



Enterprise Big Data: Case Study of Issues and Challenges for Businesses in Finance and Retail Sectors

Ehiwe Dunsen Dominic
Computer Science Department
Babcock University, Ilishan
Ogun State, Nigeria

Akinola Kayode E.
Computer Science Department
Abraham Adesanya Polytechnic
Ogun State, Nigeria

Ominike Akpovi A.
ICT Department
Petroleum Training Institute (PTI)
Delta State, Nigeria

ABSTRACT

Big data is the most common term used to describe the exponential rise and data availability, both structured and unstructured. Big data can be very important to business as the internet has become. Why? Accurate analyses can be derived from more data volume which in turn helps in making decisions confidently. And better decisions can mean greater efficiencies in operations, reduction of cost and low risk rate. Analysis of Enterprise Big Data to derive business value has become top priority for the majority of modern businesses and organizations. To most organizations, these are data that comes from online transactions, videos, audios, images, social networking interactions, science data, sensors and mobile phones and their applications etc. Stored in databases, they keep growing massively and becoming increasingly difficult to capture, store, manage, share, analyze and visualize via the use of typical database software tools. The study was conducted to analyze enterprise Big Data: Issues & Challenges for businesses in finance and retail sectors. To extract response for the study, the research design adopted was the survey method using questionnaire as the instrument for data collection. The questionnaires were administered to different categories of people in business enterprises. The results from the analysis consistently suggest that growing data set impacts over 80% of surveyed businesses.

General Terms

Management Information Systems, Big Data, Databases

Keywords

Big Data, Databases, Enterprises, Challenges, Analytics, Organization, Business

1. INTRODUCTION

"Big Data are high volume, high velocity, and high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization" [1]. At all times, big data is being generated by the things around us. Activities like digital processing and social media exchange produces big data; systems, sensors and mobile devices are involved in the transmission of big data. According to data research studies conducted by IBM, the rate of data arrival from multiple sources is alarming [6], [7]. In addition, skills and capabilities in analytics as well as advanced and optimized processing power is required for organizations.

There are several advantages that can be derived from growing data. When effectively managed via reporting and visualization, it can be a catalyst for growth for organizations. However, where an organization fails to harness the capabilities that growing datasets presents, it could fall behind

its competitors or lose deriving benefits that can favorably impact on return on investment. Different approaches are taken by different organizations to capture, analyze and make sense of the data been generated. What technology is adopted depends on a number of factors which may be peculiar to the organization concerned. With competition increasingly becoming fierce in the world of business, organizations are tasked with deriving the most benefit from the data that are being generated in the course of operations. However, when this data grows beyond the size that can be adequately stored and analyzed using the legacy systems available, how do organizations cope with the challenges and leverage the advantages growing data can provide?

The term "data" was academically distinguished from "information" for the first time [5]. He argued that the knowledge taxonomy consists of four categories comprising: data, information, knowledge and wisdom. This was given as the (DIKW) taxonomy. He went ahead to define data as all electronic signals done by computing process. While in the past, the effective management and analysis of large volumes of data was almost impossible due to resource constraints such as high cost and lack of required technology, the trend is changing rapidly in modern times; changes that comes with new realities. The rise of big data has brought both infinite opportunities and challenges in economic, cultural and social sectors. This can be observed nationally and internationally as described [3]. This has in turn, caused private businesses and even government owned organizations to be obsessed with ensuring effective responses to the risks that are derived from data generation and its management.

Enterprise big data and its analysis are at the centre of most modern businesses and organizations. Big data is generated from sources such as online transactions, emails, videos and audios, images, logs, face-book posts, search queries, medical health records, interactions on social networks, machine sensors, mobile phones and their associated applications [8], [9]. They are stored in databases thereby growing massively and becoming difficult to capture, store, manage, share, analyze and visualize via the use of typical database software tools. Traditional methods of analysis are no longer sufficient in analyzing growing data; more importantly big data which is characterized by the following properties:

a. Variety

This is the main feature that makes big data really big; coming from a variety of sources. Generally, there are three types of sources: structured, semi structured and unstructured.

b. Volume

This is the second principal feature of big data. With data volume now larger in size than terabytes and petabytes, the scale and rise of big data outstrips traditional store and analysis techniques.

c. Velocity

The third principal feature of big data. This is in reality required not only for big data but also for all processes. For limited processes, big data should be used as it streams into the organization in order to gain maximum value it.

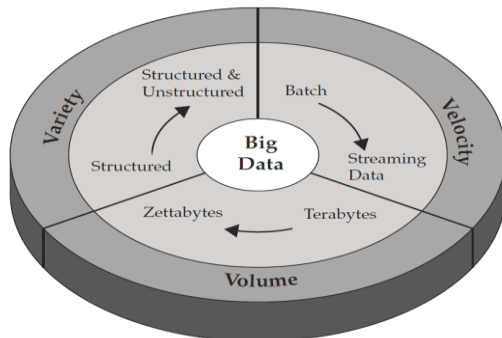


Figure 1: Characterization of Big-Data (Variety, Volume, and Velocity (v3)) [2], [4]

The effective implementation of Big Data initiatives requires an undertaking of appropriate organizational actions, including ensuring organizations are provided with all the necessary resources to enable analysis of the ever growing data sets to what they have access. In this context, the application of proper techniques and technologies is one of the key issues. In practice, however, organizations use many various techniques and technologies to aggregate, manipulate, analyze and visualize their data. These come from fields such as statistics, computer science, applied mathematics and economics. Examples of techniques that have been adopted for the analysis of data and big data are: data fusion and data integration, data mining, machine learning, predictive modeling, sentiment analysis, spatial analysis, simulation or time series analysis.

2. STATEMENT OF RESEARCH PROBLEM

Due to the popularity of Big Data, data is being viewed more as an invaluable asset. Most of the big organizations have put in place departments that are responsible for processing business data and providing needed insights. While procuring investment for Big Data isn't as challenging now as it was a few years ago, however, establishing the short-term and long-term value of such investments continue to be a huge challenge that only a few firms have been able to meet. Business managers across varying sectors and geographical locations are now facing the dilemma: "How do we ensure that enterprise big data projects deliver their true value and provide sustainable benefits?" It is important to look at the dimensions of the big data concept and how this affects businesses. The impact for businesses could take the following forms:

- Objective:** What value creations are we aiming for as a business and how do we leverage on the growing rate of data?

- Analysis task:** As business related data grows, how do we keep carrying on the following; acquisition, integration, storage, search retrieval, analysis, visualization and sharing?
- Challenge:** What are we faced with in terms of privacy, security, uncertainty, inconsistency and modelling conflicts?
- Type:** Do most of the data come in the form of image, audio, video, text, numeric, alpha-numeric data types?
- Source:** Are we generating large data volumes as a result of increase in transactions, events, social media influence or emails?

3. MATERIAL AND METHODS

3.1 Research Methodology

This section explains the research methodology and the method adopted in the gathering of data used in carrying out the study. Identification of the target population and definitions of relevant variables required in achieving the research objective was undertaken as well in this section. In addition, an analysis of the questionnaire design was carried out to highlight the research objective.

Data used for the study came from both primary and secondary data sources. The primary data was sourced through the use of questionnaires administered to the employees and managers of the sampled organizations. The secondary data was sourced from past literatures that are relevant to the study. Databases containing academic materials in forms of journals, articles, dissertations and books were accessed through electronic learning resources provided by Babcock University, Nigeria.

3.2 Research Questions

A review of literature in this study has emphasized the influence of Big Data on enterprise businesses and their performances. This relationship and its relevance therefore led to the formulation of the following research questions:

Research Hypothesis

Hypothesis 1: Relationship between enterprise business performance and Big Data.

Ho: There is no relationship between the performance of a business enterprise and the Big Data generated within the enterprise.

H1: There is relationship between the performances of a business enterprise and the Big Data generated within the enterprise.

Hypothesis 2: Relationship between enterprise business performance and Big Data across organizations within two industrial sectors (Finance and Retail).

Ho: The relationship between the performance of a business enterprise and its' generated Big Data is not similar across businesses in Finance and Retail.

H1: The relationship between the performance of a business enterprise and its' generated Big Data is similar across businesses in Lagos and Birmingham.

4. RESULTS & DISCUSSIONS

The aim of this section is to present a comprehensive analysis of the findings made by the study. A further discussion of the demographics of the sample population was also undertaken. The findings in this chapter were derived from analyzing the

primary data collected through use of questionnaires and direct contact made with a number of the respondents to the survey questions.

4.1 Analysis of Survey Results

Table 4.1: How does your enterprise generate these data, and what sources of data is generated

Data Source		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Primary Source	9	18.0	18.0	18.0
	Secondary Source	21	42.0	42.0	60.0
	Tertiary Source	20	40.0	40.0	100.0
	Total	50	100.0	100.0	

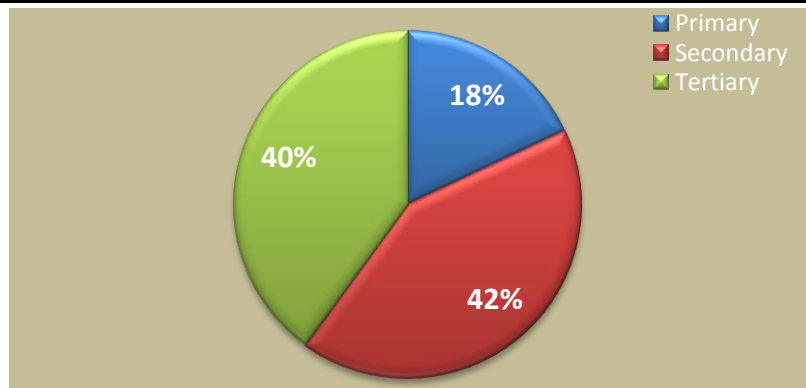


Fig. 4.1: A Pie Chart showing the percentage distribution for source of data generation

Table 4.6 and the corresponding figure shows the percentages of the sources of Big Data generated within the sampled organizations. This result shows that the bulk of data is being derived from secondary sources in these organizations. This

shows also, that within most organizations, data from activities and sources which are indirectly related to the organization constitute the bulk of their data volume.

Table 4.2: Types of Data your Organization operates with?

Types of Data Operation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Transactional Data	19	38.0	38.0	38.0
	Analytical Data	13	26.0	26.0	64.0
	Master Data	18	36.0	36.0	100.0
	Total	50	100.0	100.0	

Table 4.2 shows the composition of the type of data the sampled organizations operate with. The bulk of these data are transactional in nature. This shows therefore that transactional

day to day operational activities generate the bulk of data within the sampled organizations.

Table 4.3: Does Enterprise Big Data have significant impact on your enterprise organizational performance?

Significant Impact of Big Data

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Agree	41	82.0	82.0	82.0
	Not Sure	9	18.0	18.0	100.0
	Total	50	100.0	100.0	

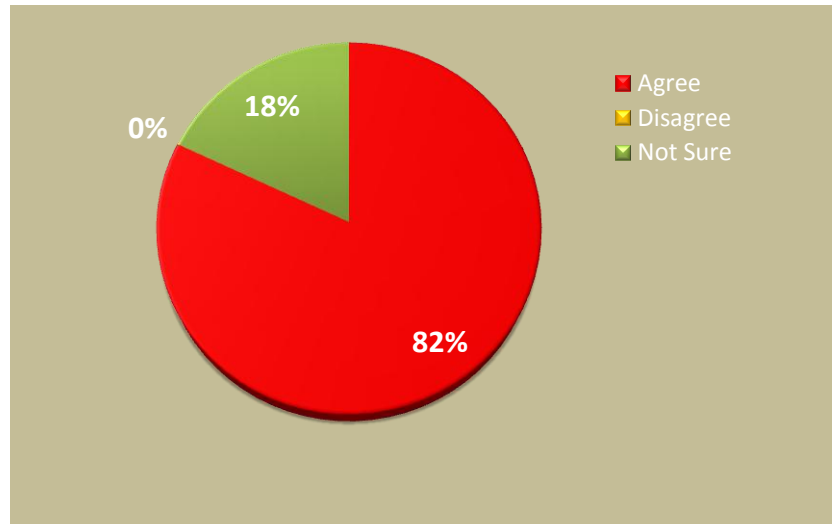


Fig. 4.3: Pie Chart representing the percentage of significant impact of Big Data on organizational performance

Table 4.3 and the corresponding figure gives the level of impact increasing data volume has on the sampled organizations. 82% of respondents' state that Big Data has impact on the operational performance of their enterprises.

While the remaining 18% responded as not sure of the significance of Big Data on their operational performance, this can be safely regarded as reflecting the opinion of the 10% who said they are unaware of the meaning of Big Data.

Table 4.4: Does Big Data pose any Challenges to your Enterprise Organizational Performance?

Challenges of Big Data on Organizational Performance

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Yes	38	76.0	76.0	76.0
No	12	24.0	24.0	100.0
Total	50	100.0	100.0	

Table 4.4 shows the number of respondents who state whether or not Big Data pose any challenge to their organization. 74% believe Big Data has impact on their organizational performance. This is a significant number and as highlighted

in the review of literature. These can be either positive or negative and of different variations depending on factors such as the type of organization, the size of data and how these data are being generated and processed within the organization.

Table 4.5: Challenges encountered by Enterprise Organization as a result of Big Data?

Challenges in Organization due to Big Data

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Storage Capacity	12	24.0	24.0	24.0
Data Management	10	20.0	20.0	44.0
Speed Capacity	9	18.0	18.0	62.0
Complexity in Processing	11	22.0	22.0	84.0
Not Sure	8	16.0	16.0	100.0
Total	50	100.0	100.0	

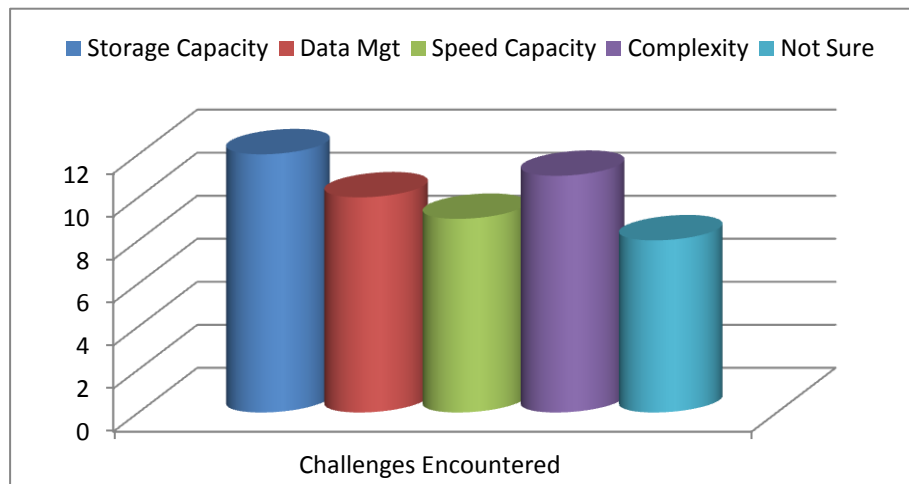


Fig. 4.5: Bar Chart presentation on the challenges Big Data pose on organizational performance

Table 4.5 and the corresponding figure shows the different categories of challenges Big Data pose to the sampled organizations. The greater challenge to organizations is in the form of storage capacity. Owing to the nature of Big Data as data whose volume is increasing in size and velocity, the capacity to store this increasing volume can be a real challenge for organizations. This obviously has an impact on the operational performance of an organization as the cost of storage would likely increase.

The complexity associated with processing is another challenge faced by organizations. As pointed out in the literature review, this capacity is determined by a number of factors. For big sized and well-funded organizations, meeting the cost of adopting better processing structures may not be much of a challenge. However, enterprises which lack the financial capacity to scale up their data processing requirement to accommodate increasing data volume can have their operational performance negatively affected.

Table 4.6: Summary of Response to Survey Information Ranging from the Questionnaire Distribution

S/N	Questionnaire Range (Q6 to Q9)	ALTERNATIVE ANSWERS & FREQUENCIES		
		YES	NO	NOT SURE
1.	Are there effective methods in place to overcome these challenges? (Q6.)	28	-	22
2.	Is your current data analytics tool effective in coping with the volume of data been generated? (Q7.)	29	5	16
3.	Are there specific roles related to Big Data management in your Organization? (Q8.)	37	7	6
4.	Are these roles clearly defined? (Q9.)	22	-	28

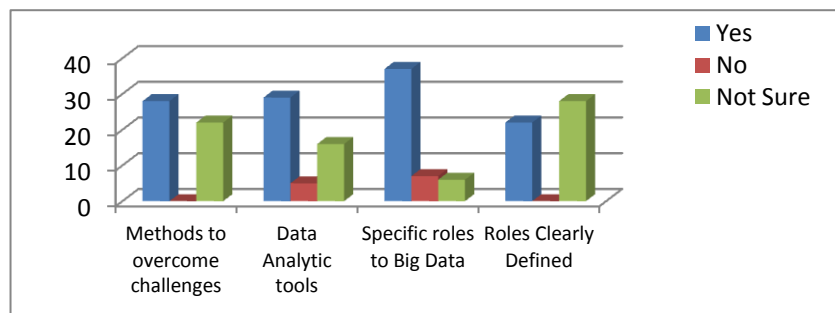


Fig. 4.6: 3-D Multiple Bar Chart presentation of response to frequencies ranging from questionnaire distribution

Table 4.11 with its corresponding figure summarizes the responses to questions 6 to 9 of the survey information. One conclusion that can be drawn from these analyses is that

majority of organizations impacted by Big Data has in place methods for handling the challenges, have defined roles associated with their data management requirements.



However, the majority of these organizations do not have these roles clearly defined in their operational structure. This therefore can be interpreted to mean that the data management

roles are subconsciously carried out by majority of organizations.

Table 4.7: With Big Data, does your organization have in place a system for sorting incoming data almost in real time by potential value, quality of data and frequency of use?

Q10: Systems for sorting incoming data				
	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	22	44.0	44.0	44.0
No	21	42.0	42.0	86.0
Not Sure	7	14.0	14.0	100.0
Total	50	100.0	100.0	

Table 4.7 shows the percentage of the readiness of the sampled organizations in handling incoming data almost in real time. A higher level 44% process their generated data almost in real time. This obviously can enable an enterprise

derive value from their data in terms of type, the frequency and the business decisions that can be made based on findings.

Table 4.8: Dedicated resources can be used to explore the possibilities of advanced analytics in Big Data for your line of business

Q11:Dedicated Resources				
	Frequency	Percent	Valid Percent	Cumulative Percent
Agree	39	78.0	78.0	78.0
Disagree	5	10.0	10.0	88.0
Not Sure	6	12.0	12.0	100.0
Total	50	100.0	100.0	

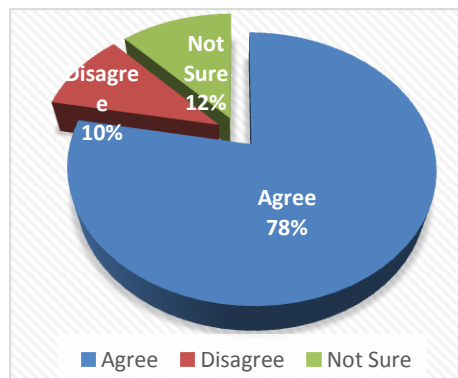


Fig. 4.8: Presentation of 3-D Pie Chart showing the percentage response on how dedicated resources can be of possibilities to advance analytics in Big Data

Table 4.8 and the corresponding figure show that 78% of respondents believe or agree that the use of dedicated resources can be beneficial in advancing analysis of their enterprise Big Data. This obviously has an impact on a

number of other organizational indices such as cost, operating expense, provision of training or the hiring of specialist staff to handle or operate these resources.

Table 4.9: Summary of response to frequencies ranging from Question 12 to 15 of the questionnaire distribution

S/N	Questionnaire Range (Q12 to Q15)	ALTERNATIVE ANSWERS & FREQUENCIES		
		YES	NO	NOT SURE
1.	Does the organization use event driven architecture to manage incoming data? (Q12.)	5	23	22
2.	Does the organization have in place specialized data services that can accommodate the different formats of data generated? (Q13.)	34	12	4



3.	Does the organization have in place specialized data services that can guarantee the security of the data generated? (Q14.)	33	5	11
4.	Does the organization have in place specialized data services that can ensure effective management of data generated? (Q15.)	29	5	16

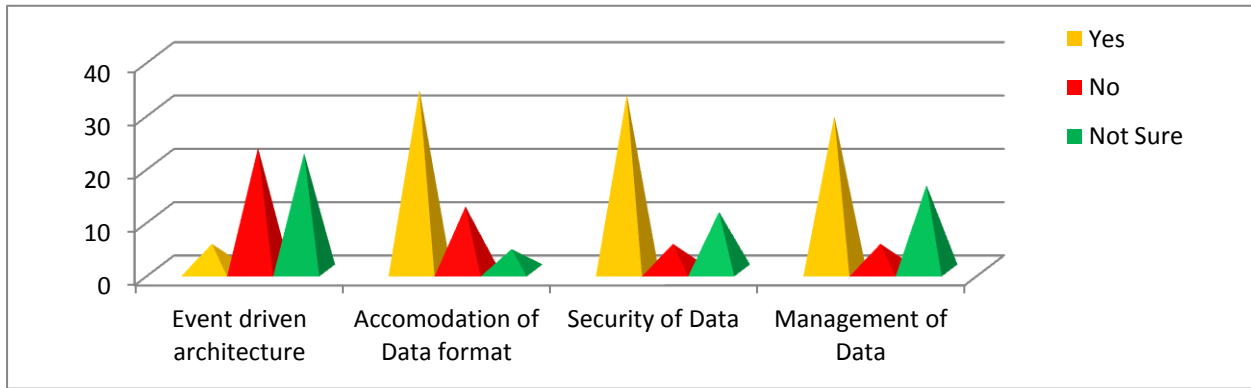


Fig. 4.9: A 3-D Multiple Bar-pyramid chart showing response to frequencies ranging from Question 12 to 15 of the questionnaire distribution.

Table 4.9 and the corresponding figure highlight the readiness of the organizations to accommodate, ensure data security and manage the different format of data that are being generated. One result glaring from the responses is that organizations are

being prepared to handle these Big Data impact on their operational performance as the Big Data phenomenon continues to gain momentum and relevance in modern business and economic settings.

Table 4.10: Summary of response to frequencies ranging from Question 16 to 19 of the questionnaire distribution

S/N	Questionnaire Range (Q16 to Q19)	ALTERNATIVE ANSWERS & FREQUENCIES		
		YES	NO	NOT SURE
1.	Is your organization currently using or considering using in-memory analytics for your organizational data? (Q16.)	8	8	34
2.	Do you have a set of documented policies regarding Big Data governance in place for your organization? (Q17.)	45	0	5
3.	Is risk management involved in your organizational Big Data governance process? (Q18.)	39	0	11
4.	Does your enterprise collaborate with third parties or outside vendors in regards to the planning and execution of your Big Data Projects? (Q19.)	41	0	9

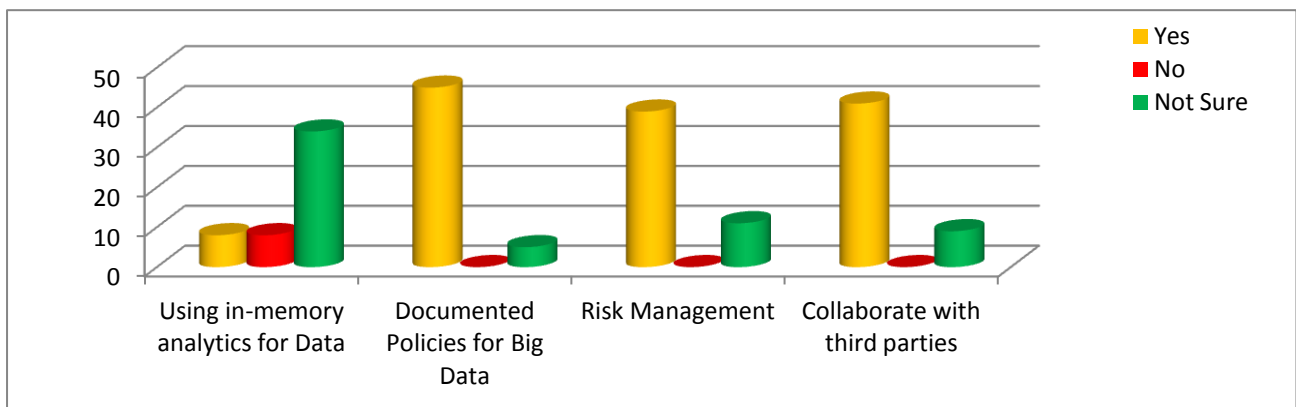


Fig 4.10: A 3-D Multiple Bar-clone chart showing response to frequencies ranging from Question 16 to 19 of the questionnaire distribution



Table 4.15 and the corresponding figure show whether or not risk management is a factor that is considered in the data management policy of the sampled organizations. Also, it is observed that documented policies are in place in most organizations in addressing their data generation, retention

and use. As data (in whichever format it is in) is the life-blood of any organization, the need to ensure that regulations are adhered to and the most beneficial outcomes are derived from these data is a priority for most organizations.

3.	Does your organization have the ability to scale in terms of storage and processing power? (Q22.)	27	0	23
4.	Is there adequate specialized knowledge available to effectively handle the volume of your organizational business? (Q23.)	31	10	9

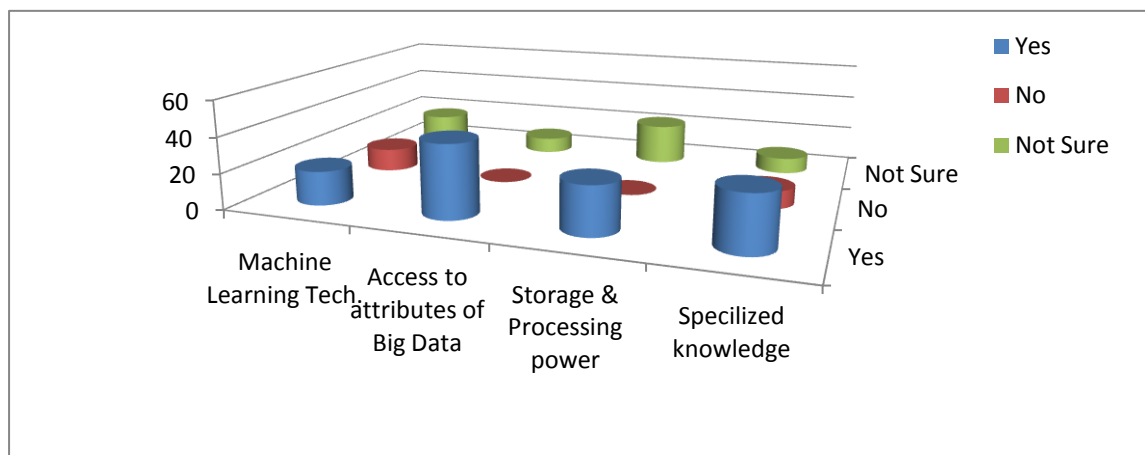


Fig 4.10 Bar graph showing responses to questions 20 to 23 of the survey questionnaire

Table 4.10 and the corresponding figure shows the readiness of organizations in scaling their data storage and processing power to effectively handle the impact Big Data has on their operational performance. One important finding of the survey is that most organizations freely give out information to the attributes of Big Data. A lot of awareness is been created to ensure people know the relevance of the Big Data concept and

the potential values that can be derived when effectively handled.

Also, about 50% of respondents currently adopt machine learning technologies in the analysis of their data. A focus on this technology can help to harness the potential benefits of increasing data and research on how this can be done is a potential area worthy of study in the near future.

Table 4.11: Does Big Data have the potential to provide insights that can transform every enterprise's organizational performance?

Q24:Big Data Potentials to provide insight					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Agree	35	70.0	70.0	70.0
	Disagree	11	22.0	22.0	92.0
	Not Sure	4	8.0	8.0	100.0
	Total	50	100.0	100.0	

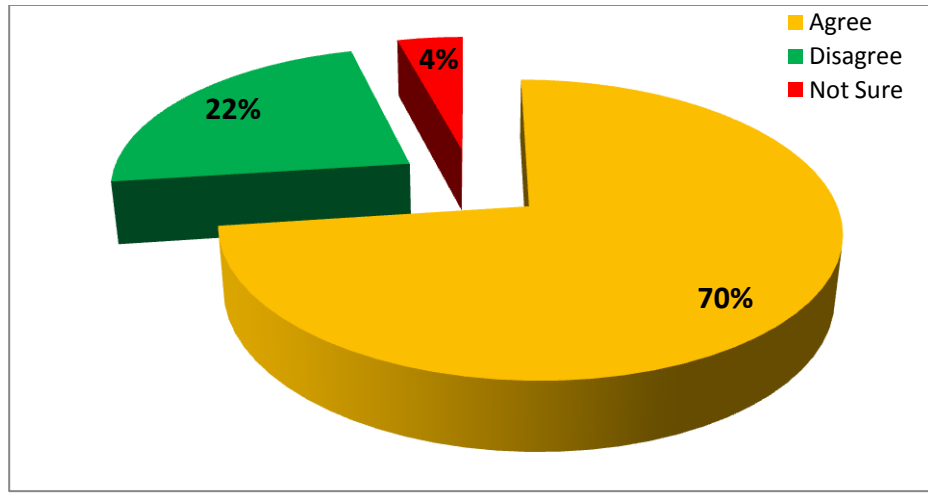


Fig. 4.11 Pie chart showing the potential of Big Data in providing insights

Table 4.17 and the corresponding figure show the potential of Big Data in providing insights to enterprise businesses. 70% of respondents agree that relevant and reliable benefits can be derived from the capture, processing and analysis of their

increasing data volumes. Where these are not effectively harnessed, it can potentially impact negatively on organizational performance.

Table 4.12: What Volume or Capacity of Big Data is generated by your enterprise organization activities monthly?

Q25: Capacity of Big Data

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid Gigabytes	19	38.0	38.0	38.0
Terabytes	22	44.0	44.0	82.0
Petabytes	9	18.0	18.0	100.0
Total	50	100.0	100.0	

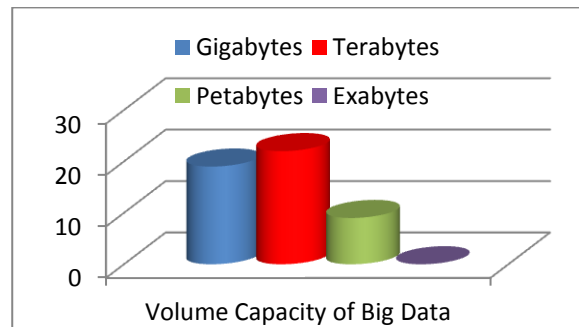


Fig. 4.12 Bar chart indicating the volume of data generated monthly within the organizations

Table 4.18 and the corresponding figure show the volume of the data that is generated within the sampled organizations. An interesting indication of this finding shows that the volume of data lies in the range of terabytes in most organizations.

4.2 Hypothesis Testing

Hypothesis 1: Relationship between enterprise business performance and big data

Null hypothesis (Ho) states that there is no relationship between the performance of an enterprise and the Big Data it generates.

Alternative hypothesis (H1) states that there is a relationship between the performance of an enterprise organization and the Big Data it generates.

The calculation of the Pearson correlation coefficient between the impact of Big Data and Enterprise business performance gives a residual value of (0.241 > 0.001). This is greater than the acceptable recommended value of 0.001 and is therefore an indication that there is a relationship between the performance of an enterprise organization and the Big Data that it generates.

Therefore, hypothesis (Ho) is rejected and (H1) accepted.

The summary of the results is contained in the tables below:



Regression Analysis between Big Data & Enterprise Performance

Descriptive Statistics

	Mean	Std. Deviation	N
Significant Impact of Big Data	1.36	.776	50
Enterprise business performance	1.80	.728	50

Correlations

		Significant Impact of Big Data	Q25: Enterprise business performance
Pearson Correlation	Significant Impact of Big Data	1.000	.780
	Enterprise business performance	.780	1.000
Sig. (1-tailed)	Significant Impact of Big Data	.	.000
	Enterprise business performance	.000	.
N	Significant Impact of Big Data	50	50
	Enterprise business performance	50	50

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	17.945	1	17.945	74.411	.000 ^b
	Residual	11.575	48	.241		
	Total	29.520	49			

a. Independent variable: Significant Impact of Big Data

b. Predictors: (Constant) Enterprise business performance.

Residuals Statistics

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	.70	2.36	1.36	.605	50
Residual	-.526	.643	.000	.486	50
Std. Predicted Value	-1.098	1.647	.000	1.000	50
Std. Residual	-1.071	1.310	.000	.990	50

a. Dependent Variable: Significant Impact of Big Data

HYPOTHESIS II: Hypothesis 2: Relationship between enterprise business performance and Big Data across organizations within two industrial sectors (Finance and Retail).

Null hypothesis (Ho) states that the relationship between the performance of a business enterprise and its generated Big Data is not similar across businesses in Finance and Retail.

Alternative hypothesis (H1) states that the relationship between the performance of a business enterprise and its generated Big Data is similar across businesses in Finance and Retail.

The computation of the Pearson correlation coefficient indicating the performance of a business enterprise across two industrial sectors (Finance and Retail) gives a value of 0.469 which is significant at the 0.01 level (2-tailed).

This is reflective of the type of impact Big Data has on an organization regardless of location. This result therefore leads to the rejection of the Null hypothesis (Ho) and the acceptance of the alternative hypothesis (H1) indicating that the relationship is similar across locations.

Regression between Significant Impact of Big Data and the Industrial Sector

Correlations

		Industrial Sector	Significant Impact of Big Data
Industrial Sector	Pearson Correlation	1	.469**
	Sig. (2-tailed)		.001
	N	50	50
Significant Impact of Big Data	Pearson Correlation	.469**	1
	Sig. (2-tailed)	.001	
	N	50	50

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



Nonparametric Correlations

Correlations

		Industrial Sector	Significant Impact of Big Data
Kendall's tau_b	Correlation Coefficient	1.000	.469**
	Sig. (2-tailed)	.	.001
	N	50	50
	Correlation Coefficient	.469**	1.000
	Sig. (2-tailed)	.001	.
	N	50	50
Spearman's rho	Correlation Coefficient	1.000	.469**
	Sig. (2-tailed)	.	.001
	N	50	50
	Correlation Coefficient	.469**	1.000
	Sig. (2-tailed)	.001	.
	N	50	50

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

Regression between Significant Impacts of Big Data in Sampled Industry (Finance and Retail) Independently.

Correlations

		Significant Impact of Big Data	Finance	Retail
Pearson Correlation	Significant Impact of Big Data	1.000	.	.
	Finance	.	1.000	.
	Retail	.	.	1.000
Sig. (1-tailed)	Significant Impact of Big Data	.	.000	.000
	Finance	.000	.	.000
	Retail	.000	.000	.
N	Significant Impact of Big Data	25	25	25
	Finance	25	25	25
	Retail	25	25	25

a. Independent variable: Significant Impact of Big Data

b. Predictors: (Constant), Finance & Retail.

5. DISCUSSION OF FINDINGS

This research used statistical analysis to provide answers to the two hypothetical questions. The first question looks at the relationship between the enterprise performance of a business or organization and the Big Data it generates. It sought to know what the strength of the relationship is and how Big Data impacts enterprises. These answers were sought by conducting a survey and getting relevant data from hundred (100) institutions regarded as big-sized, viable and able to provide insight on incoming data and how they get processed by the organization. The second question investigates Big Data impact based on the industrial sector of the enterprise. The main focus was whether or not sectorial domain impacts significantly on how enterprises cope with the challenges of Big Data and more specifically what these mean for overall enterprise performance. These two hypothetical questions were tested using Pearson correlation coefficient as well as regression analysis. The findings of the first hypothetical question show that Big Data significantly impact organizational enterprise performances. The resulting Pearson correlations coefficient residual value of 0.241 shows the high degree of impact.

The result of the second test on the hypothetical question that sought to find out Big Data impact as a result of sectorial differences also provide a Pearson correlation coefficient residual value of 0.469. As described during the analysis of the individual survey questions, it can be seen that this plays a role in Big Data governance and management.

6. RECOMMENDATIONS

From the interpretations of the survey findings above, organizations need to focus on the growing importance of Big Data. With good understanding of the sources, possible format and the frequency of data that is generated many benefits can be derived. Though the marketing hype around big data coupled with the pace of studies, reporting analysis and articles on the subject can be mind-boggling still businesses can take advantage of Big Data analytics. Doing this can help in separating fact from fiction to determine real use cases for their business. As more use cases come to light on Big Data over the coming years and there is continuous maturity of technology, Big Data will reach a critical mass. Based on this, research undertaken in this area in future studies need to focus on the trend of Big Data. It is our hope that, eventually, Big Data will become another mechanism in



the Business Intelligence (BI) ecosystem within the African business and organizational community.

7. CONCLUSION

This study presented the concept of Big Data and the impact on enterprise organizational performances. The research sought to determine the nature of the impact, challenges or opportunities that can be derived from growing data sets which are in simple term the meaning of Big Data. The results from the analysis suggest that growing data set impacts over 80% of surveyed businesses. Also, it was found that the relationship and impact on businesses are similar across geographical locations. Furthermore, the research provided recommendations on what future efforts in this regard can focus on. Doing this will help to take advantage of the potentials of Big Data analytics.

8. REFERENCES

- [1] Beyer, M. A. and Laney, D. (2012). “The Importance of ‘Big Data’: A definition,” Gartner,
- [2] Jozef Stefan Institute, Ljubljana, Slovenia Stavanger, May 8th 2012
- [3] Moon, H. and Cho, H. S. (2012). “Risk based Policy in Big Data age: Case Study of Privacy Invasion”. *Informatization Policy*.19(4) pp63-82. Tech. Rep. hBp://www.cs.uiuc.edu/homes/czhaiczhai@illinois.edu
- [4] Marko Grobelnik, marko.grobelnik@ijs.si
- [5] Zeleny, M. (1987). “Management Support Systems: Towards Integrated Knowledge Management”. *Human Systems Management*, 7(1): 59–70. 1029.
- [6] <http://ibmsystemsmag.com/power/businessstrategy/BI-and-Analytics/data-big-picture/>
- [7] <http://www.forbes.com/sites/forbesinsights/2016/02/02/h-as-big-data-taken-the-human-out-of-human-behavior/>
- [8] <https://prezi.com/hyavodg6h0nb/big-data/>
- [9] https://www.researchgate.net/profile/Abdul_Razak_T/publication/279192348_Rajakumar1/links/558d8ad608ae47a3490bccb0.pdf