that we can share and leverage of data in our industry.

- Yes
- No

8

45. I believe my business can be competitive in future.

- Yes
- No

46. I believe in digitalization which is taking place in my industry.

- Yes
- No

47. I believe I am rapidly adopting to meet the ecommerce demand of my industry.

- Yes
- No

48. I believe I understand the needs and the preferences of my customers.

- Yes
- No

49. I believe I understand the future needs and preferences of customers .

- Yes
- No

50. I believe my business has technologies and resources for the customer retention and the acquisition.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

51. I believe information management and business intelligence can improve my business.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

9

52. I believe in multiple cloud based solutions.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

53. I foresee cloud based solutions as a critical partner for my business.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

54. I understand what open source technologies are.

- Yes
- No

55. I believe that open source technologies and applications can add value to any business.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

56. I see open source technologies as the solutions to the SMEs IT needs.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

57. I believe that open source technologies are safe and secure.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly Agree

10

58. I believe that there are skills available in my business to leverage on open source technologies.

- ◯ Strongly disagree
- ◯ Disagree
- ◯ Neutral
- ◯ Agree
- ◯ Strongly Agree

59. My business has the necessary support and structure to deploy open source technologies.

- ◯ Strongly disagree
- ◯ Disagree
- ◯ Neutral
- ◯ Agree
- ◯ Strongly Agree

60. I believe open source technologies have the ability to provide my business with the competitive advantage.

- ◯ Strongly disagree
- ◯ Disagree
- ◯ Neutral
- ◯ Agree
- ◯ Strongly Agree

11

APPENDIX 2: DECLARATION OF LANGUAGE EDITING

DECLARATION OF LANGUAGE EDITING

The MBA mini-dissertation submitted by Santi Makhele

has been subjected to an English language edit by
Dr Barbara Basel

D.Litt. University of Pretoria, MA Potchefstroom University, BA UNISA
Executive Member of English Academy of Southern Africa
Lecturer in English Literature, Linguistics, Communication and Business English at Pearson Institute for Higher Education (previously CTI/Midrand Graduate Institute). Cape Town Campus.
Academic Editing – MBA Theses, MEd Theses, Master in Graphic Design Thesis, External Examiner for MEd Thesis.

1 Six Oaks,
5, Adelaide Road
Plumstead
Cape Town 7800
Tel: 021 761 4289
Cell: 082 651 1659
barbara.basel@gmail.com

Barbara Basel

19 November 2017

INVOICE

CLIENT'S	

DETAILS	Mr Santi Makhele		
Details of work completed	Editing of MBA thesis relating to implementation and use of Big Data analytics by SMEs Total 53 pages at R45 a page		**R2385.00**
EDITOR'S DETAILS	Dr Barbara Basel	I D No	3805130042087
Address:	No 1 Six Oaks 5 Adelaide Road PLUMSTEAD 7800 Cape Town		
Tel:	021 761 4289	e-mail	barbara.basel@gmail.com
Cell	082 6511695		

BANK DETAILS			
Account Holder:	B E Basel (Dr)	Bank:	First National Bank
Account No:	62091984510	Branch:	Plumstead
Type of Account:	Cheque	Branch Code:	201109
Signature:	Barbara Basel	Date:	20 November 2017