

Harnessing Big Data for Sustainable Development in Nigeria

Ojijo Odhiambo¹ & Fatima Umar¹

¹ UNDP Nigeria Country Office, Abuja, Nigeria

Correspondence: Ojijo Odhiambo, UNDP Nigeria Country Office, Abuja, Nigeria. E-mail: ojijo.odhiambo@undp.org

Received: February 10, 2019

Accepted: March 19, 2019

Online Published: May 30, 2019

doi:10.5539/jsd.v12n3p146

URL: <https://doi.org/10.5539/jsd.v12n3p146>

Abstract

Nigeria faces a myriad of development challenges in her efforts to grow the economy, create jobs and achieve the Sustainable Development Goals by 2030. Since independence, the Government has developed many Plans and Strategies, including the current Economic Recovery and Growth Plan, in an attempt to address these challenges. The ERGP, which is broadly aligned to the SDGs, is aimed at improving macroeconomic stability; fostering economic growth and diversification; improving competitiveness; fostering social inclusion; and enhancing governance and security. Recent information, communication and technological advances have led to data -from both conventional and unconventional sources- to be readily available in high volumes and velocity and in a variety of forms, or simply, to a Data Revolution. This paper examines the role of Big Data and Data Revolution in promoting sustainable development in Nigeria, as well the emerging opportunities for Statisticians in this regard. The paper posits that the attainment of the SDGs will be greatly hampered if Statisticians do not ask the right questions; access relevant data information and, crucially, perform deeper analytics around data and information. Statisticians have an important role to play in promoting Nigeria's sustainable development agenda, but only if they become more entrepreneurial; and adequately master and apply the requisite technical and non-technical skills.

Keywords: big data, data analytics, sustainable development, Nigeria

1. Introduction

Data and information have become an important driver of economic growth and sustainable development in the 21st Century. They determine the pace and direction of socio-economic development; and even political choices people make or have to contend with. Across the globe, data and information determine how, where and with whom people interact, as well as how business is conducted (Einav and Levin 2014). Indeed, some of the most valuable businesses of our time including Apple, with a market capitalization of \$926.9 billion and ranked number one in terms of profit and value; Amazon with a market value of \$777.8 ; Alphabet whose market value stands at \$766.4 billion; Microsoft with market value of \$750.6 billion; Facebook at \$541.5 billion; and Samsung with market capitalization of \$325.9 billion all deal in data in different forms (Statista, 2019). Data and information, in effect, condition the way people live their lives. They have indeed become the world's most valuable resource. Clive Humby (2006), as cited in Arthur (2013), has referred to data as the "new oil".

In September 2015, world leaders adopted Agenda 2030 and the 17 Sustainable Development Goals (SDGs), a set of universal goals to address the political, social, economic and environmental challenges facing humanity. The SDGs are indivisible and interrelated; and are aimed at ending poverty and deprivation, everywhere and for everybody, while leaving no one behind and reaching the furthest first. They represent a shift from a narrow focus on achievement of rather unambitious numerical targets to a focus on higher targets and qualitative and durable change in the lives of people, and the environment. They embody a transformative shift towards *people*, *prosperity* and *planet*; underpinned by *peace* and good governance; and *partnership* for development – the Five Ps. They are about achieving prosperous and fulfilling lives for all, in harmony with nature.

To aid development of national strategies, targeting of interventions and monitoring of progress towards the SDGs, there is an urgent need for comprehensive, accurate, up-to-date data and granular analysis of data and information. Deeper and granular analysis of comprehensive, accurate and up-to date data is key to the realization of poverty reduction (DFID, 2002). This paper discusses the role of data and statistics in promoting

economic growth and sustainable development in Nigeria, paying special attention to ‘Data Revolution’ in the specific context of the current Medium- Term Development Plan – the Economic Recovery and Growth Plan (ERGP), and the SDGs.

The rest of the paper is structured as follows: Section two traces the evolution of the concepts of Big Data and Data Revolution and the opportunities they present in the 21st century. A summary of the development challenges facing Nigeria and the national plans and strategies adopted to address them is presented in Section three while in Section four the role of Big Data and Statistics in promoting sustainable development in Nigeria is discussed. Section five presents some emerging opportunities for Statisticians in the context of the Big Data and Data Revolution. Section six concludes.

2. The 21st Century Data Revolution and Big Data

错误! 文档中没有指定样式的文字。 2.1 The Data Revolution

Historically, private businesses, governments and other organized entities have relied on data to make critical policy and strategic decisions, as well as implement and monitor progress towards set objectives. A few decades ago, data collection, collation and processing was viewed as resource-intensive and time consuming, and, in most cases, concept led- following only set rules of procedure, standards and norms. Moreover, data processing, interpretation and communication was often plagued with many challenges and the results of data analysis were available only to a limited number of people. Good quality data was therefore, a closely guarded commodity and traded, more often than not, expensively. Today however, with the advent of new and improved technology, the process of data gathering, processing and dissemination has changed dramatically. While the value of data has not diminished, the conventional methods of collection, processing and dissemination are constantly being challenged. Recent technological innovations have challenged the *status quo* regarding the generation, storage, analysis and utilization of data. As opposed to being scarce and expensive, a lot of data is now readily available and easily accessible (Kitchin, 2014). The sheer volume -referring to access to higher and more unstructured data; velocity -referring to the fast pace at which it is received; and variety -referring to the different types of data available, has led many scientists to identify the 21st century as a period of “Data Revolution”. Needless to say, the ‘Data Revolution’ was triggered by advances in the Information and Communication Technologies (ICTs) sector. Mobile devices, cloud computing and the internet of things are increasingly reshaping the way businesses are run; the way knowledge is generated and disseminated; and the way societies are governed. These changes are also influencing, in a profound manner, consumption and production patterns; communication; travel; and even leisure.

2.2 Big Data

The speed at which data is currently generated, its spread across networks and the vastness of data dynamics has led to the coining of term ‘Big Data’. (Note 1) The term ‘Big Data’ was first used in the 1990s to describe massive volumes of information (Skinner, 2013). Big data involves unconventional sources of data, often large and unstructured; data obtained from new sources such as internet feeds and clickstreams on webpages.

Eric Schmidt (2010) of Google had this to say;

“There were 5 exabytes of information created between the dawn of civilization through 2003, but that much information is now created every 2 days.”

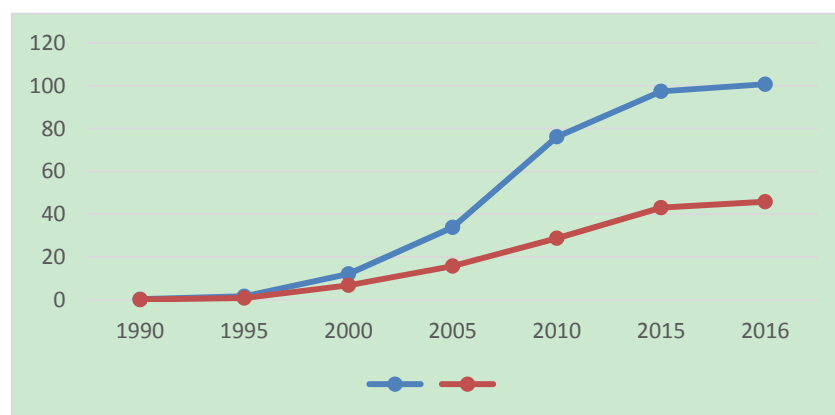


Figure 1. Trends in data availability (1990 – 2016)

Source; World Bank

Figure 1 shows the growth of data in terms of individuals with mobile cellular subscriptions as well as individuals with internet access. Between 1990 and 2016 mobile subscriptions per 100 people and individuals using the internet (as a percentage of population) registered tremendous growth. Consequently, the volume and level of detail of data currently available on individuals, societies and the economy at high speed is unprecedented.

According to Marr, 2015, Big Data can be generated in a variety of forms and from various sources including, but not limited to, the following:

- Activity data: data generated from sources such as listening to music, reading a book, web browsers, smart phones, credit card companies, online stores, etc.
- Conversation data: emails, phone conversations and social media conversations all of which leave digital footprints
- Internet of Things: smart TVs, smart watches, smart fridges, smart alarms, etc.
- Photo and Video data: online uploads of photo and videos, CCTV cameras etc.
- Sensor data: smart phones created with accelerometer to track speed and direction of travel.

In essence, 'Big Data' is characterized by the following four Vs:

- *Volume*: huge amounts of data generated by the second (zettabytes/brontobytes)
- *Velocity*: represents the speed at which it is generated
- *Variety*: different types of data generated, often unstructured- an estimated 80 percent of data generated is unstructured.
- *Veracity*: quality and order in which big data is generated is much less controllable.

2.3 Making Sense of Big Data

A conceptual question arises, however, with respect to how sense can be made of all these large volumes of data from a wide variety of sources. New methods have been designed to cope with the abundance of information as opposed to its scarcity in previous times. Some of these methods include; cloud computing, distributed systems alongside new technology which allows for leveraging and understanding Big Data in an insightful manner.

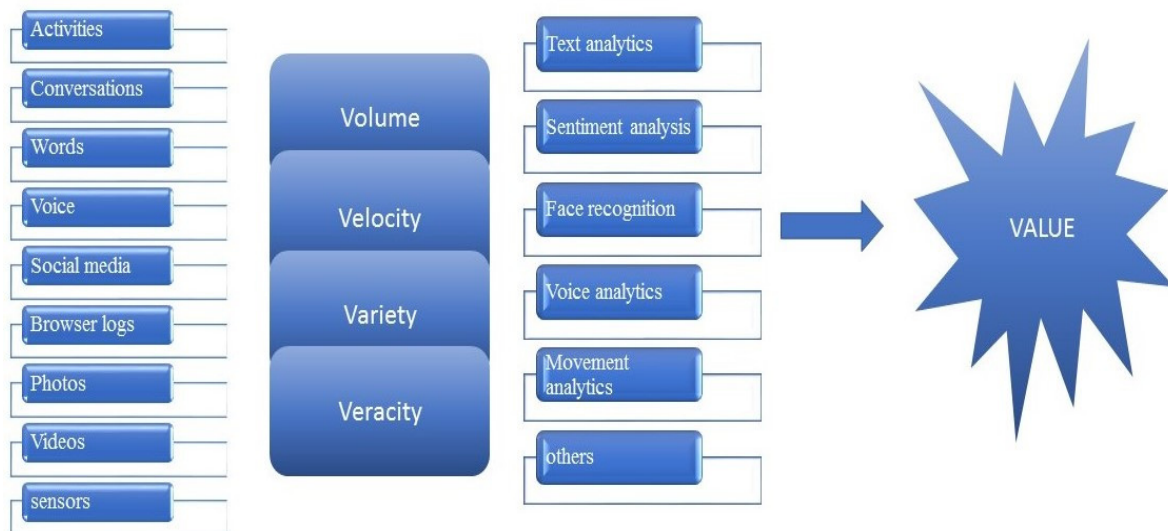


Figure 2. A conceptual framework for datafication (Note 2) of big data

2.4 Big Data, Factor Productivity, Good Governance and Innovation

Big Data has the potential to promote innovation, increase productivity and growth among and within nations and/or industries. In a study of Big Data and its relation to Competition, Innovation and Productivity, the McKinsey Global Institute (2011) (Note 3) noted that Big Data can create value and opportunities in different sectors and different geographic locations. Specifically, they noted that Big Data can lead to:

- an estimated \$300 billion potential annual value to US health care, more than double the total annual health care spending in Spain;
- €250 billion potential annual value to Europe's public sector administration, more than GDP of Greece;
- \$600 billion potential annual consumer surplus from using personal location data globally;
- 60 percent potential increase in retailers' operating margins;
- 140,000–190,000 deeper analytical talent positions, and
- demand for 1.5 million more data-savvy managers to take full advantage of big data in the United States alone

Additionally, utilizing Big Data can:

- improve transparency in public sector by making data and information easily accessible;
- enable experiments that help discover needs, expose variability and improve performance;
- enable analysis of variability in performance to enable leaders understand root causes of inefficiencies and manage performance;
- provide information that enables segmentation of populations for customized actions, ultimately, enabling policy makers to tailor policies to meet precise needs of the people and businesses to tailor their products and services to meet consumer demands and preferences;
- support decision making with calculated algorithms, minimizing risks and identifying loopholes for better results;
- promote competition among private enterprises, fostering innovation and increasing efficiency and value;
- increase operational efficiency by reducing costly errors; and
- create new opportunities for growth opportunities and consequently, increasing economic performance.

In essence, progress in virtually every field or sector is driven by ability to collect, analyze and use data and information. Indeed in 2015, Barack Obama the immediate past President of the USA stated thus: *"understanding and innovating with data has the potential to change the way we do almost anything for the better"*

3. Nigeria: Development Challenges and Strategies

3.1 The Key Development Challenges

After a period of sustained robust growth, the Nigerian economy contracted by 1.8% in 2016 (Note 4), before rebounding back immediately in 2017, growing by 0.82 percent (IMF, 2018). (Note 5) The recent recession could be attributed to a series of external shocks, including the continued decline in oil prices and associated domestic factors such as foreign exchange shortages; disruptions in fuel supply and power shortages; oil pipeline vandalism and insecurity in some parts of the country; as well as low budget implementation.

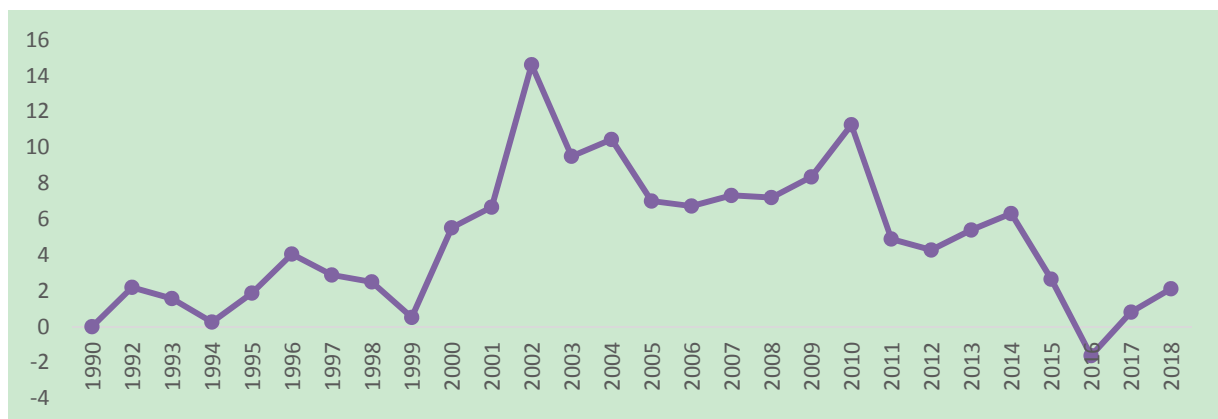


Figure 3. GDP growth rates (%) 1990 -2018

Source: IMF – 2018

As can be seen from Figure 3, prior to the recent contraction of the economy, the country had registered impressive growth, averaging 8.44% per annum during the 2000s. This impressive growth was however, not matched by significant job creation; and reduction in poverty and income inequality.

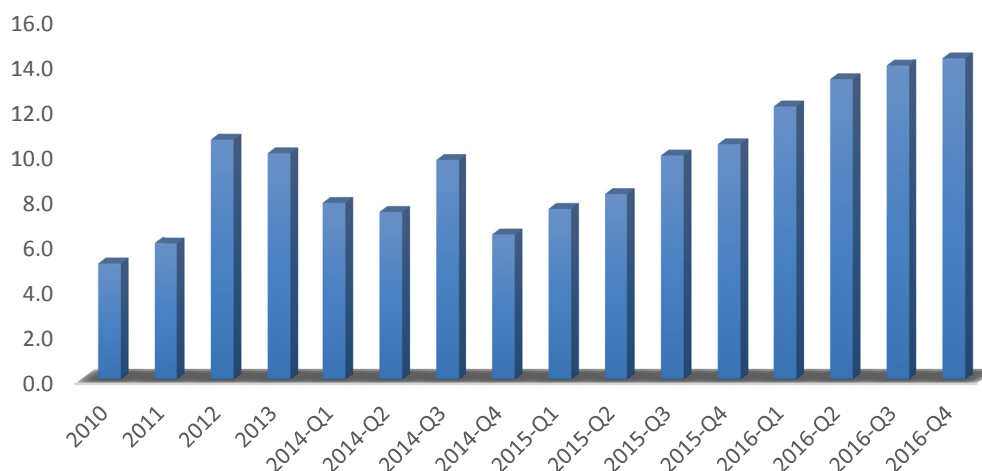


Figure 4. Unemployment trends (2010-16)

Source: National Bureau of Statistics

The rate of unemployment more than doubled over the 2014 -2016 period – from a low of 6.4 percent during Q4 of 2014 to the Q4 of 2016 figure of 14.2 percent. Although estimates by the World Bank have indicated a decline in the poverty rate over the 2004 -2011 period - from a high of 46.42 percent in 2004 to 36.19 percent in 2011, the number of people classified as poor actually increased from 63.7 million to 64.5 million during this period. In essence, poverty reduction was not commensurate with the rapid economic growth registered during this period, with Nigeria registering a Growth Elasticity of Poverty of 0.6 percent (Note 6) as compared to 1.2 percent for sub-Saharan Africa and 2.2 percent for Least Developed Countries during the same period.

3.2 A Summary of Nigeria's Development Plans and Strategies

The Nigerian government has developed, at various points in time, many Development Plans and Strategies in an attempt to address, *inter alia*, the problems of recession and low economic growth; unemployment; poverty and deprivation; and low human development. These Plans and Strategies has emphasized different policy stances and strategic interventions depending on the nature of the most pressing development challenge. For instance, the first Medium Plan of the 1960s was aimed at increasing economic growth and consumption; the economic stabilization measures of 1982 were focused on monetary and fiscal control measures; the Structural Adjustment Programme (SAP) of 1986 focused on reducing fiscal expenditure, consumption patterns as well as trade liberalization, while the National Economic Empowerment Development Strategy (NEEDS), Nigeria's home-grown PRSP of the 2000s, aimed at creating wealth, job creation, reducing poverty and adding value through enhanced public and private participation. More recently, the Government launched the Economic Recovery and Growth Plan (ERGP) covering the period 2017 -2022. The ERGP is aimed at improving macroeconomic stability; fostering economic growth and diversification; improving competitiveness; fostering social inclusion; and enhancing governance and security. The broad objectives of the ERGP are to restore growth; invest in people through social inclusion, job creation and youth empowerment and improved human capital; and build a globally competitive economy through investing in infrastructure, improving business environment and promoting digital-led growth. Its key execution modalities include stabilizing the macroeconomic environment; achieving agricultural growth and food security; ensuring energy sufficiency (power and petroleum products); improving transport infrastructure; and providing infrastructure focusing on small and medium enterprises.

4. Role of Big Data and Statistics in Promoting Sustainable Development

4.1 Role of Government in Big Data

For any country to develop, there is need for it to develop its statistical system and utilize data and statistics for policy formulation, planning and decision making. Data and statistics play an important role in the development

and implementation of national policies and plans and, ultimately, addressing national development challenges. It is with this realization in mind that the United Nations (UN) in 2010 adopted a resolution establishing the World Statistics Day to encourage governments to refurbish their statistical units in order to provide reliable data for better development planning and decision-making. The Nigerian Government established the National Bureau of Statistics (NBS) vide the Statistical Act of 2007 in order to collect, interpret and disseminate data and statistics to the public for evidence-based policy formulation and planning (Doguwa, 2009). The United States Administration under the immediate past President Barrack Obama launched the Big Data Research and Development Initiative worth \$200million with the objective of accelerating the pace of discovery in science and engineering; strengthening national security; and transforming teaching and learning (The Verge, 2018). The initiative seeks to maximize the vast amount of information generated through big data and use it to inform decision making.

At the multilateral level, the emerging role of Big Data in the 21st century led the UN to establish the Global Pulse Department whose core responsibility is to harness data and use it to accelerate development. The Global Pulse initiative is exploring how new digital data sources and real-time analytics technologies can help policymakers improve their understanding of human well-being and emerging vulnerabilities in real-time and better protect populations from shocks. The initiative was borne out of the realization that digital data offers the opportunity to gain a better understanding of changes in human well-being, and to get real-time feedback on how well policy responses are impacting on people's lives. The overarching objective of the initiative is to mainstream the use of data mining and real-time data analytics into development organizations and communities of practice. It presents an opportunity for forging public-private data sharing partnerships, generating high-impact analytical tools and approaches; and driving broad adoption of useful innovations across the UN System. In essence, Global Pulse functions as a network of innovation labs where research on Big Data for Development is conceived and coordinated in order to address 21st century development challenges across the globe.

4.2 Striking Balance between Traditional and Non-Traditional Data Sources

Traditionally, there are many data sources available to government for purposes of formulating policies and strategies and programmes aimed at promoting the sustainable development agenda. These include general administrative data, census and housing surveys; household surveys, agricultural surveys, civil registration and vital statistics, economic statistics and geo-spatial data. These various sources are discussed briefly here under:

- *Census and Housing surveys*: These involve systematic gathering and recording of demographic and housing information from the total population.
- *Household surveys*: Are designed to provide information on demographic and socioeconomic statistics from a randomly selected sample of the total population. Household surveys can also be limited to specific geographic areas.
- *Agricultural surveys*: Provide data on agricultural production systems and practices based on a random sample of the population.
- *General Administrative data*: refer to information collected primarily for administrative or management purposes including health records, tax, educational records, welfare, etc.
- *Civil registration and vital statistics*: refer to Administrative data that records vital events in a person's life including birth, marriage, divorce, adoption, and death.
- *Economic statistics (including labor force and establishment surveys and trade statistics)*: These record the performance of economic agents in relation to global, national and local markets and include measures such as Gross Domestic Product (GDP), Gross National Income (GNI), national poverty levels, household income, employment status, domestic revenues, private flows, Official Development Assistance (ODA) and trade statistics.
- *Geospatial data*: Refers to any environmental and socioeconomic data that include location-specific information. Integrating geospatial data with household surveys can enable disaggregation and analysis by spatial characteristics, for example, proximity to roads or levels of urban development.

Despite these wide sources of data and statistics available to government, systemic weaknesses have been observed in statistical systems especially in developing countries like Nigeria (Sustainable Development Solutions Network, 2015). Some of these data are also expensive to collect, a typical example being census data. Many developing countries, Nigeria included, are unable to conduct censuses as regularly as they should primarily due to the high costs involved. This calls for greater innovation in using the other sources of data

discussed in the preceding section. Big Data as discussed earlier can provide innovative means of gathering and processing vast information from multiple sources in real time. High-resolution satellite imagery, mobile devices, biometric data, smart meters, data mining and crowd-sourced citizen reporting for instance, all have the potential to influence official data collection and monitoring processes more efficiently.

4.3 What Governments can Use Big Data and Statistics for

Policy formulation, planning and programme implementation require comprehensive, accurate and up-to-date data and information. Specifically, data and statistics are important for:

- Policy formulation and planning: enabling government to evolve new development scenarios and pathways and a range of policy choices. Data and statistics can be used to efficiently allocate scarce resources and target those sectors, geographic areas and populations most in need.
- Service delivery: assisting government to improve existing services and design new service delivery mechanisms, monitoring performance and providing feedback. Reliable data enables effective regulations and a people-centered approach to service delivery to communities.
- Citizen engagement: governments can be more responsive to the needs of citizens and can be held accountable for their actions.
- Monitoring and Evaluation: to determine if project objectives have been or can be achieved.

Big data, digitization and upgrading traditional uses of information all seek to make data easily available for all to use in everyday life including finance, businesses, governance, civil participation, education and health as well as areas and sectors of the economy. There is therefore, a need for both the demand side (policy and decision makers) and the supply side (statisticians) to dialogue, continuously, to ensure that data and statistics generated are of use for policy and decision making (Shangodoyin and Lasisi 2011).

4.4 Big Data and the Sustainable Development Goals

The SDGs are a set of universal goals to address the social, economic, political and environmental challenges facing humanity. They embody the aspirations of the global community to realize zero hunger, no poverty, quality education, good health and wellbeing, gender equality, affordable and clean energy, clean water and sanitation, increased employment, address climate change, achieve inclusive economic growth among and within nations, and promote peace and security by 2030. Progress towards the 17 SDGs and 169 targets is monitored using a set of 232 indicators, as nationally contextualized. Needless to say, the comprehensive and ambitious nature of the goals calls for accurate, reliable and up-to-date data integrating new data, such as digital data, with traditional data sources to provide detailed statistics and information. Although some commentators have argued that a significant proportion of the world's population still has no access to digital data, the situation is rapidly changing because of access to digital devices. The UN Global Pulse 2018 established that access to digital devices is growing faster than access to clean water in many parts of the world. It is also estimated that growth in internet use in Africa, especially on mobile phones, could increase twenty-fold in the next five years. (Note 7) By providing solutions in real time and at lower cost, Data Revolution and Big Data could support the implementation of policies and programs aimed at accelerating progress towards the SDGs. But for this to happen, there is a need to move away from the focus on traditional data collection and processing methods to developing programs in a way that takes advantage of all available data and information. Table 1 below shows some of the ways in which Big Data can be applied across each of the 17 goals to formulate better policies, make better decisions, empower people and promote greater participation and accountability, leading to better outcomes for people and planet.

Table 1. Big data and SDGs

S/N	SDGs	Possible Use of Big Data
1	No Poverty	Spending patterns on mobile phone services can provide proxy indicators of income levels
2	Zero Hunger	Crowdsourcing or tracking of food prices listed online can help monitor food security in near real-time
3	Good Health and Well-Being	Mapping the movement of mobile phone users can help predict the spread of infectious diseases
4	Quality Education	Citizen reporting can reveal reasons for student drop-out rates
5	Gender Equality	Analysis of financial transactions can reveal the spending patterns and different impacts of economic shocks on men and women
6	Clean Water and Sanitation	Sensors connected to water pumps can track access to clean water
7	Affordable and Clean Energy	Smart metering allows utility companies to increase or restrict the flow of electricity, gas or water to reduce waste and ensure adequate supply at peak periods
8	Decent Work and Economic Growth	Patterns in global postal traffic can provide indicators such as economic growth, remittances, trade and GDP
9	Industry, Innovation and Infrastructure	Data from GPS devices can be used for traffic control and to improve public transport
10	Reduced Inequality	Speech-to-text analytics on local radio content can reveal discrimination concerns and support policy response
11	Sustainable Cities and Communities	Satellite remote sensing can track encroachment on public land or spaces such as parks and forests
12	Responsible Consumption and Production	Online search patterns or e-commerce transactions can reveal the pace of transition to energy efficient products
13	Climate Action	Combining satellite imagery, crowd-sourced witness accounts and open data can help track deforestation
14	Life Below Water	Maritime vessel tracking data can reveal illegal, unregulated and unreported fishing activities
15	Life on Land	Social media monitoring can support disaster management with real-time information on victim location, effects and strength of forest fires or haze
16	Peace, Justice and Strong Institutions	Sentiment analysis of social media can reveal public opinion on effective governance, public service delivery or human rights
17	Partnerships for The Goals	Partnerships to enable the combining of statistics, mobile and internet data can provide a better and real-time understanding of today's hyper-connected world BIG DATA SDGs How data science and analytics can contribute to sustainable development

Source: UN Global Pulse 2016

4.5 Challenges and Risks Associated with Big Data

The use of Big Data in supporting the implementation, and especially the monitoring and evaluation of the SDGs however, faces a number of challenges which will need to be addressed. These include, but are not limited to:

- *Dearth of high-quality data:* Despite recent developments in harnessing data from many sources – traditional and non-traditional – many countries still have access only to non-comprehensive, poor quality and outdated data.
- *Accuracy, Accessibility and Comparability.* High quality data, relevant to the issue at hand and accessible to those who need them are required to achieve the SDGs. Data should be relevant, accurate, timely, accessible, comparable and produced free of political interference. Data comparability is critically important because it allows for comparison across geographic space and over time in a consistent manner. Accessibility is also of significance because often, access to data and information can be restricted by legal, technical or political barriers.
- *Inadequate capture of certain groups and issues:* In many countries (especially the least developed countries of Africa, Nigeria included, and Asia), certain groups such as people living with disability and slum dwellers are not adequately captured. Equally, many issues of concern especially to women are often poorly reported. Only about 50 percent of countries, for example, had records of intimate partner violence and some of these countries' reports were inconsistent. (Note 8) Also, very little data is available on distribution of income within households and economic role of women.
- *The special case of sparse data on environmental issues.* Issues such as deforestation, desertification and chemical pollution in many countries are not properly captured/recorded.

Moreover, the use of Big Data in implementation of the SDGs is not without inherent risks. These include:

- Respect for fundamental human rights of individuals -because much of the data is generated passively, concerns have been raised regarding the implications for individuals' rights to privacy. There is thus a need to ensure that the fundamental human rights of all are respected and upheld in the process of harnessing the opportunities that the Data Revolution and Big Data present.
- The risk that Big Data could be used for harm rather than good - since information about people's movements, likes, dislikes, social interactions, health status, economic status, and so on, can easily be made available, these could be used for malicious purposes rather than the good of the people, societies and humanity.
- The era of Big Data also presents risks of breakdown of trust among institutions and people.
- Use of wrong data and information - people and planet can be harmed if unchecked data and information used for policy making turns out to be wrong.
- Risk of growing inequality - there is inherent risk of growing inequality amongst those who have access to and those without access to data. Many people, especially in developing countries such as Nigeria are already excluded from accessing relevant information due to language, poverty, education, technology, infrastructure, prejudice/discrimination and technological barriers. In Nicaragua, Bolivia and Honduras for example, the price of mobile broadband subscriptions exceeds 10 percent of monthly GDP per capita compared to France and Korea where it is less than 0.1 percent of monthly GDP per capita. (Note 9)

5. Big Data and Data Revolution – Emerging Opportunities for Statisticians

Demand for statisticians has been on the rise in the recent past and statistics could become one of the most in-demand professions in the world (Horton, 2015). This surge in demand cuts across a variety of statistical fields and opportunities abound for statisticians, especially techno-savvy ones. According to the Bureau of Labour Statistics (BLS) projections, demand for statisticians in the USA will increase by an estimated 33 percent between 2014 -2024. While this projection might not be applicable in Nigeria and other African countries, it is a useful pointer to the emerging global trend. In 2014, LinkedIn analysis established that out of its 330 million-member profiles, statistical analysis and data mining were the most sought-after skills.

Avanade (2014) has reported that 75 percent of companies will increase investments in Big Data in the coming years while CIO (2012) reported that 65 percent of companies deploy Big Data technology to boost speed and quality of business decisions and over \$500million in venture capital funds have gone into big data technology startups and vendors in recent years.

Some recent studies however, have shown that there exists shortage of talent in statistics. The talents can be grouped into three major areas:

- Deep analytical talent: individuals with advanced knowledge of statistics and machine learning and can analyze large sets of data

- Big data savvy talent: individuals with the skills to be effective consumers of big data insights, i.e., to ask the right questions for analysis, interpret and challenge the results, and take appropriate decisions
- Supporting technology: personnel who develop, implement, and maintain the hardware and software tools needed to make use of big data including databases and analytic programs

Big Data talent is grouped into deep analytical, big data savvy and supporting technology

	Deep analytical	Big data savvy	Supporting technology
Definitions	People who have advanced training in statistics and/or machine learning and conduct data analysis	People who have basic knowledge of statistics and/or machine learning and define key questions data can answer	People who serve as database administrators and programmers
Occupations (Note 10)	Actuaries Mathematicians Operations research analysts statisticians Mathematical technicians Mathematical scientists Industrial engineers Epidemiologists Economists	Business and functional managers Budget, credit and financial analysts Engineers Life scientists Market research analysts Survey researchers Industrial-organizational psychologists Sociologists	Computer and information scientists Computer programmers Computer software engineers for applications Computer software engineers for system software Computer system analysts Database administrators

Figure 5. Big data talent categories and related professions

Source; US Bureau of Labour Statistics; Mckinsey Group Institute 2011

It is projected that demand will increase in all these skill sets. As huge volumes of data become more available, demand for individuals with deep analytical skills is expected to increase, and they will be expected to provide solutions that can be easily understood by users. In order to do this, the data savvy individuals will also be required to ask questions that will drive direction of analysis and provide relevant answers to those questions. Figure. 6 below shows projections in demand for deep analytical skills in the US verses the supply by 2018. As is evident from the figure, demand is projected to outstrip supply by 50 -60 percent, pointing to an acute shortage of skilled people to perform deep analytics.

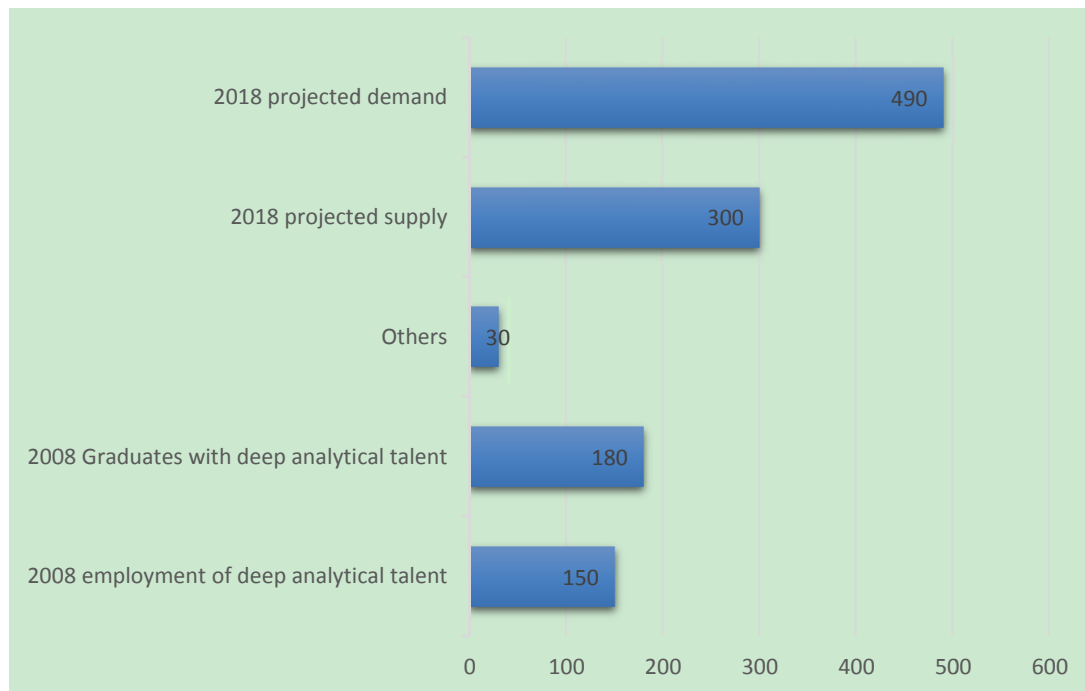


Figure 6. Supply and demand for deep analytical talent by 2018 (thousand people)

Source; US Bureau of Labour Statistics; Mckinsey Group Institute 2011

Based on the foregoing, an important question arises: how do Statisticians in Nigeria and other African countries re-tool, re-skill and re-organize themselves to address the emerging skills gap and take full advantage of these new opportunities? The answer to this question lies first in a shift from conventional thinking around statistics to cutting edge statistical analysis and innovation. Statisticians need to be more innovative, thinking outside the box and creating new ways of gathering data and more importantly, processing and interpreting them in ways that place them at the forefront of the Data Revolution. Statisticians also need to be flexible and adapt to new methods. Understanding the power of innovation is key to the future success of Statisticians. They need to recognize the fundamentals needed for innovation which include:

- A tolerance for risks; risks do not have to be big, only different. - risks are synonymous to innovation.
- Creative ideas; ideas are the basis for innovation. Ideas do not need to be big, complex or even original, they just need to address an identified problem. And to find an idea, one needs first to identify flaws.

Statisticians also need to be more entrepreneurial. They must master both technical and non-technical skills such as leadership skills, communication skills, negotiations, diplomacy and ability to debate analytical techniques (Manamley *et al.* 2016). They should develop a lifelong yearning for learning. Also, as important as data collection, interpretation and gathering is the need for transparency. Information isn't useful unless it is easily available. Statisticians need to understand this and explore the potential advantage it can bring. Sharing of information among companies and researchers could soon become the norm, thus, the need to move away from traditional data "gatekeeping" to "data facilitation". Data transparency could lead to major breakthroughs in science, research and development, economic decision making and in everyday life. From a policy advocacy perspective, there is a need to engage the relevant authorities with a view to promoting the critical role of and emerging opportunities for Statisticians in sustainable development. In particular, there is a need for government to engage statisticians and adequately re-skill and re-tool them for the emerging opportunities.

6. Conclusions

Data and information from a multiplicity of sources- both conventional and unconventional- have become integrated into almost all aspects of present-day life and are an important driver of growth and sustainable development. Data generated from a multiplicity of sources can provide a lot of information on the way people live their lives, their preferences and the choices they make. Many businesses utilize these to align their products and services towards individual preferences. Governments also use data as the basis for evidence-based policy

making, planning, implementation, monitoring and evaluation of development programmes. Nigeria's ERGP has been developed to achieve long term economic growth and social inclusion for all, as well as most of the SDG targets. However, the path to Nigeria's sustainable development is lengthy and windy, requiring commitment, dedication and targeting of development programmes. It will be impossible to achieve the SDGs in Nigeria without asking the right questions and accessing relevant information, hence the need for comprehensive, quality and up-to-date data. Data will need to be collected from various sources (administrative, social media, telephone companies, banking sector, etc.) and analyzed to provide statistics that can be used for decision making. Statisticians therefore, have an essential role to play in Nigeria's development process and the attainment of the SDGs. Big Data and Data Revolution present boundless opportunities for Nigerian Statisticians because data analysis is required in almost every occupation and statistics will be one of the most sought-after professions in the coming years. With rapidly changing technology however, Statisticians need to evolve from conventional to more advanced and rigorous analytical methods. Adaptation, re-tooling, re-skilling and greater innovation by Statisticians are thus key to effectively tapping into the emerging world of information and technology for sustainable development in Nigeria.

References

- Arthur, C. (2013). Tech Giants might be huge, but nothing matches Big Data. *The Guardian*. Retrieved March 11, 2019, from <https://www.theguardian.com/technology/2013/aug/23/tech-giants-data>
- Avanade.com. (2019). *Analytics and Artificial Intelligence (AI) | Avanade UK*. Retrieved August 7, 2018, from <https://www.avanade.com/en-gb/solutions/analytics-and-ai>
- Bls.gov. (2019). *Databases, Tables & Calculators by Subject*. Retrieved from <https://www.bls.gov/data/>
- Data Revolution Group. (2014). *A World that Counts - Mobilising the Data Revolution for Sustainable Development*.
- Data.imf.org. (2018). *IMF*. Retrieved from <https://data.imf.org/>
- Data.worldbank.org. (2019). *World Bank Open Data | Data*. Retrieved from <https://data.worldbank.org/>
- DFID. (2002). *Statistics Matter: Eliminating World Poverty*.
- Doguwa, S. (n.d.). *Statistics for National Development, 1(1)*.
- ECLAC. (2014). *Latin American Economic Outlook 2013: SME Policies for Structural Change*. p. 124. Retrieved from http://www.cepal.org/publicaciones/xml/5/48385/leo2013_ing.pdf
- Einav, L., & Levin, J. (2014). The Data Revolution and Economic Analysis. *Innovation Policy and the Economy, 14*, 1-24. <https://doi.org/10.1086/674019>
- Forbes.com. (2018). Retrieved July 16, 2018, from <https://www.forbes.com/sites/ciocentral/2012/04/24/the-web-is-much-bigger-and-smaller-than-you-think/#6570038e7619>
- Horton, N. (2015). Challenges and Opportunities for Statistics and Statistical Education: Looking Back, Looking Forward. *The American Statistician, 69(2)*, 138-145. <https://doi.org/10.1080/00031305.2015.1032435>
- Kitchin, R. (2014). *The Data Revolution* (1st ed.). UK: SAGE publications Ltd.
- Manamley, N., Mallett, S., Sydes, M., Hollis, S., Scrimgeour, A., Burger, H., & Urban, H. (2016). Data Sharing and the Evolving Role of Statisticians. *BMC Medical Research Methodology, 16(S1)*. <https://doi.org/10.1186/s12874-016-0172-9>
- Marr, B. (2015). *Big Data; Using Big Data Analytics and Metrics to make better decisions and improve performance*. West Sussex: John Wiley & Sons Ltd.
- Mckinsey Group Institute. (2011). *Big data: The Next Frontier for Innovation, Competition, and Productivity*. Retrieved July 16, 2018, from <http://www.mckinsey.com/mgi>
- Shangodoyin, D., & Lasisi, T. (2011). The Role of Statistics in National Development with Reference to Botswana and Nigeria Statistical Systems. *Journal of Sustainable Development, 4(3)*. <https://doi.org/10.5539/jsd.v4n3p131>
- Skinner, C. (2013). Big Data as Competitive Weapon. In: *BAI - Banking Strategies*. Bank Administration Institute. Retrieved July 16, 2018, from <https://www.bai.org/banking-strategies/article-detail/big-data-as-competitive-weapon>
- Statista. (2018). *Most valuable companies in the world 2018 | Statista*. Retrieved July 16, 2018, from

- <https://www.statista.com/statistics/263264/top-companies-in-the-world-by-market-value/>
- Sustainable Development Solutions Network. (2015). *Data for Development: A Needs Assessment for SDG Monitoring and Statistical Capacity Development*.
- The Verge. (2018). *Obama administration announces \$200 million 'big data' research and development initiative*. Retrieved July 16, 2018, from <https://www.theverge.com/2012/3/29/2912137/obama-big-data-research-initiative>
- Un.org. (2018). *Statistics shaping social and economic development | UN DESA | United Nations Department of Economic and Social Affairs*. Retrieved June 11, 2018, from <http://www.un.org/en/development/desa/news/statistics/statistics-shaping.html>
- Unglobalpulse.org. (2018). *Data Is at The Centre Of the Global Agenda | United Nations Global Pulse*. Retrieved June 11, 2018, from <https://www.unglobalpulse.org/news/data-centre-global-agenda>
- Walia, A., Rands, K., Kong, G., Adams, D., French, J., Berry, A., ... Hoekzema, M. (2018). *Big Data news, analysis, how-to, opinion and video*. Retrieved July 16, 2018, from <https://www.cio.com/category/big-data/>
- World Bank (2016). *Poverty Reduction in Nigeria in the Last Decade*. Poverty Work Program. Washington D.C. USA.

Notes

- Note 1. By some estimates, 90 percent of the world's current data was generated in the last 10 years alone.
- Note 2. Refers to the process of turning many aspects of our life into computerised data and transforming these into value.
- Note 3. See *Big data: The Next Frontier for Innovation, Competition and Productivity, May 2011*. McKinsey Global Institute. Available at [www](http://www.mckinsey.com).
- Note 4. The only other exception was 1991, when the economy contracted by 0.55% and stands out in stark contrast to the peak growth of 14.6% registered in 2002.
- Note 5. The economy is projected to grow by 2.2 percent in 2018.
- Note 6. A GEP of 0.6 percent implies that for every one percent growth in GDP per capita, poverty declined by 0.6 percent.
- Note 7. Unglobalpulse.org. (2018). *Data Is at The Centre Of the Global Agenda | United Nations Global Pulse*. [online] Available at: <https://www.unglobalpulse.org/news/data-centre-global-agenda>
- Note 8. Data Revolution Group (2014). *A World that Counts - Mobilising the Data Revolution for Sustainable Development*.
- Note 9. ECLAC (2014). "Latin American Economic Outlook 2013: SME Policies for Structural Change", p. 124 http://www.cepal.org/publicaciones/xml/5/48385/leo2013_ing.pdf
- Note 10. Occupations are classified by the US Bureau of Labour Statistics.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).