2021

Value of application of insights from Big Data Analytics on transformation of agriculture: case of farmers subscribed to Mkulima Techie Kenya.

Okumu, Millicent
Strathmore Business School
Strathmore University

VALUE OF APPLICATION OF INSIGHTS FROM BIG DATA ANALYTICS ON TRANSFORMATION OF AGRICULTURE: CASE OF FARMERS SUBSCRIBED TO MKULIMA TECHIE KENYA

MILLICENT OKUMU

MMA 102509

DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF MANAGEMENT IN AGRIBUSINESS AT STRATHMORE BUSINESS SCHOOL

NOVEMBER 2021

DECLARATION

I declare that this research dissertation has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Millicent Okumu

Reg No: **102509**

Signature :   Date: 01/11/2021

Approval

This dissertation was reviewed and approved by:

Dr. John Olukuru

Head of Data Science and Analytics

Lecturer, Strathmore University

Signature_____________________   Date______________________

ABSTRACT

The application of big data analytics in agriculture is a revolution that has the potential to transform agriculture from being process driven to being data driven. Transformation of agriculture is defined as a process that involves gradual shifts in the production, from a traditional concept to a modern one resulting in change from a subsistence oriented monoculture system to a diversified and market oriented production system. The objectives of the study were to demonstrate the application of Big Data Analytics, determine the effect of usage of insights from Big Data Analytics and to determine the barriers in application of Big Data Analytics within a smallholder farmer setting. The study applied a mixed methods approach to investigate the value big data analytics has on transformation of agriculture. Data was collected using structured questionnaires from 282 respondents and key informant interviews. The inferential analysis tools applied included correlation analysis and logistic regression analysis. From this study, the findings reveal that there was a significant association between insights from big data analytics and value derived which facilitated transformation of agriculture in terms of income and yield for farmers. For the practitioners, researchers and policy makers in the agricultural industry this study provide guidelines to mitigate challenges in implementation and contributing to the broader discussion on the opportunities provided by a data driven agricultural industry. The limitations also help uncover future research courses in order to achieve better knowledge.

Key words: *Big Data Analytics, Value, Agricultural transformation, Smallholder farmer*

ACKNOWLEDGEMENT

My immense appreciation goes to my family for their continuous support and encouragement to compete my MMA.

My most profound gratitude goes to my supervisor, Dr. John Olukuru for the guidance and support that made it possible to complete the research work with good insights and reviews, knowledge impact ad encouragement at every step. Your guidance throughout this process will forever remain with me.

I extend my sincere gratitude to my classmates and friends for their encouragement and valuable insight during the entire study period.

To God, this would not have been possible without His grace, mercy and favor to me.

DEDICATION

I dedicate this paper to my family and to God. You have always been there in every step of my life.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AGRA	Alliance for a Green Revolution in Africa
AGRITECH	Technology in Agriculture
ASTGS	Agricultural Sector Transformation and Growth Strategy
CTA	Technical Centre for Agricultural and Rural Cooperation
FAO	Food and Agriculture Organization of the United Nations
GDP	Gross Domestic Product
GPS	Global Positioning System
GSMA	Global System for Mobile Communications
IoT	Internet of Things
MTN	Mobile Telephone Network
NGO	Non-Governmental Organizations
SMS	Short Message Services

CHAPTER ONE

INTRODUCTION

1.1 Background

Agriculture has gone through a several revolutions that have driven efficiency, production growth, yield and increase in profits that were unattainable previously. Agricultural transformation in most areas of the world has been an important component of the broader economic transformation process (Jayne & Ameyaw, 2016). Revolutions in agriculture kick-started economic transformations in Europe, North America, South America and Asia and were facilitated by factors including technological advancement, increase in technical skills and even consumption patterns (AGRA, 2018). The agricultural industry in terms of development has been undergoing transformation from a traditional pattern to a modern one.

In Africa ,agriculture remains as a fundamental part of the economy and accounts for just over 60% of jobs across the continent and about a quarter of the continent's GDP (African Development Bank, 2016). Smallholder production systems rather than large farms dominate the agricultural industry in Africa and this is evidenced from numerous household surveys that support the idea that the median size of a crop farm is between one and two hectares in most countries in Africa (Gollin, 2014).

While agriculture is the backbone of most economies in Africa, the industry remains underdeveloped and is characterized by low production growth and a weak manufacturing sector with a rise in food imports and stagnating exports. The low growth in production can be attributed to the use of inappropriate technology, inaccessible farm inputs, weak extension support services, poor access to finance and credit, low value addition in production and poor development in marketing support mechanism for farmers (Brussels Development Briefing, 2013). Africa is increasingly becoming more dependent on imported staple food with an import bill estimated at 35 billion US dollars annually. There's a consensus for a need of agricultural transformation in Africa (FAO, 2017).

As it is in many countries in Africa, agriculture remains vital to the growth of the Kenyan economy and is key to the country's food security and poverty reduction efforts (D'Alessandro et al., 2015). The sector is pivotal for the country to achieve the formidable goals established in the government's Vision 2030. Agriculture accounts for about 51% of GDP, 26% directly and 25% indirectly through its linkage with other sectors. It is also responsible for most of the country's export and accounts for up to 65% merchandise exports in 2017(World Bank, 2019).

1.1.1Transformation of agriculture

According to Timmer(1988), the process of agricultural transformation evolves through at least four phases that are roughly definable. The process begins when productivity per worker rises. The increase in productivity creates a surplus, which marks the second phase, and can be tapped directly, through taxation or indirectly through interventions by the government into rural-urban terms of trade. The surplus is then utilized in developing the sectors that are non-agricultural. The third phase is marked by the progressive integration of the agricultural sector into the macro economy through improved infrastructure. With the success of this phase, the fourth is barely noticeable. When integration is not successfully accomplished, most governments encounter serious problems of resource allocation and even problems beyond their borders because of pervasive attempts by high-income countries to protect their farmers from foreign competition.

Divanbeigi & Saliola, (2016) argue that agricultural transformation is shaped by three interrelated processes, first higher yields and lower costs from the existing and new farming lands increase agricultural productivity. Second, the types of agricultural products change from subsistence to cash crop, from staple to intermediate inputs and from low value/low risk to high value/high risk varieties and this change is reflected in the evolution of agricultural commodities in global markets. In the third process, agricultural market transactions become more integrated with the rest of the economy, more dependent on finance and more oriented on international trade.

According to Boettiger, Sara, Denis, Nicolas & Sanghvi (2017), the dynamics of agricultural transformation include higher productivity in farms and greater markets. Increase in productivity means that larger markets are served and agro processing expands. Some farmers make the choice to use less time farming and take up jobs that offer better incomes or economic opportunities while others stay fully engaged in farming and adopt new technologies and even expand operations. Agriculture sector modernizes and becomes more efficient and less labour intensive. Overall these dynamics help in tracking agricultural transformation in any country.

Mukasa et al., (2017) further agrees with these views, that agricultural transformation is defined as growth in productivity and an increase in uptake of modern technologies and inputs. It also involves solving imperfections in the financial markets, land and labour that impede adoption of technology. It further involves improved efficiency and value addition in connecting farmers to markets, as well as enhanced resilience in the face of growing risk due to climatic, market and political shocks. ACET (2017) define agricultural transformation as a process that incorporates two main processes; modernizing farming by boosting productivity and running farms as modern businesses and strengthening the links between farms and other economic sectors.

FAO (2017) discusses in detail transformation process in agriculture to be associated with seven trends. These trends are as follows: First, farmers in production become more commercialized while others move out of farming to take advantage of better economic opportunities. At the second stage, the farms transition from producing a diversity of goods to becoming more specialized, taking advantage of regional comparative advantage and in the process become dependent on markets. At the third stage, the ratio of agribusiness value added to farm value added rises over time as more economic activities take place in upstream input manufacture and supply downstream trading, processing and retailing. At the fourth stage, farms begin to supply the agricultural sector to capture economies of scale in production and marketing and the mean farm size rises with the exit of people from rural farms and consequently farm consolidation. The technologies in farm production evolve to respond to changes in factor prices (land, labour and capital) as a country develops at the fifth stage and there is transition from shifting cultivation to focus on intensive, sustainable and management-intensive cultivation of specific fields. This marks the sixth stage. At the final stage the agro-food system becomes more integrated into the wider economy.

Generally, agricultural transformation has a set of components which involve gradual shifts in the production, from a traditional concept to a modern one resulting in change from a subsistence oriented monoculture system to a diversified and market oriented production system (Yangfen et al., 2010). Agricultural transformation has further been defined as a process through which farms gradually move from highly diversified, subsistence oriented systems of production to more specialized and business oriented production processes (Mukasa et al., 2017). Heads of states in Africa through the Malabo Declaration as well as Agenda 2063 have a vision with regards to what transformed agriculture in Africa should look like (FAO, 2017). The vision would see the move from the current low total factor productivity to a high income industrial state. Agricultural sector will then be modernized and will continue to produce food at a low cost (African Union Commision, 2014). Africa has agricultural advantages such as plenty of natural resources, a young and growing labour force and a surge in urbanization. The continent also has an estimated six hundred million hectares of arable land that has not been cultivated (Boateng, 2017).

Agricultural transformation is by no means a short process and it is highly unlikely that one can observe the process full in the course of this research (Agricultural Technology Adoption Initiative, 2016). To address this, the study identifies key metrics that can indicate whether agricultural transformation in being achieved.

Agricultural productivity can be used as a measure of agricultural transformation (Boettiger, Sara; Denis, Nicolas; Sanghvi, 2017). Agricultural productivity is defined as the ratio of outputs to inputs and larger values of this ratio are associated with better performance. Through measurement of agricultural productivity, farm incomes can be assessed more accurately. The relationship between agricultural productivity and agricultural transformation is at the centre of many debates (GSARS, 2017). The Malabo Declaration (African Union Commision, 2014) for example places agricultural productivity growth at the centre of Africa's objective to achieve agricultural –led growth .

The significance of these two concepts stem the fact that they form a basis for the scope and define the focus of the paper. This study shall adopt the definition of agricultural transformation according to Divanbeigi & Saliola, (2016) specifically the first stage of agricultural transformation that is achieved by higher yields and lower costs of production. Therefore, productivity will be used as a measure of transformation. Productivity in agriculture is defined as the ratio of measure of output to a volume measure of input. A measure of productivity that is often cited in agriculture is agricultural output (GSARS, 2017). It is an ideal measure in this study because it involves collecting raw data on the physical quantities produced from a farm and farm gate prices or a proxy, to value production.

1.1.2 Agricultural Transformation in Kenya

The performance of the agricultural sector has been erratic with productivity of foods falling rapidly in relation to the growing demand and this has resulted in rising import bills to fill up the food deficit. Furthermore, climate change is increasingly a threat to agricultural output with negative implications to food security (D'Alessandro et al., 2015).

The government of Kenya, having acknowledged the importance of agriculture to the economy, has formulated the Agricultural Sector Transformation and Growth Strategy (ASTGS) with a view to transform the agricultural sector in order to make it a regional powerhouse. ASTGS is anchored on three outcomes which includes increasing small-scale farmer incomes, increasing agricultural output and value addition and boosting household food resilience. ASTGS have six flagships that sit under these anchors and three enablers (Ministry of Agriculture, Livestock, 2019).The transformation is expected to create a vibrant, commercial and modern agricultural sector that would contribute to the country's sustainable economic development.

Leveraging on modern technology has been identified as one of the tools that can be applied to spin off a wide range of agricultural applications in order to provide weather updates, access to finance for farmers and driving logistical efficiencies for input suppliers and buyers(World Bank, 2019).Digitization in agriculture refers to the use of digital technologies, data or innovative business models to transform practices in the agricultural value chain and address the challenges in agricultural production, post-harvest handling, access to the market, financing and management in the supply chain with the aim of improving income for farmers especially the smallholder farmers, improve food security as well as mitigate effects of agriculture to the environment as a result of intensification (CTA, 2019). Digitization of agriculture will transform agriculture and cause a substantial change in production of food and farming over the coming years and as such the ASTGS document includes digitization of agriculture as an enabler to the transformation of agriculture (Ministry of Agriculture, Livestock, 2019).

1.1.3 Trends in Digital Transformation of Agriculture

The basis of the current digital transformation was laid down almost two decades ago along with the use of first precision agriculture technologies on farms. Precision farming now comprises of GPS for yield mapping (Kosior, 2018).In the past decade implementation of technology has given farmers the opportunity to increase their yield by maximizing their output and automating input through use of technologies in agriculture. Classical extensive agriculture was replaced by precision agriculture. Systems of planting and cultivating crops changed from being manual and labour intensive to procedures that use robots and satellites. Instead of ploughing, sowing and fertilizing manually it has become possible for farmers to calculate the number of resources necessary for each crop in the field.

Internet of Things has simplified and streamlined the collection and distribution of resources using sensors on equipment. Sensors spread strategically in fields together with image recognition technologies allow farmers to view crops from anywhere in the world and get up to date information in real time (PA Consulting, 2017). Drones and are used to spray chemicals and robots used to weed. Robots are also used for transplanting, adding a new level of efficiency to the traditional methods while others are used for harvesting.

Through Machine learning predictions on the traits and genes that will be best for crop production is done therefore giving farmers the best breed for their climate and location. Machine learning algorithms is also used where consumers are purchasing products and show which products are being purchased the most and which ones are falling thus creating proficient and effective forecasting for future farming.

Collectively these technologies are generating an enormous amount of data from a wide range (Delgado et al., 2019). Firms have established themselves as data holders and data specialist in the emerging digitization of agriculture. The firms range from start-ups to dominant players in machinery, fertilizers, chemicals and seeds such as AGCO, John Deere, Monsanto, Sygenta among others and use the data they have to understand the value of off-farm data, harnessing its power through big data analytics(Pham & Stack, 2018). Big data and big data analytics is a strong catalyst for increase in productivity, improved income for farmers and ensuring that technologies help farmers become more efficient and sustainable thus agricultural transformation (Maru et al., 2018).

1.1.4 Big Data Characteristics

Bekmamedova and Shanks (2014) define big data as data storage, management, analysis and visualization of large and complex sets of data and focuses on data management techniques that supersede traditional relational systems and are better suited to manage large volumes of data. Further, (Chi et al., 2016) characterize data in five “V” dimensions: volume, velocity, variety, veracity and valorisation. Volume refers to the size of the data collected. Sources of data such as connected smartphones, sensors and other devices, combined with rapidly developing information and communication technologies have contributed to the tremendous generation of data. Velocity is characterized by the high speed of data generation and the timeline within which the data remains useful. Variety means having data from multiple sources, with different formats, from different disciplines and several application domains.In agriculture, there’s no standard way by which data is captured and that results to difficulty in harmonization and compilation of data from the various sources difficult. Veracity refers to the quality of the data in terms of accuracy, reliability and confidence while valorisation refers to the ability to disseminate the knowledge. Rodriguez, de Voil, Rufino, Odendo, and van Wijk (2017) argue that although big data is described using the five "V"s, big data analysis does not need to satisfy all the five dimensions. Wolfert, Ge, Verdouw, and Bogaardt (2017) , explain that despite the characterization of Big data by high volume, velocity, veracity, valorisation and variety big data is not about the volume rather about the capacity to search, aggregate and cross reference large sets of data in reasonable time and capability to extract information and insights where there were challenges previously, either technically or economically to achieve this. Stubbs (2016) further suggests that the term big data as it is applied in agriculture is less about the size of the data and more about the combination of technology and advanced analytics that creates a new way of processing information in a way that is more useful and timely.

1.1.5 Big Data Analytics in Agriculture , It's Potential Use and Transfromation of Agriculture

Lamba and Dubey (2015)defined big data analytics as application of multiple analytic methods that address the diversity of big data to provide actionable descriptive, predictive and prescriptive results.

The descriptive, predictive and prescriptive aspects of big data analytics hold the potential to make changes possible so that aggregation of data fragments from farmer's profile, farm satellites and weather can be used to bring together resources and service providers to support farmers and agriculture in general. Descriptive analytics gives insight into the past and provides answers on what happened. Predictive analytics involves forecasting techniques looks at what can happen in the future, therefore giving answers on what could happen. Prescriptive analytics uses optimization and simulation algorithms to offer recommendations based on possible outcomes and gives answers on what should be done (IBM, 2017).All aspects of big data analytics work together to improve efficiency and minimize loses (FAO, 2019). Big Data analytics in agriculture would thus be useful in the following ways in agriculture: surveillance , extension services, supply chain management, access to market and financial services (CTA, 2019).

Insights from Big data analytics provide new and useful data for decision making, farm management and solving problems by use of extension services. The new abundance of information from farms provided by the different technologies could improve the ability of farmers to make profit maximizing decisions. However, pooling the datasets of hundreds or thousands of farms has a much greater potential in terms of value to both the individual farmers and the agricultural industry as a whole. Agricultural big data that has been combined into an aggregate form has the potential to reveal more insights (Coble et al., 2018). Compilation of farm level data from many farms within a specific geographical region, which in itself is big data, is likely to yield insights from which new management practices could be developed(Lassoued et al., 2021).

Initiatives such as Esoko, ICow, M-Farm, Twiga Foods and Digifarm in Kenya and Farmerline in Ghana, MLouma in Senegal, FarmCrowdy in Nigeria, Troto Tractor are examples of some of the agri tech companies and start-ups using mobile application and services in Africa to give services to farmers. The initiatives are making use big data by aggregating farm related data from different sources in order to derive insights from the datasets so as to transform agriculture among smallholder farmers by enabling them to produce optimally and make farming a profitable business (GSMA, 2018).

Table 1.1 Agricultural Platforms that use Big Data Analytics

Platform Name	**Usage**
Icow	The objective of iCow is to increase the productivity of farmers by giving them access to knowledge and expertise. Since its launch in 2011, iCow has enrolled an average of 15,000 active farmers
Esoko	This platform keeps farmers informed about market prices, weather forecast, extension advise through customized messages and has recently ventured into additional services financial services such as digital credit, insurance and payment and has reached 1 million farmers across Africa.
M-Shamba	This is an interactive platform that provides information to farmers through the use of a mobile phone using SMS to provide information on production, harvesting, marketing, credit, climate and weather.
iShamba	iShamba is a call centre of agricultural experts where farmers can send questions via SMS or call in to speak to an expert for instant help. Farmers that subscribe to the services receive agri tips on crop and livestock, market prices and weather updates.

Farmis	The platform provides market information on fresh produce to farmers in the value chain. In addition to accurate and timely market information, FARMIS also offers weather and crop advisory services that can be tailored to a farmer's needs, considerably reducing the cost of extension services
We-farm	This is a free peer-to-peer service that enables farmers to share information through SMS, without the internet and without having to leave their farm. Through We-Farm, farmers can ask questions on a wide variety of subjects and receive crowd-sourced responses from other farmers.

The challenges in agriculture are linked and many models in the market are seen to provide services for particular sections of the agricultural value chain. However, in order to maximize the returns from agriculture, problems need to be solved in an integrated manner.

1.1.6 Mkulima Techie Platform

Mkulima Techie in Kenya is a unique example of a platform that attempts to bring in an integrated approach to provide services across the value chain using big data analytics (Sawant et al., 2016). The platform offers farmers a one stop access to quality farm inputs at discounted prices, input loans and extension services. Additionally, the platform provides yield cover insurance. The platform has about 1,038,817 registered users. Of the registered users, 42,000 are active users making it one of the most used platforms by smallholder farmers. This can be attributed to the fact that it is a one stop service provider, unlike other platforms, as it provides access to input, finances and information. The platform applies predictive, prescriptive and descriptive aspects of big data to offer services. Predictive analytics is applied in supply chain management and this has helped to get agricultural input to the destination faster and more cost-effectively. The predictive and prescriptive aspect of Big data analytics has also been used by the platform for credit risk assessment as well as for insurance pay outs. Most of the other platforms offer one service. E-Soko for example is a platform that links farmers to the market, M-Shamba provides information on production, harvesting, weather and where to access market or credit. I Shamba provides extension services where farmers can send SMS or call to ask questions and get to speak to an expert for instant help. These examples show the uniqueness of the Mkulima Techie platform that provides multiple solutions and hence the reason why this study focuses on the platform as a case study.

1.2 Problem Statement

The concept of extracting value from these huge data sets in order to unlock the transformative potential of Big Data analytics is limited (CGIAR, 2016). The use of Big Data analytics stands poised to be the next great revolution of agriculture, with the potential to help producers of all sizes and enterprises produce more with less (Boettiger, Sara; Denis, Nicolas; Sanghvi, 2017). With the rapid developments in technology that have taken place, traditional practices in agriculture are still being followed in several countries especially the developing countries. Use of insights from big data analytics in agriculture fundamentally diminishes the likelihood of crop failure and may help to optimize farm production processes, improve risk management and enhance strategic decision making capabilities(Arjun et al., 2016). Theoretically, the application of insights from big data analytics in influencing decision making across the entire agriculture value chain can potentially lead to greater efficiencies and increased income. However actually knowing the value of these benefits from the use of these insights from big data analytics is lacking(Kosior, 2018). The agricultural sector is well placed to leverage on big data analytics. Many institutions in the agricultural sector are exposed to large amounts of data collected on input supply, agricultural finance, insurance, on-farm production, crop aggregation, long-distance transport, processing, retailing among others(Omo-Ojugo, 2018).With the continuous data streams and ever changing information about climate, crops, farm equipment and other condition monitoring sensors, that is approximated at 30,000 gigabytes of data per second, big data analytics has the potential to pull all applicable data from multiple sources to develop insights useful for transformation of agriculture (Addison & Msengezi, 2018).

While several agro-tech companies provide solutions in agriculture based on insights from big data analytics the key challenge is an illustration of its true value that gives accurate and actionable insights to farmers. Most companies have developed a push perspective where the functionality of the system is shown but the value to the end user is not shown. The existing gap in big data analytics is the realization of the value of big data that is obtained from actionable information from these growing piles of data that farmers can use to boost their productivity (Ribarics, 2016).

The purpose of this study, therefore, is to establish a connection between big data analytics and the value obtained from a practical use of big data analytics. This was done through looking at the case of farmers subscribed to an identified agritech firm in order to establish the extent to which both the firms and farmers subscribed to their services were able to derive value from Big Data Analytics and how that has influenced transformation in agriculture.

1.3 General Objective

The study aims to establish the value of big data analytics on transformation of agriculture.

Specific Objectives

i. To demonstrate the current application of big data analytics in agriculture on the *Mkulima Techie* Platform
ii. To determine the effect of usage of big data analytics on transformation of agriculture on the *Mkulima Techie* Platform
iii. To determine the barriers of application of Big Data Analytics on transformation of Agriculture on the *Mkulima Techie* platform

1.4 Research Questions

i. In what ways is Big Data Analytics currently being used on the *Mkulima Techie* platform?
ii. In what ways has the application of Big Data Analytics facilitated the transformation of agriculture on the *Mkulima Techie* platform?
iii. What are the limiting factors that hinder the realization of the full potential of Big Data Analytics on the *Mkulima Techie* platform?

1.5 Scope of the Study

This study sought to investigate the value of big data analytics on transformation of agriculture using a case of farmers subscribed to Mkulima Techie Platform in Kenya and key informants who are members of staff of the Mkulima Techie platform. While Big Data Analytics in Agriculture enhances collaboration among different nodes in the agricultural value chain, this study is limited to farmers and members of staff of the Mkulima Techie platform so as to understand how they interact with Big Data to reach specific outcomes related to the farms and the organization. With a target population of 42,000 farmers subscribed to this platform and are active users, a sample of 396 formed the unit of analysis of the study.

The choice of Mkulima Techie platform was based on the fact that the platform has the highest subscription of the many digital platforms available and it combines multiple solutions across the value chain, applying insights from Big Data Analytics, which brings out its uniqueness among the other service providers. The study incorporated seven indicators that measured transformation in agriculture. These were access to input, financial access, advisory and information services, farm management, decision making, farm problem solving and barriers in application of big data analytics. In as much as transformation involves the entire agricultural ecosystem the study focuses on farmers that are subscribed to the platform and members of staff of the platform in order to give an in depth understanding of the application of insights from big data from a production perspective. The study adopted descriptive research design and used descriptive and inferential statistics for data analysis.

1.6 Significance of the study

This study is relevant in many aspects as data in agriculture is needed by stakeholders in this sector ranging from those who make decisions in the government to players in the private sector. To the policy makers, the study aims to guide policy formulation for the creation of an enabling environment where an effective use of big data analytics can be applied... Decision makers such as managers are beginning to view the quality of data products and appreciate data that is accurate, consistent, timely and complete to make decisions. Beyond the farms, there's an interaction with infrastructure, markets, policy and more and with the multiplication of the number of users the data ecosystem for food security continues to become complex. The finding of this research could bring together stakeholders from weather and climate change, insurance and finance, land identification to see how much an integration of these different datasets put together could transform agriculture. Academicians and scholars would use this information to conduct future studies as it will build to the body of knowledge in the Kenyan context

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter presents a review of literature and examines both theoretical and empirical literature. The chapter also includes a conceptual framework and identifies the research gaps in the existing literature as well as operationalization of variables.

2.2 Theoretical review

This section brings forth relevant theories that show the association between big data and its role in transformation of agriculture.

2.2.1Agriculture Transformation Theory of Schultz

Agriculture transformation theory of Schultz coined in 1964 stressed that the key to transformation lied in emphasizing technological change in agriculture (Wharton, 2019). According to Schultz, technology in agriculture would be described as traditional if it remained unchanged for a long time. This state, when agriculture is no longer progressive comes about when people using various inputs under such technology have fully explored the pros and cons of using this inputs and in general the productivity of these inputs and their cost have become equal to each other. When year after year a farmer, under such circumstances, would get the same returns, they would be bound to adjust their investment in various factors in such a way to limit the marginal productivity of each factor so that it finely balanced with its price and the balance would stay as long as this method remained unchanged. He called this a state of "stationary equilibrium" with no investments and net savings. He argued, therefore, that there was need to accelerate output growth through technological interventions. He envisioned a crucial role for investment in non-traditional inputs such as knowledge and education and improvement in the quality of material inputs and people to help shift agriculture to a firmer footing and capitalize on agriculture (Pasa, 2017). He suggests three ways of increasing production, by making use of resources that have not been utilized, by allocating resources optimally so as to maximize on production and by replacing old factors with new ones with higher output-input ratios. He ruled out the first two options by urging that factors of production in traditional agriculture are fully utilized and that resources in agriculture are always perfectly allocated. He suggested that adoption of new factors of production was necessary. But even so he further suggested that the process should be demand driven and not command driven.

The role of the theory in the research study was to provide the basis from which the study's variables were conceptualized and which are part of the constructs of the framework and on which the study was anchored .In examination of the extent of the application of insights of big data in terms of non traditional input as factors of production, the theory provides guidelines such as use of resources, optimal allocation of resources and adoption of new factors of production which big data in agriculture provides and their link to transformation of agriculture.Therefore, the theory was a representation for all the variables in under study and was useful in understanding how these elements work together to achieve the desired performance, which is transformed agriculture.

2.2.2Diffusion of Innovations Theory

The diffusion of innovations theory is useful in providing an account of how technological innovation such as Big Data move from the stage of invention to widespread use or not. Diffusion of Innovation Theory developed by Everett M. Rogers in his book Diffusion of Innovations in 1962 (Orr, 2003).

The theory is characterized by four elements namely innovation, communication channel, time and social systems all of which influence the adoption or rejection of an innovation by a particular group. Innovation is defined as an idea or practice that is perceived as new by an individual or any other unit of adoption. Innovations may have been invented a long time ago but if the unit of adoption perceives it as new, it becomes innovation to them. Its characteristics, such as relative advantage, compatibility, malleability and complexity influence the likelihood that it will be adopted (Micheni, 2015). Diffusion is described as the practise of communication of an innovation through some channels over a period of time among members of a social system and hence essential for the subsequent acceptance or rejection of an innovation (Chen et al., 2015). Adoption is defined as the choice an individual makes to use of an innovation as the best course of action available. Change is measured by the number of people, groups or individuals adopting the innovation over time and is influenced by perceived attributes of innovation, type of innovation decision, communication channels, nature of social systems and the extent of change agents' promotion efforts (Micheni, 2015).

As such the Diffusion of Innovations theory is very comprehensive and its concepts provide a framework to make an analysis of the diffusion of Big Data. The theory explains the perceived features of an innovation and how communities view Big Data Analytics, how innovations are communicated and shared and the consequences of adoption in terms of the costs involved, benefits and its value and impact to communities(Micheni, 2015). The Diffusion of Innovation Theory provides a general framework and guidance on adoption and utilization of Big Data analytics in agriculture.

These theories serve as a theoretical guideline for studying factors shaping the adoption and utilization of Big Data in agriculture. The theory of Schultz emphasized that the key to transformation in agriculture lies in emphasizing change in technology in agriculture and therefore need to accelerate growth in output through technological intervention. The diffusion of innovations theory claimed that the degree of adaption or rejection of technology is dependent on five stages that an individual pass through: awareness stage, interest stage, evaluation stage, trial stage and adoption stage. The theories suggest that there must be an interdependent relationship between transformation of agriculture and agricultural technology for agriculture modernization to take place. The study therefore seeks to show the value of Big Data Analytics adoption in agriculture in terms of its impact and role in transforming agriculture and this can be a vital component on ensuring successful adoption and utilization of Big Data Analytics.

2.3 Empirical Review

The literature reviewed examines previous empirical data that evaluates how big data is applied in agriculture

2.3.1 Application of Big Data Analytics and Transfromation of Agriculture

Tatge (2016) outlines four sectors from which farmers can indirectly and directly benefit from data from their farms as seed, extension services, crop insurance and farm equipment. The article shows that seed companies consider data a key component in developing new seeds and chemicals that match controlled lab results. Seed companies would want to get data from farms and make a comparison between the maximum potential of yields from a farm and the actual results to understand the difference so as to ensure higher produce for farmers. Secondly, timely data for agronomist would allow them to make better recommendations. Data collected over years helps agronomists look at trends in specific fields, therefore helping them to get a reference point on a farm’s yield and improve them compared to the average yield in a particular area. Big data analytics would even make it possible to show an association of data on weather patterns, equipment data and other sources of data in order to come up with better yield prediction models. Thirdly, he argues that real time data can enable crop insurance companies to access a farmer's profile, which leads to better strategies in profiling risks and more accuracy in yield projections per season which translates to better management of loss reserves in farmer's balance sheets. Lastly, the equipment manufacturers can identify engine optimizations in real world use through the data collected from their newer models as well as the wear and tear on the equipment. Dealers can offer predictive maintenance, so that parts are replaced well before time to ensure there is no downtime in the middle of a busy season.

Bronson & Knezevic (2016) in their study on big data in food and agriculture reported that John Deere invited farmers to subscribe and pay in order to access information that could help them make decisions in the farm. This information was acquired from sensors that were fitted in tractors and the sensors streamed data about soil and conditions of crops. The data collected by John Deere helped farmers to decide for example, where to plant crops but the data was not openly accessible to farmers.

I.Protopop and A.Shanoyan (2016) provided an analysis of illustrative cases of smallholder oriented big data applications and gave an overview of Kilimosasa in East Africa. Kilimosasa sort to provide insurance for the smallholder farmers and this allowed them to qualify for micro-loans and ability to invest in production as farmers were now resilient to weather shocks. They further looked at AgriVAS that offered services such as agronomy information and advisory services to via mobile phones to farmers in remote areas. The cost was billed to customers and the information helped farmers to enhance their yields. AgriVAS also provided digital finance services via M-Pesa. Big Data Analytics facilitated the process of determining loan sizes accurately and the borrowing terms for farmers in smallholder setting.

Himesh et al., (2018) made an assessment on how big data was driving new revolution in agriculture and reviewed Plantix, a pest and management tool that is based on Big Data Analytics provided diagnosis and steps to mitigate diseases and helps disseminate best practice methods to reduce use of pesticides by farmers. The App also featured a library of diseases which farmers could refer to in the case that there was no connectivity as well as prescriptions for over 200 crops.

(Pham & Stack, 2018) established that DuPont through its platform called Encirca offered agronomic recommendations, fuel and fertilizer price tracking and customized weather reports. While agreeing with the findings from Bronson & Knezevic (2016), Pham & Stack(2018) further reported that data transfer from the tractors was done via a system called JDLink and an application called Mobile Farm Manager facilitated them to provide information to farmers on weather tracking and prescription services. The study further posited that Monsanto began selling a prescription service called Field script in select states that grew corn. Farmers were asked to turn in over two years of historical field data and in return they received a prescription on planting which had recommendations on seed variety and a planting program that was designed to work with Precision Planting equipment. Farmers received prescriptions sent on their phones once they reviewed and accepted the recommendations. At the beginning of the planting season, farmers downloaded the program and the planters automatically carried out the planting prescription.

Harvesting Inc. that is more generally known as Harvesting used Big Data Analytics to provide timely, cost-effective and accurate insights into crop activity of individual farms worldwide. According to (Dean, 2018), Harvesting delivered crop data to clients through an agro-lending software that was designed to reduce the information asymmetries and transaction costs of serving smallholder farmers. Data collected by Harvesting was reported to flow into a credit risk scoring system that incorporated alternative data including Harvesting's remote sensing data on crop activity, to generate credit scores for borrowers. The credit scoring improved the efficiency of the loan approval process. Once the loan was made, farmland monitoring enabled lenders to monitor a farm remotely and cost-effectively and intervene with borrowers pro-actively.

Kosior (2018) in a study on opportunities and challenges of digital transformation in Agri-Food Sector also reviewed some of the applications of Big Data analytics and found out that FarmFacts provided detailed weather forecast and predictions concerning pest infestations and fungus while Climate Fieldview offered a digital platform with tools for yield analysis and advanced crop health imaging. Further, Big Data analytics enabled dedicated platforms to provide information concerning access to inputs and finance and risk mitigation measures.

In their review of global good practices of cases that use Big Data Analytics to provide services to farmers (FAO & ITU, 2019) reported that Olam Farmer Information System(OFIS) connected farmers to the global economy through delivering mutual benefits for smallholder farmers and customers. They reported that the farm data gathered enabled them to provide more tailored support to the farmers. The personalized plans enabled them to offer advice on the exact amount of agricultural input required, the number of shade trees that needed in order to protect the crop as well as how to prune properly. If yield did not appear as expected, then

Olam compared the yield of one farmer with other farmers' performances in the region and identified what the barriers were. Olam was also exploring possibilities of linking farmers' phones to a digital wallet to create a banking ecosystem. It is through this that they would ensure financial inclusion for those farmers who had been previously overlooked by the financial system by ensuring they provided access to crop insurance, saving facilities and peer to peer lending.

The identified studies highlight the approaches of different interventions that have been put in place by different firms using big data analytics in terms of providing access to financing, access to input and access to extension services. These findings are significant in that they provide further credence in the value offered by application of big data analytics at the firm level in agriculture that can lead to transformation of the industry.

2.3.2 Effects of usage of Big Data Analytics and Transformation of Agriculture

Kshetri (2014) established that farmers were able to upload pictures and videos related to a problem they faced which were then analysed by experts who offered customized advice. The researcher reviewed academic literature, policy documents and other reports on trends in Big Data utilization in agriculture and farming activities in developing countries and provided insight into how farmers benefitted from utilizing big data. The author further reported that farmers were able to reduce waste of inputs such as fertilizer while achieving increased agricultural productivity. While agreeing with these findings, (Constantiou & Kallinikos, 2015) added that enhanced decision making and a more informed strategizing were values that could be gained from big data.

Poppe, Wolfert, Verdouw, & Renwick, (2015)in their paper that sort to answer the question of how big data exchanged between farmers and their business partners would change the nature of farming, pointed out that Big Data analytics helped farmers to achieve profitability and efficiency through cost reductions and obtaining better prices for their products. This was made possible through access to optimal decisions and through improved management control. While acknowledging that farming is experiencing a digital mutiny (Bronson & Knezevic, 2016) reported that farmers were now able to make decisions based on information from massive sets of data and analytics. Similarly, (Ribarics, 2016) in a study on big data and its impact on agriculture, reported that big data analytics could enable higher yield and reduced support costs. These findings were in conformity with those of the aforementioned researchers.

Wolfert et al., (2017) while seeking to gain insight into the big data application in farming established that big data was used to provide predictive insights in farming operations, drive real-time operational decisions on crop yield and harvesting, while helping farmers to improve profitability and efficiency as they reduce their cost of production. Himesh et al.,(2018) while in agreement with these findings added that big data analytics was set to transform traditional farming into a modern data driven one and could address issues around farm productivity and profits. They established that big data analytics helped farmers in terms of farm management that helped to reduce their cost of input and maximize on farm productivity and profitability by facilitating them to achieve high yield, high quality, efficient and ecologically sustainable agriculture even solve problems in the face of climate change and market vulnerabilities. Farmers were able to get information that helped them in farm planning, crop choice, irrigation management of inventory, equipment optimization and financial records. The information also helped farmers to make informed decisions on crop planning, irrigation, fertilizer application, disease control, harvesting and post-harvest processing and therefore reduce risks.

Kosior, (2018)conducted a study on digital transformation in the agro-food sector through reviewing literature and collection of use cases deployed in agribusiness in selected research projects and further confirmed that Big Data helped farmers to reduce cost of input, improve yield and profitability. Farmers also benefitted by having access to detailed weather forecast, predictions on pest infestations and fungus and this helped them to solve problems in the farm on time. The study revealed that farmers were supported in multiple stages of the production cycle starting from preparation of fields, planting, care of crops, harvesting, drying and monitoring and control. Further, a study of the grain industry in Australia outlined the benefits of application of big data in agriculture and considered this one of the most important developments in this industry. In their study (Jakku et al., 2018),showed that big data had the potential to transform Australian agriculture through improving decision making in the farm as well as the broader industry level, analysis and prediction that leads to increase efficiency in farm management, which in turn helped the agricultural industry to overcome challenges in productivity and sustainability. Further in their findings, they reported that farmers noted an increase in production and thus an increase in profits as well as an improvement in farm management and decision making. Along the supply chain the study further identified that Big Data enabled optimisation and improved decision making at the industry level.

According to (ICRISAT, 2019) article on how big data delivered benefits to farmers, they established that Big Data analytics provided knowledge systems that gave decision support to farmers under uncertainties like rainfall especially for those whose farming activities are rain fed. Delgado, Short, Roberts, & Vandenberg, (2019) agreed that Big Data Analytics contributed to ensuring farmers had decision making support tools that ensured increased yields and therefore increased profits as well as increased sustainability of the agricultural systems. Kamilaris, Kartakoullis, & Prenafeta-Boldú, (2017) on their review of big data analytics in agriculture found out that big data analytics helped farmers in planning, solving problems in the farm and also improved productivity and this is in keeping with that of the aforementioned researchers.

From the empirical findings on the influence of big data analytics has had in transformation of agriculture, there is an indication that there is lack of substantial evidence from a developing country, specifically on the exact benefits that big data analytics has had for farmers in the proposed area of study. This study serves to address this challenge by offering additional empirical evidence on the influence big data analytics has had in transforming agriculture in terms of farmers realizing the value of big data analytics in agriculture.

2.3.3 Barriers of Usage of Big Data Analytics and Transfromation of Agriculture

It is noted that use of Big Data is more rampant in the developed countries compared to developing countries(Kshetri, 2014)**.**The researcher outlined several challenges faced by both the farmer and the firms in terms of use of big data analytics. The researcher found out that for the smallholder farmers' challenges included the affordability of specialized machines, specialised seeds and computers or tablets for farming activities. The researcher also pointed out that there were challenges on accuracy of the insights given and whether it was actionable, with the fear that smallholder farmers often could not access the data and had no means of interpreting it especially in the case of dissemination of climate data for instance. The researcher further established that lack of human resource and expertise posed a major barrier in the implementation of Big Data projects. European Union economies also reported a shortage in man-power in data analysts. Lack of infrastructure to collect information, he argued, was another major barrier that hampered the use of Big Data. On matters related to potential misuse of information at the firm level that affected farming activities, the researcher established that there were concerns that seed companies prescriptive planting programs had vested interest in higher crop yields associated with use of Big Data use in the case where the firm might have influenced farmers to buy specific seeds, sprays in order to profit from the cost of their service.

Stubbs(2016) in a study on big data in the United States Agriculture outlined several challenges in the use of big data including resources because the analysis of big data is complex, cumbersome and resource intensive. The process of collection, management and analysis was reported to be costly. Arguments were made on the greater returns on investment. Data ownership according to this study was also a major challenge that had been widely discussed. The study suggested that the producer from which data is collected should be the owner. The answer on ownership of data however generally lies on who owns and control its value but the study assessed that the debate still leaves a lot of questions un-answered. Issues on data privacy also come up as a challenge. While most concerns are in respect to competition in the market as information related to yield and performance are of incredible value. Technology system failure and technology adoption also come up as challenges. Basis failures in the system continued to bring challenges in big data and in some cases wrong data or data that was inaccurate that lead to poor decisions. In terms of adoption, while some producers were open to adopting technology in their operations, the aging producer population were noted to be slow in adopting technology. Furthermore, the author reported that an unbalance in the information could distort the market place for both input markets by creating market advantage for some input companies and the commodity market as well. Challenges on the affordability of big data analytics products was also reported even with the anticipation of higher returns.

Günther, Rezazade Mehrizi, Huysman, & Feldberg, (2017) in their findings noted that organizations could be hesitant in sharing or exchanging data with network partners due to privacy or security concerns or when the data analytics was considered a source of competition and sharing it would put at risk an organization's uniqueness. This explained why some organizations controlled access to data. Furthermore, the authors establish that collecting and synthesizing data from many different contexts involving many different stakeholders is a complex process and it is difficult to ensure that data collected is of high quality. Lack of quality, they note, has a severe consequence for organizations especially if the data translates into faulty actions.

Wolfert et al., (2017) found in literature and categorized challenges into technical and organizational ones. The authors suggest that technical challenges are solvable if enough business opportunities are realized so that there is a return on investment. However, in terms of revenue, especially in developing countries there is a challenge in making big data based solutions affordable for farmers. Data privacy and security, as is in other findings comes up in this study as challenges. Additionally, data availability and quality are also noted in this research as challenges. Kamilaris et al., (2017) further observed that barriers such as limited reliability, variety and time relevance were discussed in the papers their study looked at.

2.4 Research Gaps

The Agricultural Transformation Theory of Schultz and the Diffusion of innovation theory affords researchers, policy makers and policy implementers guidelines on the potential benefits and the challenges in diffusion of insights from big data analytics in agriculture. Empirically, most of the studies and articles cut across the different ways in which big data analytics has been used in agriculture. There has been a lot of research assessing application of insights from big data analytics in transformation agriculture in developed countries. Currently the Big Data applications discussed are taking place in Europe and America with only few examples in Africa There is relatively limited attention to Big Data Analytics-based solutions and the actual value derived that is focused on smallholder farms in developing countries. Although findings from most of the studies undertaken indicate positive responses to investment in Big Data in agricultural system, the studies done in a smallholder farmers setting are too few to give a conclusive result. There is need for evidence about the effect on the ground so that the contribution of big data analytics in agriculture can be clear. Motivated by this gap, this study therefore seeks to establish the value that big data analytics has had in the transformation of Agriculture in a smallholder farmer setting.

Ha: There is a positive significant relationship between value derived from applications of big data and transformation of agriculture.

2.5 Conceptual Framework

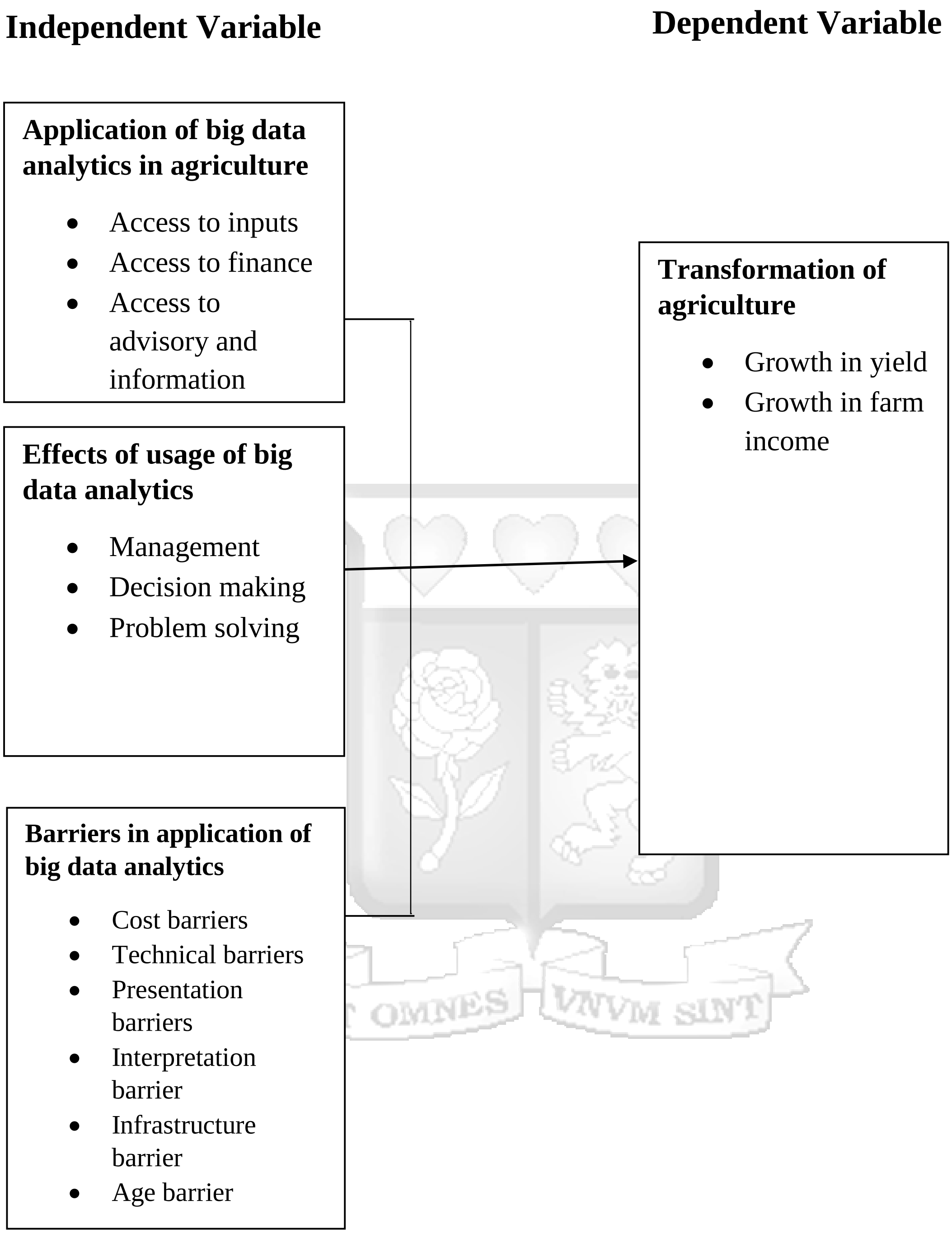

Figure 2.1 Conceptual framework

2.6 Operationalization of Variables

Table 2.1 Operationalization of variables

Independent Variables	Sub Variables	Measurement	Source
Application of Big Data Analytics in Agriculture	Access to inputs Financial access Advisory and information services	Qualitative	Jia, Hall, & Song, (2015) Sykuta, (2016) Sonka, (2016) Wolfert et al.,(2017) Jakku et al.,(2018)
Effects of usage of Big Data Analytics on transformation of agriculture	Farm management Decision making Farm problem solving	Quantitative measurements: Five Point Likert Scale 1=Strongly disagree 2=Disagree 3=Neutral 4=Agree 5=Strongly Agree	Kshetri(2014) Tatge(2016) Sonka, 2016
Barriers of application of Big Data Analytics	Challenges in implementation of BDA	Qualitative and quantitative measure Five Point Likert Scale 1=Strongly disagree 2=Disagree 3=Neutral 4=Agree 5=Strongly Agree	Kshetri(2014) Stubbs(2016) Günther, Rezazade Mehrizi, Huysman, & Feldberg, (2017) Wolfert et al.,(2017) Kamilaris et al., (2017)

Dependent Variables	Sub Variables	Measurement	Source
Transformed agriculture	Yield growth Growth in farm income	Quantitative measures of yield in kilograms and amount in Ksh	FAO(2017) Mukasa et al., 2017 Boettiger, Sara; Denis, Nicolas; Sanghvi, (2017) Kosior, (2018) Jakku et al., (2018

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Introduction

This chapter describes the research design that will be applied, the target population and the sampling as well as the methods of data collection and analysis. The chapter also details validity and reliability considerations as well as ethical considerations.

3.2 Research Design

The study applied a descriptive research design with a combination of both qualitative and quantitative data. The use of both qualitative and quantitative methods build on factors and dimensions that have recurrently been noted in literature reviewed as being important contributors of application of big data analytics in agriculture. The study was aided by the use of a case which presented an in depth understanding of the objective of the study (Creswell, 2013).

3.3 Population and Sampling

The study involved surveying a sample of the registered and active users of the platform as well as members of staff of the Mkulima Techie platform. The total number of people registered on the Mkulima Techie platform is 1,038,817. Of the registered users, 42,000 farmers are the active users of the platform. A sample of the 396 farmers were targetted respondents to the questions on the questionnaire. Mkulima Techie has 91 members of staff. These members of staff formed the target population of the key informants.

The study adopted a mix of sampling methods to select the respondents from Mkulima Techie. The sampling methods included purposive sampling and multistage cluster sampling. Purposive sampling was used to identify the key informant interviewees from Mkulima Techie staff members. The study adopted the multistage cluster sampling that entailed cluster sampling and simple random sampling from the selected cluster in order t reach the target sample size of the active users of the Mkulima Techie platform. The study sample size was determined using the sample size calculation formula by Slovin (Slovin, 1960).

$$n = N/(1 + \mathrm{N}e^2)$$

Where n is the sample size, e is the margin of error (which is 0.05 with a confidence level of 95%) N is the population of the farmers under the study.
By substitution the sample size (n) is computed as 396.

Cluster sampling will be used in order to determine the sample size of farmers who will be the respondents and is appropriate in this case because the population is large and the farmers are already organized in smaller geographical areas in which simple random sampling can be conducted. This method is similar to that used by Azumah, Donkoh, & Awuni, (2018) in their study based in Ghana on the effectiveness of technology in agriculture.

3.4 Data Collection

The study relied on structured questionnaires and key informant interviews for collection of primary data with the aim of addressing the research objectives. For the questionnaires the questions were evaluated using Likert scales with the exception of demographic data. The questionnaire was divided into five sections, the first addressing the general information about the respondent and the subsequent four sections addressing the independent variables of the study.

Key informant interviews was conducted with a purposively selected sample of the staff members of Mkulima Techie with the use of an interview guide (see Appendix 3). Due to Covid- 19 restrictions one on one interviews were not conducted, instead interviews were conducted via virtually. Information was recorded using an audio recorder and later transcribed for analysis. The secondary data were collected from multiple sources including monthly and quarterly reports, third party news articles, website information and other documents given by Mkulima Techie. Follow up email and phone conversations were also used. Data was documented using field notes and audio recordings (Creswell, 2013).

A questionnaire (see Appendix 2) that addresses the objectives of the study and answers the study questions was administered to the selected farmers who are the respondents.

3.5 Data Analysis

The process of analysis is designed to understand the data collected in context (Ranney et al., 2015). The qualitative and quantitative data will be analysed as follows:

3.5.1 Analysis of qualitative data

The qualitative data results are presented and illustrated in a descriptive form. Field notes from the interviews and documents reviewed were coded to derive concepts. The field notes collected and audio were transcribed then coded. The codes were developed into themes according to their appearance within the individual topics and locations. The identified themes were then contextualized according to the literature reviewed then be interpreted.

Thematic Analysis

Phase 1

- Involves engagement with the data by reading and re-reading in order to extract possible codes
- Transcription of the data will also be done at this point
- Records of raw data in form of field notes and recordings will also be kept

Phase 2

- Generation and documentaion of initial main codes will be done. For example if a respondent says "The use of big data has made it easier to estimate the amount of money to loan a farmer X without necessarily asking for a lot of documents and collateral therefore reducing the number of days needed to process the loan" . Here the codes can be identified as 'easier to estimate loan amount and shorter loan processing period.
- Key elements in the data will be recorded

Phase 3

- Potential themes will be selected and data that is relevant to each of these themes will be gathered from the data collected from each respondent. For example the theme' Easy and effective loan processing process" can be developed from the above code and therefore any mention of this from the data will fall into this theme.

Phase 4

- The themes will then be reviwed making sure that the application has a relation to the codes that had been extracted in Phase 2 and the entire data set.
- An on going analysis to refine the specifics of each theme will be done in order to generate clear definitions and names of each theme to ensure they have a specific identity.
- A thematic map of the analysis will thereafter be generated

Phase 5

- Selection of vivid and compelling extract examples will be done and this will be related to the research questions and literature in order to produce a scholarly report on the analysis

Figure 3.1 Thematic analysis

3.5.2 Analysis of quantitative data

The quantitative data collected through questionnaires was encoded in SPSS and also analysed with SPSS. Descriptive and inferential statistics was used to address the research objectives. Descriptive statistics was computed and reported where appropriate and provided contextual understanding of the variables assessed. Inferential statistics was subsequently conducted to assess the linear relationship between the variables under the study.

The study denotes the value a farmer derives from the platform as follows:

Logistic Model:

$$logit\ [P(y\ = 1)] = \beta_0 + \beta_1 \times Applications\ of\ BDA + \beta_2 \times Effects\ of\ usage\ of\ BDA + \beta_3 \times Barriers\ of\ application\ of\ BDA + \beta_i \times \sum Control\ variables + \varepsilon$$

Where $y = \begin{cases} 1 = \text{Derives value from the platform} \\ 0 = \text{Does not derive value from the platform} \end{cases}$

Where y is the value that a farmer derives from joining the Mkulima Techie platform and is a function of applications of BDA, effects of BDA and Barriers of BDA.

This model is similar to that applied by Turland & Slade (2020) in their study on farmers' willingness to participate in big data platform.

Merging of the results of the analysed data of both qualitative and quantitative data took place at the results point and a joint result presented to show an in depth picture of the cases using narratives, tables and figures where applicable (Schoonenboom & Johnson, 2017). This method of analysis is similar to that applied by Sattar et al., (2017) in their study on adoption of suitable agricultural practices.

3.6 Research Quality

The quality of the research is based on the reliability and validity aspects. Validation refers to the trustworthiness and authenticity of the study while reliability refers to a measure of stability of responses. The validity and reliability strategies applied for the qualitative data include triangulation, which is use of multiple sources, was used to provide validity to the findings, use of rich and thick description to describe in detail the study and continuously comparing data with the established themes. For the quantitative data a descriptive analysis of all data was applied (Creswell, 2013). A pilot study was conducted with twenty respondents. The feedback received from the participants was used to redesign the questionnaire accordingly where necessary.

3.7 Reliability

In this study, the questionnaires were subjected to an overall reliability analysis of internal consistency. The Cronbach alpha which was a coefficient of internal consistency was used to quantify the reliability of the questionnaire. Internal consistency quantifies the associations that exist between the various items on the same test and whether various items that are suggested to measure the same general construct result to similar scores. Castillio (2009) presents the decision rules as follows: >0.9 – Excellent, >0.8 – Good, >0.7 – Acceptable, >0.6 – Questionable, >0.5 – Poor and <0.5 – Unacceptable. In this study, the acceptable value of 0.7 was taken as the cut–off of reliability. The reliability test results showed that all the variables were reliable as shown by the associated Cronbach alphas that were greater than 0.7.

Table 3.1 Reliability Test Results

Reliability Statistics		
Variable	Cronbach's Alpha	N of Items
Application of BDA	0.935	9
Effects of usage of BDA	0.926	8
Barriers of application of BDA	0.879	7

3.7 Ethical Consideration

The research aims to adhere to ethical requirements while conducting the study. In the current study, the researcher will seek approval for the study from the National Commission for Science, Technology and Innovation and Strathmore University's ethics board. The researcher will also seek consent from the management of Mkulima Techie before commencement of data collection.

Confidentiality will be maintained by ensuring anonymity and by avoiding deception. The objectives of the research will be given in detail by the researcher to the respondent and confidentiality of the data collected guaranteed to them. Clarity in data collection techniques and methods will be observed. Objectivity during the consultation will be kept by the researcher to avoid favouritism in data analysis. Work from other academics and researchers will be quoted and referenced. Participation in the study is on a voluntary basis. The collected data will not be used for reasons other than those outlined for the study objectives.

CHAPTER FOUR

DATA ANALYSIS AND PRESENTATION OF RESEARCH FINDINGS

4.1. Introduction

This chapter presents the analysis of data, findings, and interpretation based on the objectives of the study. The results are presented in tables, charts, diagrams and quotations. As indicated in Figure 1.2 a thematic approach was used to analyse the qualitative data. The analysed data was arranged under themes that reflected the research objectives. The demographic information of the farmers, descriptive statistics of the findings and correlation and regression analysis as well as a summary of the chapter is provided.

4.2. Response rate

By substitution the sample size (n) was computed as 396. A total of 282 questionnaires were properly filled and returned while 6 respondents took part in the key informant interview. This represented an overall successful response rate 73% as shown on Table 4.1. A response rate of 60% and above is deemed sufficient for statistical analysis (Cooper Donald & Schindler, 2011) . Therefore this response rate was deemed satisfactory in ensuring validity of the findings.

Table 4.1 Response rate

	Response frequency	**Percentage**
Responded	282	73%
Did not respond	18	6%

Source: Primary data, (2020)

The veracity of the responses was ensured as the researcher employed the use of research assistants in administration of questionnaires to qualified respondents through the 7 sub counties that had farmers that subscribed to the Mkulima techie platform.

Table 4.2 Distribution of Respondents

Subcounty	Ward	No. of farmers
Tigania West -	Akithi,	60
	Mbeu	45
Tigania East	Muthaara	30
Igembe Central	Igembe East	40
Igembe North	Ntunene	40
Buuri,	Kiirua Naari	30
	Ruiri/Rwarera,	20
	Kibirichia	30
Imenti North	Nyaki East	60
Imenti Central	Kiagu	45
		400

4.3 Respondent Demographic Information

This section presents the demographic characteristics of the respondents including the size of the farm, the type of farming practiced and the number of years a farmer has been practicing farming. Responses to these questions are captured below.

4.3.1. Size of the farm in acres

The study sought to obtain the number of acres farmed by the respondents as this would provide profile of the farmer. The findings are presented in Figure 4.1

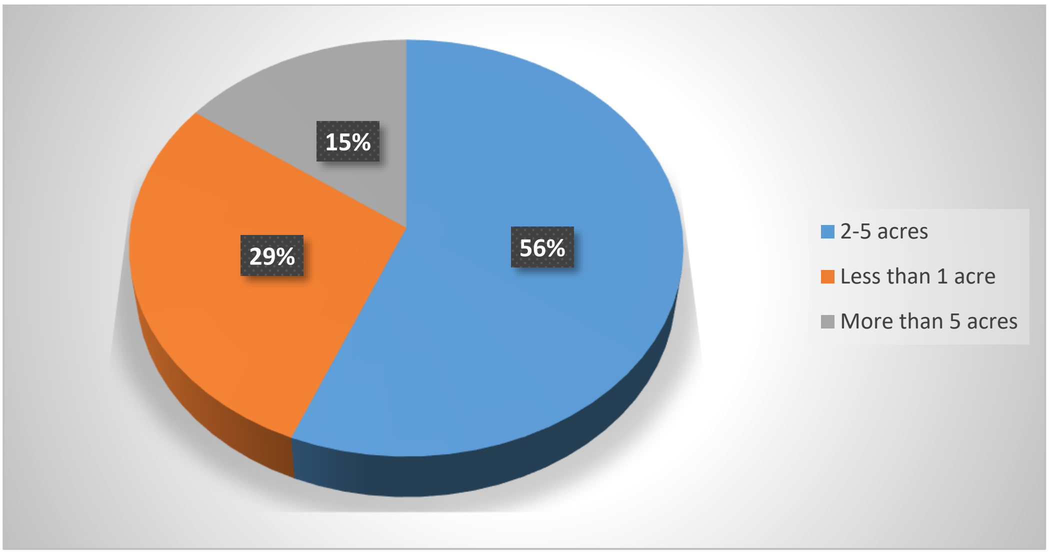

Figure 4.1 Size of the farm in acres

56% of the respondents had between 2 and 5 acres while 29% had less than an acre. 15% of the respondents had more than 5 acres. The findings of the study are reflective of a smallholder farmer population.

4.3.2. Type of farming practiced by respondents

Further the study sought to establish the type of farming practiced by the respondents. The results are as shown in Figure 4.2

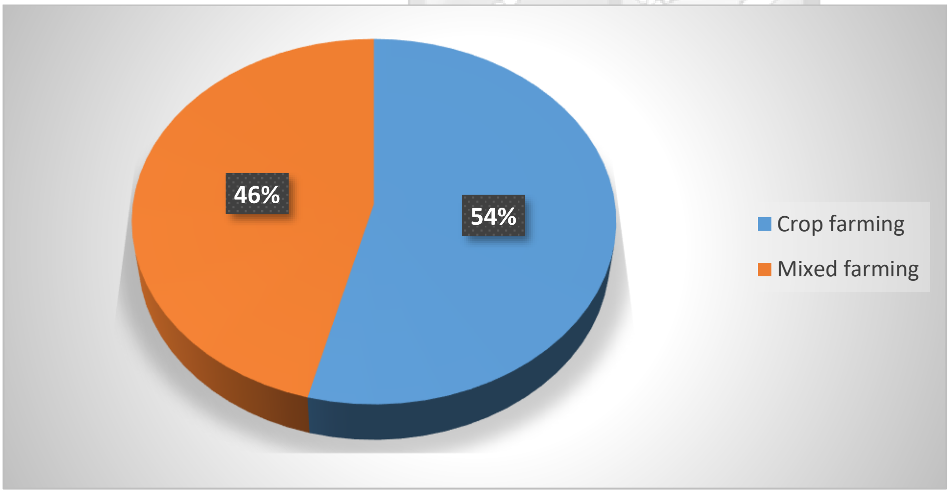

Figure 4.2 Type of farming practice

A majority of the respondents were crop farmers (54%) whereas farmers practicing mixed farming made up 46% of the sample size. Mkulima Techie platform supports crop farmers and are yet to establish a platform that supports livestock farmers.

4.3.3. Years of experience in farming

Knowing the number of years of experience in farming for the respondents served a comparative function where the farmers could evaluate the value, if any, that has been derived since they joined the platform. The findings observed are presented in the Figure 4.3

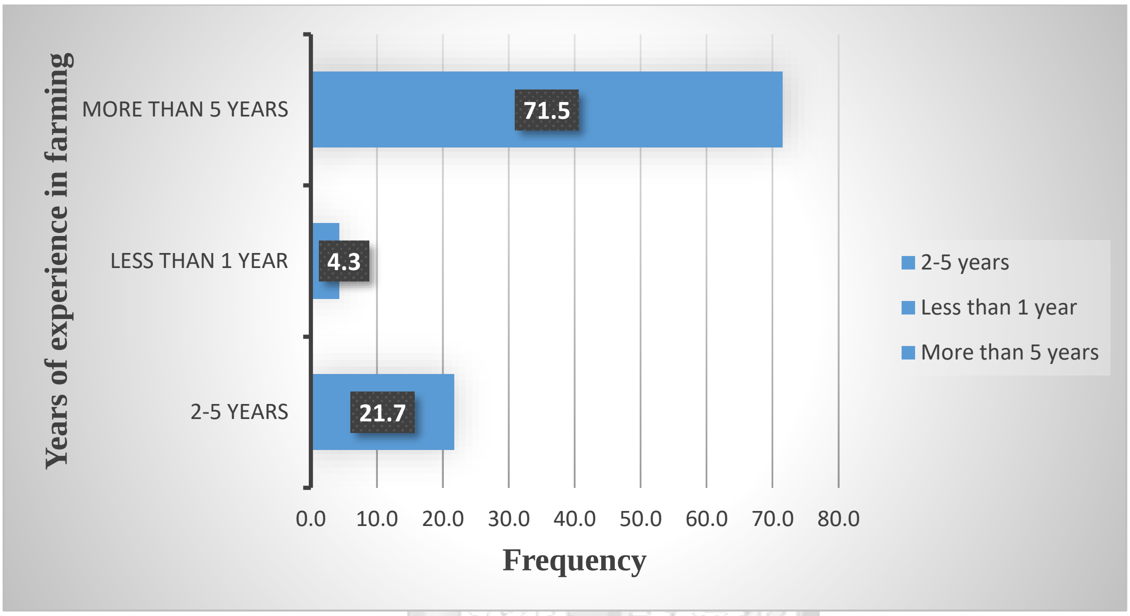

Figure 4.3 Years of experience in farming

In reference to their number of years in farming experience, 71.5% of the farmers had farmed for more than 5 years whereas those involved for less than a year made up 4.3% of the respondents. The number of respondents with 2-5 years of experience was significantly higher than those with less than a year experience in farming as seen in Figure 4.3. The findings indicate that the respondents selected were equipped with the information needed for the study since they could make comparison of years before and after joining the platform thus indicating validity in the study findings.

4.4 Application of Big Data Analytics and Transfromation of Agriculture

Respondents were asked to state the ways in which they have been using the *Mkulima Techie Platform* with the options being for farm input (fertilizer, seeds and pesticides), Finances (insurance), Advisory services or None of the services. The findings are presented in the graph below.

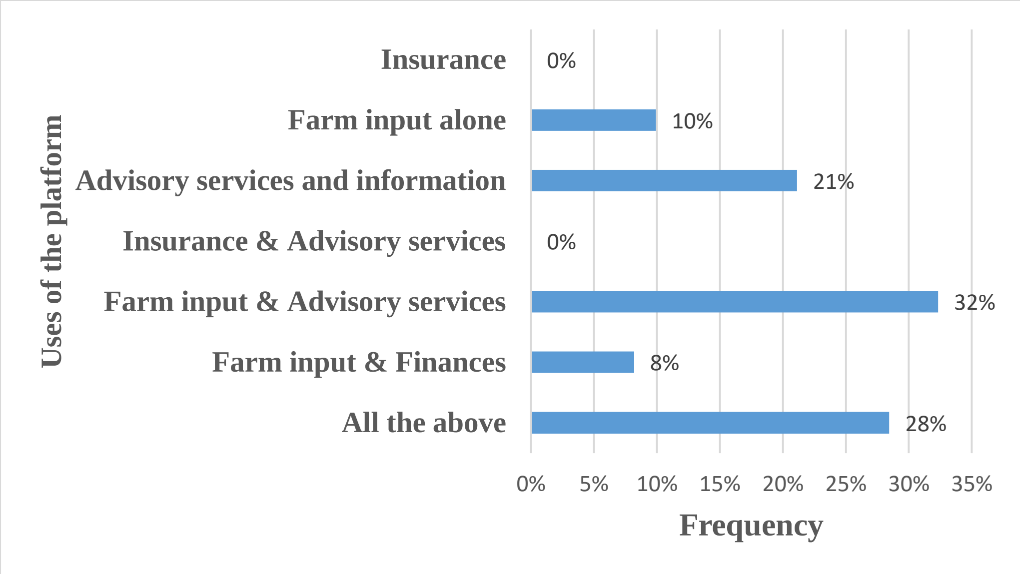

Figure 4.4 Usage of Mkulima Techie platform

Most farmers (32%) used the platform to access farm input and advisory services while 28% of the farmers used the platform to access all services provided. The study established that very few farmers used only one service. It was found out that of the farmers that used only one service, advisory services and information was the most used (21%)

In the study, the researcher further asked the respondents to give a comparison of their experience in receiving farm input, financing and advisory services and information and indicate the level to which they agreed with various statements using a 5- point Likert scale ranging from "very poor" to "excellent". The researcher assessed the level of agreement using a list of 9 statements. The findings are indicated in the table below:

Table 4.3 Effects of use of the Mkulima Techie platform

	N	Minimum	Mean	Std. Deviation
Fertilizer	282	1	3.46	1.660
Seeds	282	1	3.30	1.593
Pesticides	282	1	3.38	1.588
Herbicides	282	1	3.41	1.594
Loans	282	1	2.51	1.715
Insurance	282	1	1.96	1.509
Agronomy information	282	1	3.22	1.567
Advisory services)	282	1	3.45	1.569
Weather reports	282	1.00	2.9504	1.61764

Source: Primary data 2020

According to the findings, a majority of the farmers had a better experience in accessing fertilizer after joining the platform in comparison to before they subscribed to it as evidenced by a mean of 3.46 and a deviation of 1.66 from the mean, as well as advisory services as evidenced by a mean of 3.45 and a standard deviation of 1.569. They also had a better experience accessing herbicides as shown by a mean 3.41, pesticides as shown by a mean of 3.38 and seeds as shown by a mean of 3.30 in that order. The study established that insurance was the least accessed service with a mean of 1.509 and a standard deviation of 1.96. This may suggest that the farmers did not fully understand what insurance was especially in this case where it is bundled with the loan.

When it comes to the areas of application of big data, the key informants identified the following as areas where they used big data analytics.

4.4.1 Data Capture

In order to get inputs, the key informants explained that the farmer had the option of downloading the *Mkulima Techie* Application on a smart phone or use a USSD code, follow instructions provided to complete the self –registration process.

> "The farmer provides us with their ID number, County, Sub County and Ward names and also gives information on what crops they farm and the size of land they have. Once

they have done that they will receive an SMS on their phones confirming the registration and a farmer number that is unique to them."

One key informant further explained that after receiving this information from the farmer, it was used for geotagging.

"After self-registration process we are able to add geographic information including the latitude and longitude coordinates and other positional data. This information is used not only for location but further for radius mapping of the farmers to an input distributor. Farmers are able to redeem their loan vouchers from the distributor in order to receive input as well as get advisory and agronomy advice on their crops. The input distribution centres sometimes also act as an aggregate centre for yield collection at the end of the season."

Another key informant explained that this model helped farmers to access quality input without having to cover a long distance in search of inputs as famers were mapped to a distributor within at least a 5 km radius.

4.4.2 Redeeming E-Vouchers

The key informants explained that loans were given for inputs and this allowed farmers to get input at the beginning of the season. The repayment of the loan was done in small bits during the season or at the end of the season once a farmer had harvested and sold their produce.

"We noted that in many cases the cash that a farmer received as credit is not used for its intended purposes; rather it is used for other purposes such as for sorting family needs. Therefore, the loan is given it terms of a redeemable e-voucher and includes a facilitation fee. The loan amount also includes a mandatory insurance cover for the farmer that covers the crops in case of damage by natural calamities such as floods or drought. The farmer is also covered in case of death or permanent disability"

4.4.3 Time Specific Advisory Service

In terms of advisory services and information, the key informant interviewees confirmed that the farmers would receive information on their phones depending on the stage of growth of the crop, on what actions to take, for example what agro chemicals or fertilizer to apply.

"The information at the moment is general to farmers in a particular area and not really personalized or specific to a farmer. They are generalized according to a particular agro ecological zone but the information is time specific"

Another key informant revealed to the study that application of big data analytics had enabled the organization to provide time specific extension services.

> "With geomapping, we have the exact coordinates of farmers and thus we are able to develop precision agriculture tools, enter precise planting dates and send precise weather update that is accurate to about 5km of a farmer's location and agronomic information on what a farmer should do at a particular stage in the crop cycle"

4.5 Effects of usage of Big Data Analytics and Transformation of Agriculture

Determining the influence of Big Data Analytics on transformation of Agriculture was the second aim of the study. A 5- point Likert scale was developed which asked the respondents to indicate to what extent they agreed with the statements on Table 4.2 below. The findings are as indicated in the table:

Table 4.4 Comparison of experiences before and after joining the Mkulima Techie platform

STATEMENT	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Mean	Std. Deviation
Joining the Mkulima Techie platform has contributed to improvement in the decision-making process in the farm	15	6	9	33	38	3.73	1.395
Joining the Mkulima Techie platform has facilitated making of real-time operational decisions in terms of crop planning, irrigation techniques, fertilizer application	24	6	16	23	31	3.33	1.546
Joining the Mkulima Techie platform has contributed to reduction in waste of fertilizer	23	7	9	28	33	3.39	1.564
Joining the Mkulima Techie platform has contributed to informed strategizing	20	10	10	28	31	3.40	1.511
Joining the Mkulima Techie platform has helped to reduce the cost of production	20	6	12	28	34	3.52	1.493
Joining the Mkulima Techie platform has resulted in better preparation in anticipating weather changes and plan accordingly	27	13	11	24	24	3.05	1.557
Joining the Mkulima Techie platform has enhanced on-time problem solving abilities in terms of pest and disease control	19	9	9	29	35	3.52	1.500
Joining the Mkulima Techie platform has enhanced On time delivery of farm input	25	5	13	23	33	3.33	1.581

Source: Primary data, (2020)

The findings show that 38% of the farmers strongly agreed with the statement that joining the platform contributed to improvement in making decisions in the farm as shown by a mean of

3.73. 35% of the farmers strongly agreed that their ability to solve problems in the farm in terms of pests and diseases was enhanced as evidenced by a mean of 3.52. 34% of the farmers also strongly agreed and that joining the platform had contributed to a reduction on the cost of production. The study further established that joining the platform had contributed to the farmers having informed strategies as well as reduction in waste of fertilizer as evidenced by the mean of 3.40 and 3.39respectively. 27% of the farmers strongly disagreed with the statement that joining the platform had resulted to them having better preparation in anticipation of weather changes and thus plan accordingly.

With regards to the effects of use of big data to the organization, the key informants' interviews revealed that the use of big data analytics had enabled the organization to make data driven decisions, improve in management of the organization and solve problems within the organization which in turn meant that the farmers were able to benefit. The study identified the following themes:

4.5.1 Product design and development

The key informants noted that the use of prescriptive aspect of big data had enabled the organization to transform itself into a data – driven company. This in turn enabled the organization to design and develop products that met the customer preference and expectation. This was achieved by continuously monitoring customer behaviour, transforming the observed reality together with insights from data into comprehensible information for decision making. As such, the organization was able to transform how they issued insurance pay-outs, the pricing mechanism, credit scoring process and provision of time specific extension services to farmers.

4.5.2 Credit Cycle and Credit Scoring

The key informants informed the study that they had been able to apply data analytics to derive creditworthiness of farmers without relying on traditional credit records and traditional methods as such as getting collateral. A key informant explained that alternative data sources such as spending and transaction data, top ups on mobile money were used to mine and derive credit history for farmers and therefore set credit limits without exposing the organization to additional risks.

> "Credit scoring was initially based on credit records such as CRB and a farmer's performance on M-Pesa transactions to determine credit worthiness. We found out that we could not use just this to give a credit score to a smallholder farmer. We now disburse loans based on the input mix required in the specific value chain that a farmer is in and the size of land the farmer farms. For farmers who have been in the program

for several cycles we are able to pick up on a farmer's behaviour for example history of planting, loan repayment, consistency in production and integrate that with other data points to determine the credit score."

A key informant further informed the study that with regards to the credit cycle, the descriptive aspect of data analytics was applied to identify bottlenecks in loan repayment. This process revealed why loan repayment was poor and therefore advised the design in product change.

> "The use of big data analytics informed us to make the move to a value chain based loan. We started off with a short term loan that needed to be repaid within 30 days. We looked at the trend in repayment and realized that the default rate was high and feedback was that the period was too short. With this information and the analysis of production patterns using data collected from over a couple of years on weather patterns and production cycles from different regions, we found out that most crops crop seasons are about four months. We realized that it made no sense to ask farmers to repay before the end of that period. With the insight provided by the data we had, we were able to make the decision to extend the repayment period to 60 days then to 90 days and we have been able to improve the repayment rate."

4.5.3 Pricing mechanism

With the application of predictive and prescriptive analytics, the respondents informed the study that the organization was able to have influence the pricing mechanism. The use of big data analytics enabled the organization to make predictions on the quantity of input required thus enabling them to negotiate directly for better prices with the suppliers. Further the suppliers were willing to deliver the input to the distributors therefore reducing the costs on transportation while shortening the supply chain, resulting in getting the input at a price slightly lower than the normal market price.

> "The cost of input when we started was the same as the market price because we had very many players within the input supply chain who would add their margins and therefore make the price high. With application of big data, we are able to know the amount of input we need based on the data we have from farmers. We then make a bulk order from the suppliers and negotiate for discounts. We have recruited agro vets within different areas who become our local agents thus reducing the number of players within that supply chain"

The key informants further informed the study of that the application of big data analytics had advised the organization on the design to apply in input provision.

> "We used to give all the input at the beginning of the season therefore giving the farmer a great burden when it came to repayment especially in case of unprecedented calamities. Now we give the input depending on the crop stage and that way we can monitor the progress. From such we have been able to realize a higher repayment rates."

4.5.4 Insurance payouts

In terms of insurance pay-outs, the study was informed that application of big data analytics enabled the organization to design the insurance product by working out premiums using a predictive model and farmers would get paid when the number of occurrences exceeded a predefined threshold.

> "Crop cut method is used for yield estimation. The exercise involves having a number of farmers selected randomly in a particular agro ecological zone and a mock harvest is done. A predictive model is used to forecast actual yield per acre then a comparison is made between the actual yield and that from the model to determine whether there is a need to do any pay-outs for losses. Further, there is basis risk fund that takes care of unique cases in an area that otherwise got good yields but particular farmers did not because of calamities such as flooding in farms that are close to a river and they experienced loss because the river burst its banks."

4.5.5 Data driven decision making

The key informants further informed the study that application of big data analytics enabled the organization to come up with tools and techniques to make data driven supply chain decisions. This resulted in having the process better and faster and of benefit to the farmers. Use of predictive analytics enabled the organization to do proper forecasting and demand management which informed planning, production and scheduling processes in the organization. This in turn facilitated improvements in logistics, waste minimization, better partner engagement and even last mile efficiency. The key informants informed the study that use of data analytics had facilitated coordination among the supply chain partners by increasing visibility through use of customized dashboards. These dashboards have easy to use data views that track the different conditions in the fields enabling collaboration among the different stakeholders in the supply chain thus proving useful to all parties throughout all stages. This helped the partners to optimize fleet management, increase delivery reliability and cut costs. The advanced mapping of locations facilitated the achievement of last mile efficiency.

"In the beginning we would load up trucks with fertilizer, seeds and pesticides and had no means of verification on who would pick it up. Now from the self-registration we are able to know where the farmers are, what they are planting and make proper follow ups with the unique identification codes that the farmers get when they register. We also get the precise amount of input needed and advice our partners on the stock needed"

An interviewee further informed the study that inputs were now easily accessible to farmers and that it was delivered on time.

"With the data we have we are able to know precisely what input goes where and the quantities required. With data from the farms, from soil testing and geotagging and also the value chains that are popular with farmers within a particular region, we are able to get from our distributors more responsive chemical fertilizers as well as improved seed varieties. The input is certified and is the kind that does well in that particular region. We are able to track progress on uptake of input and therefore advice the distributors on when it is necessary to restock. And in most cases the advice is precise as it is data driven".

4.5.6 Risk Management and mitigation of fraud cases

The key interviewees further acknowledged that the application of use of big data analytics had proven useful in managing risks and fraud mitigation.

"When we started we had very many cases of fraud. Non-performing loan was at 69%. Anyone could register on the platform and access loans. In the case that a particular product was available for farmers in a particular region even farmers from other regions who heard this would travel to that area and access the products then fail to make repayment. Now we have a geotagging system and only farmers who are geotagged and mapped to an input distributor and a village advisor have access to the loan."

Reduction in fraud cases

"We have managed to reduce significantly the cases of fraud. Some people would take up several sim cards and use all of them to register for loans. We now through the self-registration platform are able to give unique identification numbers linked to an individual farmer and thus significantly reduced the number of multi registration cases."

"We have developed a yield prediction model from our rich database that is accurate and as such we can tell the yield per farmer therefore reducing the number of cases where farmers would aggregate produce from other places and sell to the organization because we offer better prices"

4.6 Barriers of Application of Big Data Analytics and Transformation of Agriculture

This section sought to establish the barriers that hindered the realization of the benefits of application of big data analytics. The statements sort to give information on the affordability of the devices and services used, accuracy ,interpretation and ease of use of the information provided as well as ease of adoptio of the goods and services recommended. The study findings are presented in Table 4.3 below.

Table 4.5 Barriers of Application of Big Data Analytics

STATEMENTS	Not at all (1)	Small extent(2)	Indifferent (3)	Moderate extent (4)	Great extent (5)	Mean	Std. Deviation
I am able to afford the recommended farm input	23.0	17.7	4.3	25.9	29.1	3.20	1.578
I am able to afford a device from which I can receive services from Big data platform	26.2	8.5	7.8	24.5	33.0	3.29	1.619
I am able to receive accurate information that helps me make decisions in the farm	20.9	14.9	10.3	26.2	27.7	3.25	1.517
I am able to interpret the weather information sent to me	27.0	16.0	13.5	21.3	22.3	2.96	1.533
I find it easy to use the information given to me via the platform	21.3	14.9	9.9	27.3	26.6	3.23	1.514
I am able to afford the services given to me by Mkulima Techie platform	18.4	6.7	10.3	24.5	40.1	3.61	1.513
I am able to adopt the technology recommended by the Mkulima Techie platform	23.4	8.5	10.3	25.5	32.3	3.35	1.567

Source: Primary data, (2020)

27% of the respondents indicated that the interpretation of weather information sent to the farmers was a challenge as indicated by a mean of 2.96 and a standard deviation of 1.533 while the price of the recommended input was not affordable to the farmers as shown by the mean of 3.20 and a standard deviation of 1.578. The farmers also indicated that the information received was not accurate enough to help them make decisions at the farm as shown by the mean of 3.23 and a standard deviation of 1.514.

The farmers agreed that they were able to afford the services given to them on the platform as indicated by a mean of 3.61 and standard deviation of 1.513 and that they were able to adopt the technology recommended to them by the platform.

The key informants informed the study of several barriers that they were facing in application of big data and the findings were arranged in the following themes:

4.6.1 Slow adoption of technology

A key informant informed the study that most farmers still wanted to maintain their traditional farming methods despite receiving educational material and extension service on their phones.

> "When I went out to meet the farmers one time, they told me that despite getting a message that there was going to be little rain, they still went on to plant because it was their planting season and did not think much on the effect the small amount of rain would do to their crops." This is an indication that the farmers are unable to effectively use information on weather alerts despite the organizations ability to develop, generate and disseminate this information.

4.6.2 Data sharing

Another key informant acknowledged that another major barrier was the hesitation in sharing of data. The respondent informed the study that the policies around data sharing were rigid and limiting.

> "The data sharing policies has made access to information a challenge because before sharing or receiving any kind of data with our partners we have to seek approval which in most cases wastes time. Other than that our partners have to sign Non -disclosure agreements before they can access data. And even with these protocols in place we can only share aggregated data, the kind that cannot uniquely identify a farmer"

4.6.3 Data quality

Another key informant informed the study of the challenge of getting quality data from farmers.

> "The self -registration portal is our primary source of data and in some cases farmers put in data that is inaccurate and unreliable, which then forces us to go apply a verification process."

The respondents further informed the study that it took a lot of time and resources to clean and verify the data to ensure it was fairly accurate and that the process was majorly manual, tedious and time consuming.

4.6.4 Disintegration of data

A key informant noted that there was lack of integration of data sets both internally and from the different organizations that they partner with. There was lack of data connectivity between the different platforms and tools used internally, which the respondent considered critical especially now that data was being collected at scale. The respondent noted that the reconciliation process of data from the different platforms in order to get information was time consuming and tedious.

> "Many players in the field have many different small sets of data that if combined could really influence the way in which the data can be used to transform agriculture, yet each of them ties their value to these sets of data and are very rigid when it comes to sharing."

4.6.5 Digital divide

Even with the availability and use of USSD codes, a key informant informed the study of pockets of people who did not have a phone or proper network especially the aged farmers and this was a barrier to the application of big data for these farmers.

4.6.6 Data security and privacy

Data collected from farmers may contain some personal information of an individual producer and may include their names, address, property location which may lead to knowing a producer's income and value of their farm land. The general worry for both the farmers and the Mkulima Techie staff was that of breech of privacy should the data get in to the hands of unauthorized third parties. The farmers expressed concerns on the data being used for regulatory enforcement purposes such as repossession in the case that they failed to repay loans.

The value of derived by the farmer was measured using average annual income and average yield per year. Study first determined whether there was a significant difference between average annual income before and after use of technology.

Table 4.6 Comparison of income and yield before and after joining the Mkulima Techie Platform

Descriptive Statistics					
Variable	N	Minimum	Maximum	Mean	Std. Deviation
Before Joining Mkulima Techie platform					
Estimated average annual income	268	0	350,000.00	34,052.99	50,004.85
Average yield per year	274	0	170,000.00	1,819.15	10,413.33
After Joining Mkulima Techie platform					
Estimated average annual income	273	0	330,000.00	34,896.01	49,164.90
Average yield per year	274	0	280,000.00	2,846.24	17,262.95
Valid N (listwise)	268				

Source: Primary data, (2020)

From the Table 4.4, the estimated mean annual income before joining Mkulima Techie platform was Ksh. 34052.99 and this was lower than the estimated mean annual income after joining Mkulima Techie platform which was Ksh. 34896.01. The estimated average yield per year before joining Mkulima Techie platform 1819.15 kgs. This was lower than the estimated average yield per year after joining Mkulima Techie platform which was 2846.24 kgs.

From the findings, we may deduce that it is likely that application of big data had influenced the increase in yield and income and that the respondents had derived value from the *Mkulima Techie Platform.*

4.7 Inferential Statistics

In addition to descriptive statistics, Wilcoxon Signed Ranks Test was used to tests whether Value derived from application of Big Data analytics was statistically significant. The findings are shown below:

Table 4.7 Value derived from application of Big Data Analytics

Mkulima Techie platform: Before - after		
	Z	**P VALUE**
Estimated average annual income	-4.956^{b}	0.000
Average yield per year	-5.989^{b}	0.000
a. Wilcoxon Signed Ranks Test		
b. Based on negative ranks.		

Source: Primary data, (2020)

Since significant differences was observed before and after use of joining Mkulima Techie platform hence from this information, a categorical variable was created to denote farmers who derived value by use of the technology and those who did not derive value defined by

$$y = \begin{cases} 1 = Does \text{ Mkulima Techie platform derives value } - Yes \\ 0 = Does \text{ Mkulima Techie platform derives value } - No \end{cases}$$

Where derive value means that the farmers average annual income or average yield per year was greater after use of the technology that before and vice versa.

4.7.1 Correlation Analysis

Correlation analysis was used to determine the association between dependent and interdependent variables and the magnitude of this association. This study's aim was to determine whether the value derived by a farmer was related to the use of big data analytics. Since both variables are categorical, chi square test of independence was used to test the association between these two categorical variables.

4.7.1.1 Chi Square Tests of Independence

Chi-square tests for associations were performed to check for the association between value derived by the farmer and use of big data analytics. A chi square tests was run between the dependent variable (y) and application of big data analytics, effect of application of big data analytics and barriers of usage of big data analytics. The results are presented in Table 4.6. There was significant association between the dependent variable and application of BDA in agriculture($chi\ square = 93.808, p < 0.01$), effect of use of BDA in transformation of agriculture ($chi\ square = 53.961, p < 0.01$) and Barriers of application of BDA($chi\ square = 77.803, p < 0.01$).

Table 4.8 Chi-Square Test Results

Chi-Square Tests			
Variable	**Pearson Chi-Square Value**	**df**	**Asymptotic Significance (2-sided)**
Application of BDA	93.808[a]	48	0.000
Effects of usage of BDA	53.961[a]	30	0.005
Barriers of application of BDA	77.803[a]	31	0.000

Source: Primary data, (2020)

4.8 Logistic Regression Analysis

To establish the application of BDA, effects of usage of BDA and Barriers of applications of BDA on dependent variable (y), the study computed a logit regression model that was used as defined by the following equation.

4.8.1 Logistic Model:

$$\begin{aligned} logit\ [P(y = 1)] &= \beta_0 + \beta_1 \times Applications\ of\ BDA + \beta_2 \times Effects\ of\ usage\ of\ BDA \\ &+ \beta_3 \times Barriers\ of\ application\ of\ BDA + \beta_i \times \sum Control\ variables \\ &+ \varepsilon \end{aligned}$$

Where $y = \begin{cases} 1 = \text{Derives value from the platform} \\ 0 = \text{Does not derive value from the platform} \end{cases}$

Deriving value meant the average income per year was higher after use of the Mkulima Techie platform compared to before subscribing to it.

4.8.2 Omnibus Tests of Model Coefficients

Before looking at the effects of independent variables on dependent variable, there was need to test whether the model improved ability to predict the outcome. This was done by comparing a model without any of the variables (Baseline or "Intercept only" model) against one with independent variables (Final) as shown in Table 4.7. This comparison was to see which model significantly improved the fit to the data.

Table 4.9 Assessment of the model to explain variation in the dependent variable

Omnibus Tests of Model Coefficients				
		Chi-square	df	Sig./Pvalue
Step 1	Step	97.603	25	0.000
	Block	97.603	25	0.000
	Model	97.603	25	0.000

Source: Primary data, (2020)

The *Omnibus Tests of Model Coefficients* is used to check that the new model (with explanatory variables included) is an improvement over the baseline model. It uses chi-square tests to see if there is a significant difference between the Log-likelihoods (specifically the -2LLs) of the baseline model and the new model. If the new model has a significantly reduced -2LL compared to the baseline, then it suggests that the new model is explaining more of the variance in the outcome and is an improvement. Here the chi-square is highly significant ($\mathrm{chi-square} = 97.603, \mathrm{df}\ = 1, \mathrm{p} < .001$) so our new model is significantly better.

A further analysis of the goodness of fit was assessed using the Pseudo R squared as shown in Table 4.8. The R^2 values tell us approximately how much variation in the outcome is explained by the model as interpreted in a linear regression model.

Table 4.10 Model Summary

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	97.237[a]	0.356	0.595
a. Estimation terminated at iteration number 8 because parameter estimates changed by less than .001.			

Source: Primary data, (2020)

The *Model Summary* as shown above provides the -2LL and pseudo-R^2 values for the full model. The -2LL value for this model (97.237) is what was compared to the -2LL for the previous null model in the 'omnibus test of model coefficients' which told us there was a significant decrease in the -2LL, that is our new model (with explanatory variables) is significantly better fit than the null model. The Nagelkerke's R^2 suggests that the model explains roughly 59.5% of the variation in the outcome hence the fitting model is good.

4.8.3 Interpreting how independent variables contributes to variations in the dependent variable.

Table 4.2 Coefficient in the model

Variables in the Equation							
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	Application of BDA	-14.178	4.123	11.821	1	0.001	0.000
	Effects of usage of BDA	-0.669	0.493	1.844	1	0.174	0.512
	Crop farming(1)	0.263	0.660	0.159	1	0.690	1.301
	Time(1)	2.895	1.056	7.513	1	0.006	18.075
	Less than 1 acre(1)	-0.173	1.010	0.029	1	0.864	0.841
	2-5 acres(1)	0.385	0.943	0.167	1	0.683	1.470
	Barriers of application of BDA	-0.698	0.581	1.442	1	0.230	0.497
	Farm inputs	1.005	0.723	1.934	1	0.164	2.733
	Finances	-5.708	1.903	9.000	1	0.003	0.003
	Advisory	5.264	1.901	7.672	1	0.006	193.350
	Joining the platform (yes)	6.300	1.514	17.325	1	0.000	544.513
	Fertilizer	1.724	0.937	3.383	1	0.066	5.608
	Seeds	0.349	0.731	0.228	1	0.633	1.417
	Pesticides	-1.573	2.079	0.572	1	0.449	0.207
	Herbicides	5.980	2.163	7.643	1	0.006	395.559

	Loans	2.370	0.680	12.154	1	0.000	10.698
	Insurance	1.418	0.613	5.349	1	0.021	4.128
	Agronomy information	2.724	0.820	11.028	1	0.001	15.236
	Advisory services)	1.252	0.694	3.253	1	0.071	3.497
	Timely and reliable	-0.463	0.395	1.376	1	0.241	0.629
	Cost effective	-0.076	0.473	0.026	1	0.873	0.927
	Accurate	1.366	0.493	7.683	1	0.006	3.921
	Personalized	-1.874	0.497	14.223	1	0.000	0.154
	Actionable	0.408	0.466	0.765	1	0.382	1.504
	Constant	2.464	1.901	1.680	1	0.195	11.746
a. Variable(s) entered on step 1: Application of BDA, Effects of usage of BDA, Crop farming, Time, Less than 1 acre, 2-5 acres, Barriers of application of BDA, SecB21, SecB22, SecB23, SecE5, Fertilizer, Seeds, Pesticides, Herbicides, Loans, Insurance, Agronomy information, Advisory services), Timely and reliable, Cost effective, Accurate, Personalized, Actionable.							

Source: Primary data, (2020)

The table above provides the regression coefficient (**B**), the Wald statistic (to test the statistical significance) and the all-important Odds Ratio (**Exp (B)**) for each variable category.

From the table above, the coefficient of applications of BDA is negative and significant$(\text{Wald} = 11.821, \text{df} = 1, p < .05)$. The Exp (B) column (the Odds Ratio) tells us that for every one-unit increase in applications of BDA, the likelihood of a farmer deriving value decreases (by 0.001 times). Under applications of BDA, the following variables were found to have a significant effect on the value derived from the use of *Mkulima Techie Platform* ; Finances$(\text{Wald} = 9.0, \text{df} = 1, p < .05)$, advisory $(\text{Wald} = 7.672, \text{df} = 1, p < .05)$ Farm Input$(\text{Wald} = 7.643, \text{df} = 1, p < .05)$, loans$(\text{Wald} = 12.154, \text{df} = 1, p < .05)$, insurance$(\text{Wald} = 5.349, \text{df} = 1, p < .05)$, and advisory and information services$(\text{Wald} = 11.028, \text{df} = 1, p < .05)$. Use of *Mkulima Techie Platform* in finance is less likely lead to derivation of value by 0.003 holding other factors constants whereas use of *Mkulima Techie Platform* in advisory is likely to lead to derivation of value by 193.53 times holding other factors constant. Use of *Mkulima Techie Platform* to access Farm Input is likely to lead to derivation of value by 395.6 times holding other factors constants. Use of *Mkulima Techie Platform* in loans is likely lead to derivation of value by 10.7 holding other factors constants.

CHAPTER FIVE

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

5.1 INTRODUCTION

This chapter gives a discussion and summary of the findings in relation to the research problem and research objectives by showing how the data collected answered the objective questions and how the results from the findings relate to the existing body of knowledge. The chapter's aim is to highlight the conclusion, recommendation and limitations of the study.

5.2 DISCUSSIONS

This section aims to provide a discussion of the findings on the various objectives in the study. The section is structured into sub sections, with each sub section addressing the objectives of the study that had been set forth.

5.2.1 Application of Big Data Analytics and Transformation of Agriculture

Protopop & Shanoyan(2016) in a study of Big data and Smallholder farmers provided illustrations of cases of big data applications. As set forth by these authors, Big Data analytics was applied to enhance farmers' access to credit, insurance, input, agronomic and market information. The results of this study are in line with these findings. Most respondents from the study indicated high ratings on access to farm input and advisory services and information. Even for the farmers that only used one service, advisory services and information was the most used. This finding was corroborated by findings from key informants who informed the study that farmers received time specific advisory service on their phones based on the stage of growth of the crop. The Agriculture Transformation Theory is significant in assessing the importance of investment in knowledge and education would facilitate transformation in agriculture. Big Data Analytics enabled the platform to develop precision agricultural tools to facilitate this.

Insurance was the least used service on the platform according to the findings of the study. This observation can be attributed to the limited information available on insurance for smallholder farmer. Patt, Suarez, & Hess,(2010) highlighted in their study that farmers did not necessarily understand the core concept of insurance. Sibiko, Veettil, & Qaim, (2018)in further elaboration of the topic showed that most farmers were hesitant to take up insurance unless it was bundled up with other benefits. This finding therefore, in relation to diffusion of innovation theory, indicates that most farmers suffered an insufficiency of information and thus did not understand that insurance was part of the loan bundle. The theory gives guidelines on communication channels to ensure that the innovation is accepted. More research can be done to investigate the other causes of low understanding of insurance for smallholder farmers despite it being beneficial to them because smallholder farmers face the adverse effects of crop losses due to adverse weather conditions.

5.2.2 Effect of use of Big Data Analytics and Transformation of Agriculture

In terms of effects of use of insights from Big Data Analytics, results from the study indicated that application of Big Data Analytics facilitated farmers' ability to manage their farms better, make decisions for the different farm activities and solve problems in their farm systems. The study findings are in line with those of (Kshetri, 2014) and (Kosior, 2018) who indicated that insight derived big data analytics facilitated farmers to enhance decision making and have more informed strategies as well as helping them in operational decisions such as crop planning, fertilizer application and disease control. Findings from key informants informed the study that use of big data analytics had enabled the platform to come up with better pricing mechanisms and better ways of deriving credit scores and credit cycles. Further they were able to make data driven decisions with regards to planning, forecasting and demand management. The farmers in turn, from the results of the study agreed that they were able to make decisions regarding farm activities and solve problems that related to pest and diseases. The findings also show that they were able to reduce costs of production and come up with informed strategies for their farms. This can be attributed to the application of prescriptive and predictive models that facilitated on time delivery of farm inputs, demand forecasting and management and shorter supply chain. These findings were in line with those of (Ghalehkhondabi et al., 2020) and (Hassen & Chen, 2020) who found out that application of big data analytics was a valuable tool that facilitated demand prediction, inventory management, product development and pricing mechanism as well as supply chain performance optimization and improvement in the decision making process.

Contrary to the views put forth by an article by ICRISAT, (2019) that noted that use of big data analytics provided farmers with decision support tools that helped farmers make decisions under uncertainties like rainfall, this study findings revealed that farmers were not able to use the information provided to make better preparations in anticipation of weather changes. A case in point is where farmers received information on their phones that they should expect little rain within that season, but they still went ahead to plant crops that required a substantial amount of rain, considering that all the respondents practiced rain fed agriculture. (Lioutas et al., 2019) suggested that this can be attributed to the limitation that the application of big data analytics at farm level was highly dependent on a farmers' motivation, goals and intentions and that the community and rules, such as traditions highly influenced the way in which big data could be fully exploited at farm level. In this case, farmers were highly influenced by routines, that planting season begins at a particular month, and it would not matter if they had information beforehand that there would be little rain, they still went on to plant. The diffusion of innovation theory concepts may provide guidelines on adoption and utilization of insights from Big Data. These findings may be indicative of a need for further studies on ways of enhancing behaviour change among farmers from process driven to data driven agriculture in order to optimize on insights from Big Data Analytics.

5.2.3 Barriers of Application of Big Data Analytics and Transformation of Agriculture

The results of this study showed that data sharing posed a major barrier in realization of the full potential of Big Data analytics as some of the data sharing policies were rigid. This finding is similar to that of (Günther et al., 2017) who noted that organizations were hesitant to share or exchange data with their network partners citing security concerns while others feared that they would be putting at risk their organization's uniqueness as data analytics was considered a source of competition. This study also showed that rigidity in data sharing contributed to the lack of connectivity between the different data sets and that the reconciliation process was both tedious and time consuming. This is similar to the findings of (Günther et al., 2017) whose study also noted that synthesizing data from many different stakeholders was complex and it was difficult to ensure high data quality.(Lioutas et al., 2019b) further elaborated that data cleaning process was very subjective and would potentially lead to serious biases putting in doubt the accuracy of big data terms of data security, this study revealed that the general worry for both the farmers and the staff of the platform was breech of privacy of the data collected as it contained sensitive information that is unique to an individual and should it land in the hands of an unauthorized third party then it could be misused. This finding is similar to that of (Kshetri, 2014) whose findings showed that there were concerns about potential misuse of information that could affect farmers and farming activities. The findings may be indicative of need to have further studies on the willingness of organizations in the agriculture sector to share data for professional gain in the sector.

5.3 CONCLUSION

This study's aim was to establish the value of big data analytics on transformation of agriculture. The approach involved establishing correlations and relationships between insights from big data analytics and transformation of agriculture. To determine the value derived by the farmer, the study measured the difference in the average annual income and the average yield per year before and after joining the Mkulima Techie platform. The correlation analysis showed that there was a significant association between insights from big data analytics and value derived. Deriving value meant that there was an increase in the average income and yield per year after joining the platform compared to before joining it.

Finding from the study showed that since subscribing to the *Mkulima Techie Platform* the respondents had better access to farm input, financing and advisory and information services. The services provided from the platform have enabled the farmers to adjust their farm management practices, guided their decision making process on the farm and enhanced their ability to solve problems within the farm. The platform attributed its ability to help smallholder farmers access input, finance and advisory services to application of insights from Big Data analytics. Credit scoring models are data driven, enabled them to provide farmers with a customized bundle for input, insurance and agronomic advice, this way transforming traditional driven agricultural systems to smarter systems that are driven by data.

While the findings from this study are in line with those observed from the previous research, this study revealed and helped to fill the gaps in the available literature by showing the real promise of data driven agriculture in action. The findings therefore address the main objective of the study which sought to establish the value of big data analytics on transformation of agriculture. Significant differences were observed before and after joining the *Mkulima Techie Platform* where value was derived as shown by the increase in average yield and annual income.

Applying the theories of diffusion of information theory and agriculture transformation theory of Schultz to the study provided theoretical guidelines to depict the factors that shape the adoption and utilization of insights from Big Data Analytics in agriculture. The theories shed more light on the interdependence relationship between transformation of agriculture and technology in agriculture for modernization of agriculture to take place. This guide brings out the vital components that facilitate successful adoption and optimization of the benefits of insights from Big data analytics showing that technologies, can be used to transform practices by making it collaborative, flexible and accessible.

5.4 Limitations of the study

The study was conducted in the time of Covid – 19 pandemic when cessation of movement into and out of Nairobi had just been lifted and service provision to farmers had been suspended temporarily. As such it was anticipated that this may not have been the right time to have people participate in the study. Some potential participants were not able to take part in the study. With the help of leaders of the organized farmer groups, the researcher was able to administer the questionnaires to the respondents. The findings still provide interesting insights in a smallholder farmer context.

5.5 Recommendations and Areas of further research

This study has shown that the existing data is highly fragmented and that different stakeholders hold sets of it, that they cling on to because they consider it a source of competition and that sharing it would jeopardize an organization's uniqueness. Not all the stakeholders have equal access to this data. Even with the promulgation of the Data Protection Act 2020, there is no regulation that has been specifically tailored for the agricultural industry. Some of the information collected in agriculture such as crop yield data, farm input application or land location do not fall into categories of personal information that is protected by this law yet if unauthorized third parties gained access to such information it would be a security breach. The study recommends that there should be regulations tailored for the agricultural industry that govern data sharing. Perhaps the various parties in the agriculture industry could establish guidelines for data such as producer data and develop privacy and security guidelines for data from the farm. Further, guidelines on data sharing with third parties should be developed so that these parties are held to some standards. The major essence of these guidelines would be to allow data sharing while ensuring better protection of the agriculture data than the current system affords. Farmers should also be involved in the process as they play a double role as the producers of this data and the consumers of it.

The findings also showed that despite receiving valuable insights, farmers still practice process driven farming decisions. This calls for need to get feedback from farmers on their perception of information and ways in which this information is disseminated so as to incorporate farmers' views into the technology used and information provided. Use of model farms to show best practices would also go a long way in encouraging behaviour change among the farmers. Perhaps the targeting criteria could also be looked into so that subscription is not necessarily open to all, including those who are necessity driven but rather to those who are opportunity driven, that will take the insight seriously and act on them accordingly.

With the advent of COVID-19, the agricultural sector suffered damage, with farmers not being able to access input and finances because of the pandemic situation. The containment measures caused disruptions in farm activities, access to agricultural inputs and finance. Restrictions on movements prevented some farmers from selling produce in markets while for others it prevented them from sourcing labour. Big Data analytics can play a role in reducing intermediaries, enabling farmers to connect directly with suppliers and with consumers by facilitating communication as well as use of social media networks to facilitate aggregation, transport and sale of produce

REFERENCES

ACET. (2017). *Agriculture Powering Africa's Economic Transformation*. 1–191.

Addison, C., & Msengezi, C. (2018). *Farmer organisations and precision agriculture data services*. https://ictupdate.cta.int/en/article/farmer-organisations-and-precision-agriculture-data-services-sid01f3106f8-adba-4b7b-b8d0-d56305668a01

African Development Bank. (2016). *Feed Africa : Strategy for Agricultural Transformation in Africa 2016-2025. May*.

African Union Commision. (2014). *Malabo Declaration on Accelerated Agricultural Growth and Transformation for Shared Prosperity and Improved Livelihoods. https://au.int/sites/default/files/documents/31247-doc-malabo_declaration_2014_11_26.pdf*. https://au.int/sites/default/files/documents/31247-doc-malabo_declaration_2014_11_26.pdf

AGRA. (2018). *Africa Agriculture Status Report 2018: Catalyzing Government Capacity to Drive Agricultural Transformation*.

Agricultural Technology Adoption Initiative. (2016). *EVIDENCE FOR TRANSFORMATION : FRAMING A RESEARCH AGENDA IN AGRICULTURE FOR DEVELOPMENT*.

Arjun, S., Joshi, A., Das, H. P., & Amutha, R. (2016). *Big Data Analytics for Agricultural Development in India*. *4*(22), 1–5.

Azumah, S. B., Donkoh, S. A., & Awuni, J. A. (2018). The perceived effectiveness of agricultural technology transfer methods: Evidence from rice farmers in Northern Ghana. *Cogent Food & Agriculture*, *4*(1), 1–11. https://doi.org/10.1080/23311932.2018.1503798

Bekmamedova, N., & Shanks, G. (2014). Social media analytics and business value: A theoretical framework and case study. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 3728–3737. https://doi.org/10.1109/HICSS.2014.464

Boateng, G. (2017). *Agricultural transformation in Africa: The myths, key issues, and the new pathway – ACET*. http://acetforafrica.org/media/blogs/agricultural-transformation-in-africa-the-myths-key-issues-and-the-new-pathway/

Boettiger, Sara; Denis, Nicolas; Sanghvi, S. (2017a). Readiness for agricultural transformation. *McKinsey Quarterly*, *November*. https://www.mckinsey.com/industries/chemicals/our-insights/readiness-for-agricultural-transformation?cid=soc-app

Boettiger, Sara; Denis, Nicolas; Sanghvi, S. (2017b). Readiness for agricultural transformation | McKinsey & Company. *McKinsey Quarterly*, *November*. https://www.mckinsey.com/industries/chemicals/our-insights/readiness-for-agricultural-transformation?cid=soc-app

Bronson, K., & Knezevic, I. (2016). Big Data in food and agriculture. *Big Data & Society*, *3*(1), 205395171664817. https://doi.org/10.1177/2053951716648174

Brussels Development Briefing. (2013). *Drivers of Success for Agricultural Transformation in Africa*. *33*, 1–7.

CGIAR. (2016). *CGIAR Big Data Coordination Platform Proposal CGIAR Big Data Coordination Platform Leveraging CGIAR data : Bringing big data to agriculture , and agriculture to big data March 2016*.

Chen, H.-M., Kazman, R., & Matthes, F. (2015). *Association for Information Systems AIS Electronic Library (AISeL) Demystifying Big Data Adoption: Beyond IT Fashion and Relative Advantage Demystifying Big Data Adoption: Beyond IT Fashion and Relative*

Advantage. http://aisel.aisnet.org/digit2015%5Cnhttp://aisel.aisnet.org/digit2015%5Cnhttp://aisel.aisnet.org/digit2015/4

Chi, M., Plaza, A., Benediktsson, J. A., Sun, Z., Shen, J., & Zhu, Y. (2016). Big Data for Remote Sensing: Challenges and Opportunities. *Proceedings of the IEEE*, *104*(11), 2207–2219. https://doi.org/10.1109/JPROC.2016.2598228

Coble, K. H., Mishra, A. K., Ferrell, S., & Griffin, T. (2018). Big data in agriculture: A challenge for the future. *Applied Economic Perspectives and Policy*, *40*(1), 79–96. https://doi.org/10.1093/aepp/ppx056

Constantiou, I. D., & Kallinikos, J. (2015). New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology*, *30*(1), 44–57. https://doi.org/10.1057/jit.2014.17

Cooper Donald, & Schindler, P. (2011). *Business Research Methods(11th ed)*. McGraw-Hill. moz-extension://4e493f4f-b290-4392-9afa-ed20419aa9e2/enhanced-reader.html?openApp&pdf=http%3A%2F%2Fsutlib2.sut.ac.th%2Fsut_contents%2FH139963.pdf

Creswell, J. (2013). *Qualitative Inquiry and Research Design:Choosing Among Five Approaches*.

CTA. (2019). *The Digitalisation of African Agriculture Report 2018-2019* (1st Editio).

D'Alessandro, S., Caballero, J., Lichte, J., & Simpkin, S. (2015). *Kenya: Agricultural Sector Risk Assessment*. *96289*, 126. www.worldbank.org

Dean, M. (2018). *Launching into space: using satellite imagery in financial services – Finance in a Digital Africa*. Mastercard Foundation. https://www.financedigitalafrica.org/2018/07/12/launching-into-space-using-satellite-imagery-in-financial-services/

Delgado, J. A., Short, N. M., Roberts, D. P., & Vandenberg, B. (2019). Big Data Analysis for Sustainable Agriculture on a Geospatial Cloud Framework. *Frontiers in Sustainable Food Systems*, *3*(July). https://doi.org/10.3389/fsufs.2019.00054

Divanbeigi, R., & Saliola, F. (2016). Regulation and the Transformation of Agriculture. *Journal of Agriculture and Environmental Science*, *September*, 1–22.

FAO. (2017a). Agricultural Transformation in Africa: The role of natural resources. *Agricultural Transformation in Africa*, *31*(1).

FAO. (2017b). Agricultural Transformation in Africa. *Agricultural Transformation in Africa*, *31*(1).

FAO. (2019). *Digital technologies in Agriculture and Rural Areas-Status Report*. https://doi.org/10.4324/9780429507533-13

FAO, & ITU. (2019). E-Agriculture in Action: Big Data for Agriculture. Bangkok. In *Journal of Chemical Information and Modeling* (Vol. 53, Issue 9). https://doi.org/10.1017/CBO9781107415324.004

Ghalehkhondabi, I., Ahmadi, E., & Maihami, R. (2020). An overview of big data analytics application in supply chain management published in 2010-2019. *Production*, *30*. https://doi.org/10.1590/0103-6513.20190140

Gollin, D. (2014). *Smallholder agriculture in Africa*. *October*.

GSARS. (2017). *Productivity and Efficiency Measurement in Agriculture*. *February*, 1–77. https://doi.org/10.1016/S0933-3657(12)00077-2

GSMA. (2018). *Start-Ups and Mobile in Emerging Markets: Insights from GSMA Ecosystem Accellerator*. *February 2012*. https://sustainabledevelopment.un.org/partnerships/goal17/

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., & Feldberg, F. (2017). Debating big data: A literature review on realizing value from big data. *Journal of Strategic*

Information Systems, *26*(3), 191–209. https://doi.org/10.1016/j.jsis.2017.07.003

Hassen, A., & Chen, B. (2020). *Big Data Analytics for Agriculture Input Supply Chain in Ethiopia: Supply Chain Management Professionals Perspective*. https://www.diva-portal.org/smash/record.jsf?pid=diva2:1451736

Himesh, S., Prakasa Rao, E. V. S., Gouda, K. C., Ramesh, K. V., Rakesh, V., Mohapatra, G. N., Kantha Rao, B., Sahoo, S. K., & Ajilesh, P. (2018). Digital revolution and Big Data: A new revolution in agriculture. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, *13*(021), 1–7. https://doi.org/10.1079/PAVSNNR201813021

I.Protopop and A.Shanoyan. (2016). Big Data and Smallholder Farmers : Big Data Applications in the Agri-Food Supply Chain in Developing Countries. *International Food and Agribusiness Management Review*, *19*(A), 19.

IBM. (2017). *Descriptive , predictive , prescriptive : Transforming asset and facilities management with analytics Choose the right data analytics solutions to boost service quality , reduce operating costs and build ROI*. 8.

ICRISAT. (2019). *Building trust is key for Big Data to deliver benefits to farmers – ICRISAT*. https://www.icrisat.org/building-trust-is-key-for-big-data-to-deliver-benefits-to-farmers/

Jakku, E., Taylor, B., Fleming, A., Mason, C., Fielke, S., Sounness, C., & Thorburn, P. (2018). "If they don't tell us what they do with it, why would we trust them?" Trust, transparency and benefit-sharing in Smart Farming. *NJAS - Wageningen Journal of Life Sciences*, 100285. https://doi.org/10.1016/J.NJAS.2018.11.002

Jayne, T. S., & Ameyaw, D. (2016). *Africa's Emerging Agricultural Transformation: Evidence, Opportunities and Challenges*. 2–23.

Jia, L., Hall, D., & Song, J. (2015). The conceptualization of data-driven decision making capability. *2015 Americas Conference on Information Systems, AMCIS 2015*, *2011*.

Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. (2017). A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture*, *143*(October), 23–37. https://doi.org/10.1016/j.compag.2017.09.037

Kosior, K. (2018). Digital Transformation in the Agri-Food Sector –Opportunities and Challenges. *Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu*, *XX*(2), 100–106. https://doi.org/10.5604/01.3001.0011.8122

Kshetri, N. (2014). The emerging role of Big Data in key development issues: Opportunities, challenges, and concerns. *Big Data & Society*, *1*(2), 205395171456422. https://doi.org/10.1177/2053951714564227

Lamba, H. S., & Dubey, S. K. (2015). Analysis of requirements for Big Data Adoption to maximize IT Business Value. *2015 4th International Conference on Reliability, Infocom Technologies and Optimization: Trends and Future Directions, ICRITO 2015*. https://doi.org/10.1109/ICRITO.2015.7359268

Lassoued, R., Macall, D. M., Smyth, S. J., Phillips, P. W. B., & Hesseln, H. (2021). Expert insights on the impacts of, and potential for, agricultural big data. *Sustainability (Switzerland)*, *13*(5), 1–18. https://doi.org/10.3390/su13052521

Lioutas, E. D., Charatsari, C., La Rocca, G., & De Rosa, M. (2019a). Key questions on the use of big data in farming: An activity theory approach. *NJAS - Wageningen Journal of Life Sciences*, *90–91*(April), 100297. https://doi.org/10.1016/j.njas.2019.04.003

Lioutas, E. D., Charatsari, C., La Rocca, G., & De Rosa, M. (2019b). Key questions on the use of big data in farming: An activity theory approach. *NJAS - Wageningen Journal of Life Sciences*, *90–91*(April), 100297. https://doi.org/10.1016/j.njas.2019.04.003

Maru, A., Berne, D., Beer, J. De, Ballantyne, P., Pesce, V., Kalyesubula, S., Fourie, N., Addison, C., Collett, A., Chaves, J., Maru, A., Berne, D., De Beer, J., Ballantyne, P., Pesce, V., Kalyesubula, S., Fourie, N., Addison, C., Collett, A., & Chaves, J. (2018).

Digital and Data-Driven Agriculture: Harnessing the Power of Data for Smallholders. *F1000Research*, *7*(525), 38. https://doi.org/10.7490/F1000RESEARCH.1115402.1

Micheni, E. M. (2015). Diffusion of Big Data and Analytics in Developing Countries. *The International Journal Of En Gineering And Science (IJES)*, 44–50.

Ministry of Agriculture, Livestock, F. and I. (2019). *Agricultural Sector Transformation and Growth Strategy 2019-2029*. 1–216.

Mukasa, A. N., Woldemichael, A. D., Salami, A. O., & Simpasa, A. M. (2017). Africa's Agricultural Transformation: Identifying Priority Areas and Overcoming Challenges. *Africa Economic Brief*, *8*(3).

Omo-Ojugo, E. (2018). Relevance of Big Data Analytics in Agriculture: Focus on Nigeria Agricultural Sector. *International Journal of Science and Research (IJSR)*, *7*(9), 1–11. https://doi.org/10.21275/ART2019910

Orr, G. (2003). *Diffusion of Innovations , by Everett Rogers (1995)*. *1995*, 1–5.

PA Consulting. (2017). *TRANSFORMING AGRICULTURE WITH DATA-DRIVEN INSIGHTS*.

Pasa, R. B. (2017). Technological Intervention in Agriculture Development. *Nepalese Journal of Development and Rural Studies*, *14*(1–2), 86–97. https://doi.org/10.3126/njdrs.v14i1-2.19652

Patt, A., Suarez, P., & Hess, U. (2010). How do small-holder farmers understand insurance, and how much do they want it? Evidence from Africa. *Global Environmental Change*, *20*(1), 153–161. https://doi.org/10.1016/j.gloenvcha.2009.10.007

Pham, X., & Stack, M. (2018). How data analytics is transforming agriculture. *Business Horizons*, *61*(1), 125–133. https://doi.org/10.1016/j.bushor.2017.09.011

Poppe, K., Wolfert, S., Verdouw, C., & Renwick, A. (2015). A European Perspective on the Economics of Big Data. *Farm Policy Journal*, *12*(1), 11–19. https://www.researchgate.net/publication/278300518

Protopop, I., & Shanoyan, A. (2016). Big Data and Smallholder Farmers : Big Data Applications in the Agri-Food Supply Chain in Developing Countries. *International Food and Agribusiness Management Review*, *19*(A), 19.

Ranney, M. L., Meisel, Z. F., Choo, E. K., Garro, A. C., Sasson, C., & Morrow Guthrie, K. (2015). Interview-based Qualitative Research in Emergency Care Part II: Data Collection, Analysis and Results Reporting. *Academic Emergency Medicine*, *22*(9), 1103–1112. https://doi.org/10.1111/acem.12735

Ribarics, P. (2016). Big Data and its impact on agriculture. *Ecocycles*, *2*(1), 33–34. https://doi.org/10.19040/ecocycles.v2i1.54

Rodriguez, D., de Voil, P., Rufino, M. C., Odendo, M., & van Wijk, M. T. (2017). To mulch or to munch? Big modelling of big data. *Agricultural Systems*, *153*, 32–42. https://doi.org/10.1016/j.agsy.2017.01.010

Sattar, R. S., Wang, S., Muqadas, M., Ashraf, M. F., & Tahir, M. N. (2017). Qualitative and quantitative approaches to study adoption of sustainable agricultural practices: A research-note on mixed method approach. *International Journal of Agricultural Extension and Rural Development*, *5*(2), 539–544. www.internationalscholarsjournals.org

Sawant, M., Urkude, R., & Jawale, S. (2016). *Organized Data and Information for Efficacious Agriculture Using PRIDE ^{TM} Model*. *19*, 115–130.

Schoonenboom, J., & Johnson, R. B. (2017). How to Construct a Mixed Method Research Design. *Kolner Zeitschrift Fur Soziologie Und Sozialpsychologie*, *69*, 107–131. https://doi.org/10.1007/s11577-017-0454-1

Sibiko, K. W., Veettil, P. C., & Qaim, M. (2018). Small farmers' preferences for weather index insurance: Insights from Kenya. *Agriculture and Food Security*, *7*(1), 1–14.

https://doi.org/10.1186/s40066-018-0200-6
Slovin, E. (1960). *Slovin's Formula*. 0–2.
Stubbs, M. (2016). *Big Data in U . S . Agriculture*.
Tatge, J. (2016). *Data is the New Cash Crop: Understanding the Market for Farm Data*. https://dirtforfarmers.com/data-is-the-new-cash-crop-understanding-the-market-for-farm-data-59056dda285d#.yzglkmkid
Timmer, C. P. (1988). The agricultural transformation. *Handbook of Development Economics*, *1*(January 1988), 275–331. http://www.dipsa.unifi.it/romano/PDF/Timmer HDE ch08.pdf
Turland, M., & Slade, P. (2020). Farmers' willingness to participate in a big data platform. *Agribusiness*, *36*(1), 20–36. https://doi.org/10.1002/agr.21627
Wharton, C. R. (2019). Transforming Traditional Agriculture. *Subsistence Agriculture & Economic Development*, 363–365. https://doi.org/10.4324/9781315130408-37
Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M.-J. (2017). Big Data in Smart Farming – A review. *Agricultural Systems*, *153*, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023
World Bank. (2019). *Growth Unbundling the Slack in Private Sector Investment: Transforming Agriculture Sector Productivity and Linkages to Poverty Reduction*. *19*.
Yangfen, C., Yansui, L., & Keshuai, X. (2010). Characteristics and mechanism of agricultural transformation in typical rural areas of Eastern China: A case study of Yucheng City, Shandong Province. *Chinese Geographical Science*, *20*(6), 545–553. https://doi.org/10.1007/s11769-010-0430-4

APPENDICES

Appendix 1: National Commission for Science, Technology and Innovation(NACOSTI) Permit

REPUBLIC OF KENYA

NATIONAL COMMISSION FOR SCIENCE,TECHNOLOGY & INNOVATION

Ref No: 636695

Date of Issue: 15/May/2020

RESEARCH LICENSE

This is to Certify that Miss.. MILLICENT AMONDI OKUMU of Strathmore University, has been licensed to conduct research in Embu, Meru on the topic: VALUE OF BIG DATA ANALYTICS ON TRANSFORMATION OF AGRICULTURE: A CASE OF FARMERS SUBSCRIBED TO MKULIMA TECHIE IN KENYA for the period ending : 15/May/2021.

License No: NACOSTI/P/20/4941

636695

Applicant Identification Number

Director General
NATIONAL COMMISSION FOR SCIENCE,TECHNOLOGY & INNOVATION

Verification QR Code

NOTE: This is a computer generated License. To verify the authenticity of this document, Scan the QR Code using QR scanner application.

THE SCIENCE, TECHNOLOGY AND INNOVATION ACT, 2013

The Grant of Research Licenses is Guided by the Science, Technology and Innovation (Research Licensing) Regulations, 2014

CONDITIONS

1. The License is valid for the proposed research, location and specified period
2. The License any rights thereunder are non-transferable
3. The Licensee shall inform the relevant County Director of Education, County Commissioner and County Governor before commencement of the research
4. Excavation, filming and collection of specimens are subject to further necessary clearence from relevant Government Agencies
5. The License does not give authority to tranfer research materials
6. NACOSTI may monitor and evaluate the licensed research project
7. The Licensee shall submit one hard copy and upload a soft copy of their final report (thesis) within one of completion of the research
8. NACOSTI reserves the right to modify the conditions of the License including cancellation without prior notice

National Commission for Science, Technology and Innovation off Waiyaki Way, Upper Kabete,
P. O. Box 30623, 00100 Nairobi, KENYA
Land line: 020 4007000, 020 2241349, 020 3310571, 020 8001077
Mobile: 0713 788 787 / 0735 404 245
E-mail: dg@nacosti.go.ke /
registry@nacosti.go.ke Website:
www.nacosti.go.ke

Appendix 2: Strathmore University Introductory Letter

Strathmore
UNIVERSITY

7th May 2020

Ms Okumu, Millicent
millicent.okumu@strathmore.edu

Dear Ms Okumu,

RE: Value of Big Data Analytics on Transformation of Agriculture: A Case of Farmers Subscribed to Mkulima Techie in Kenya

This is to inform you that SU-IERC has reviewed and **approved** your above research proposal. Your application approval number is **SU-IERC0772/20.** The approval period is **7th May 2020 to 6th May 2021.**

This approval is subject to compliance with the following requirements:

i. Only approved documents including (informed consents, study instruments, MTA) will be used
ii. All changes including (amendments, deviations, and violations) are submitted for review and approval by SU-IERC.
iii. Death and life threatening problems and serious adverse events or unexpected adverse events whether related or unrelated to the study must be reported to SU-IERC within 72 hours of notification
iv. Any changes, anticipated or otherwise that may increase the risks or affected safety or welfare of study participants and others or affect the integrity of the research must be reported to SU-IERC within 72 hours
v. Clearance for export of biological specimens must be obtained from relevant institutions.
vi. Submission of a request for renewal of approval at least 60 days prior to expiry of the approval period. Attach a comprehensive progress report to support the renewal.
vii. Submission of an executive summary report within 90 days upon completion of the study to SU-IERC.

Prior to commencing your study, you will be expected to obtain a research license from National Commission for Science, Technology and Innovation (NACOSTI) https://oris.nacosti.go.ke and also obtain other clearances needed.

Yours sincerely,

Dr Virginia Gichuru,
Secretary; SU-IERC

Cc: Prof Fred Were,
Chairperson; SU-IERC

STRATHMORE UNIVERSITY INSTITUTIONAL ETHICS REVIEW COMMITTEE (SU-IERC)
07 MAY 2020
TEL: +254 (0)703 034 000
P. O. Box 59857 - 00200
NAIROBI - KENYA

Ole Sangale Rd, Madaraka Estate. PO Box 59857-00200, Nairobi, Kenya. Tel +254 (0)703 034000
Email info@strathmore.edu www.strathmore.edu

Appendix 3: Questionnaire

Dear Sir/Madam

RE: CONSENT FOR PARTICIPATION IN THE STUDY

Consent for participation in the study

I am a student at Strathmore Business School and would wish to conduct a research in your firm. My research assessment is on the Value of Big Data Analytics on Transformation of Agriculture. The study targets farmers that have subscribed to the Mkulima Techie platform.

The study conducted by the researcher is done purely for academic purposes. The information obtained will be treated with the utmost confidentiality and the interviewee is free to stop the process at any time during the exercise.

There are no risks in taking part in this study. All the information you provide will be treated as confidential and will not be used in any way without your express permission. Participation in this study is entirely voluntary. Even if you decide to take part at first but later change your mind, you are free to withdraw at any time without explanation.

All research records will be stored in securely locked cabinets. That information may be transcribed into our database but this will be sufficiently encrypted and password protected. Only the people who are closely concerned with this study will have access to your information. All your information will be kept confidential.

The interviews will take no more than 15 minutes to complete. Any facilitation and assistance you give in the study will be highly appreciated. You can reach me on 0728 773 758 if you require any clarification or additional information about the research. You can also contact my supervisor at Strathmore Business School by email jolukuru@strathmore.edu

I agree to participate in the research study. I understand the purpose and nature of this study and i am participating voluntarily. I understand that I can withdraw from the study at any time.

Agree ☐

Disagree ☐

Instructions: Please be as honest and objective in responding to the questions and mark the relevant box with a tick

Confidentiality: The responses you provide will be strictly confidential. No reference will be made concerning any individual(s) in the report of the study.

SECTION A

Farm Profile

Kindly indicate the following:

1. What is the size of your farm in acres?

   Less than 1 acre ☐

   2-5 ☐

   More than 5 acres ☐

2. What type of farming have you been practicing?

   Crop farming ☐

   Livestock farming ☐

   Mixed farming ☐

3. For how many years have you been engaged in farming?

   Less than 1 year ☐

   2-5 years ☐

   More than 5 years ☐

SECTION B: APPLICATION OF BIG DATA ANALYTICS

Kindly indicate the following:

4. In what ways do you use the Mkulima Techie platform? Tick where applicable

   Farm Input (Fertilizer, seeds, Pesticides, Herbicides) ☐

Finances (Insurance ☐

Advisory services and information ☐

None of the above ☐

Kindly indicate, on a level of 1(Very Poor) to 5 (Excellent), your level of agreement with the following statement:

5. In comparison to before you joined the Mkulima Techie platform, how would you describe your experience in accessing the following services?

	Very poor	**Poor**	**Average**	**Good**	**Excellent**
Fertilizer					
Seeds					
Pesticides					
Herbicides					
Loans					
Insurance(crop/livestock)					
Agronomy information(pest management,disease mitigation)					
Advisory services(prescriptions on planting methods)					
Weather alerts/reports					

SECTION C:

6. Kindly indicate your level of agreement with the following statements based on your view:

1=Strongly disagree, 2= Disagree, 3= Neutral, 4= Agree, 5= Strongly agree

	1	2	3	4	5
Joining the Mkulima Techie platform has contributed to improvement in the decision-making process in the farm					
Joining the Mkulima Techie platform has facilitated making of real-time operational decisions in terms of crop planning, irrigation techniques, fertilizer application					
Joining the Mkulima Techie platform has contributed to reduction in waste of fertilizer					
Joining the Mkulima Techie platform has contributed to informed strategizing					
Joining the Mkulima Techie platform has helped to reduce the cost of production					
Joining the Mkulima Techie platform has resulted in better preparation in anticipating weather changes and plan accordingly					
Joining the Mkulima Techie platform has enhanced on-time problem solving abilities in terms of pest and disease control					
Joining the Mkulima Techie platform has enhanced On time delivery of farm input					

SECTION D: BARRIERS

7. Kindly indicate on a scale of 1-5 your level agreement with the following statements in relation to question above:

	Not at all(1)	To a Small extent(2)	Indifferent(3)	Moderate extent(4)	Great extent(5)
I am able to afford the recommended farm input					
I am able to afford a device from which I can receive services from Big data platform					
I am able to receive accurate information that helps me make decisions in the farm					
I am able to interpret the climate information sent to me					
I find it easy to use the information given to me via the platform					
I am able to afford the services given to me by Big data platform					
I am able to adopt the technology recommended by the Mkulima Techie platform					

SECTION D: VALUE ON TRANSFORMATION OF AGRICULTURE

8. What was your annual income before you joined the Mkulima Techie platform? Ksh ___________________
9. After joining the Mkulima Techie platform what has been your average annual income? Ksh ___________________
10. What was the average yield per season in the farm before you joined the Mkulima Techie platform? ___________________ Kilograms.
11. What has been the average yield per season in the same farm since you joined the Mkulima Techie platform? ___________________ Kilograms.

Thank you for your participation

Appendix 4: Interview Guide
Respondent background

- What is your role in the organization?
- How long have you been with the company and how long in the industry?
- What is your involvement with the firm technology?

Context questions

Interventions based on application of Big Data Analytics

1. In how many areas does your firm operate?
2. How many farmers have you reached so far and what are their localities?
3. What are your current combination of data points?
4. Is there any cross-organizational data exchanges and are you able to seek partnership where possible?
5. How do you leverage on the data collected by others?
6. How does your organization currently analyse its data? (descriptive, predictive, prescriptive, or analytics)
7. How do you ensure quality and usability of the data you collect?

Influence of use of big data in transformation of agriculture

8. What are examples of relevant uses of data within your service provider model? In which areas have you benefitted most(products, services, processes or any other)
9. How is value created using big data in the company? Is the value created internally or externally or both?
10. What targeting criteria do you consider as you deliver your service and what implications does this have on your beneficiary?
11. What are some of the changes you have realised as a result of application of big data?
12. What are some of the benefits your customers have realized as a result of application of Big Data?
13. How has the decision making process been since you started applying Big Data analytics?
14. In what ways are you using farmer profile data and employing analytics to meet your goals and the needs of the farmers?
15. What framework/protocol is in place, if any, to capture data about the farm itself?

16. What are the ways in which farmers interact with the data you have for greater impact?
17. Do farmers currently have access to and use their own data in any way? If not, what new technologies might enable and empower farmers to effectively engage with their data?

Barriers

18. What are some of the obstacles you have encountered in implementation of the big data analytics project?(regulatory, legal, ethical) . Does this inhibit what can be done with big data?
19. What gaps in skills exists? What skills should be acquired internally by staff and what should be outsourced to consultants?
20. What risks have you identified for the collection and use of farmer data?
21. What misuse of data might be possible that would harm the farmer as well as your organization?
22. What challenges are your clients facing in use of big data?
23. How can these be mitigated?

Thank you for taking your time to answer the questions.