A BUSINESS INTELLIGENCE SYSTEM USING SOCIAL MEDIA DATA

BY

EBOLO CHRISTIANA ISI

PSC1707444

A PROJECT SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE AWARD OF BACHELOR OF SCIENCE (BSc.) DEGREE IN COMPUTER SCIENCE TO THE DEPARTMENT OF COMPUTER SCIENCE, FACULTY OF PHYSICAL SCIENCES. UNIVERSITY OF BENIN

DECEMBER 2022

ATTESTATION

I, EBOLO CHRISTIANA ISI, an undergraduate student in the department of Computer Science, faculty of Physical Sciences, University of Benin, Edo State, with matriculation number PSC1707444 attest to have done this project in partial fulfilment of the requirements for the award of the Bachelor of Science (BSc.) Degree in Computer Science, University of Benin.

_____ _____

EBOLO CHRISTIANA ISI Signature/Date

(Project Student)

CERTIFICATION

This is to certify that EBOLO CHRISTIANA ISI, an undergraduate student in the department of Computer Science, faculty of Physical Sciences, University of Benin, Edo State, with matriculation number PSC1707444 did this project in partial fulfilment of the requirements for the award of the Bachelor of Science (BSc.) Degree in Computer Science, University of Benin; under my supervision.

_____ _____

PROF G.O, EKUOBASE Signature/Date

(Project Supervisor)

APPROVAL

This project report was prepared by EBOLO CHRISTIANA ISI, an undergraduate student in the Department of Computer Science, Faculty of Physical Sciences, University of Benin, Edo State, with matriculation number PSC1707444 is hereby approved in partial fulfilment of the requirements for the award of Bachelor of Science(BSc.) Degree in Computer Science.

_____ _____

PROF G.O, EKUOBASE Signature/Date

(Project Supervisor)

_____ _____

PROF (Mrs) A.O, EGWALI Signature/Date

(Head Of Department)

DEDICATION

This project is dedicated to God Almighty, who is and forever will be my source of courage and strength.

ACKNOWLEDGEMENT

My sincere gratitude goes to God Almighty, for granting me the grace and mental prowess to complete this project. This project completes another milestone in my academic career. I sincerely appreciate the continuous support of my parents, Dr. and Mrs. Jimmy Ebolo.

I also wish to appreciate the Head of Department, Computer Science, Prof. (Mrs) A. O. Egwali. It is pertinent at this juncture to appreciate the efforts of my project supervisor, Prof. G. O. Ekuobase, for his support and guidance throughout the course of this project.

I also wish to appreciate the lecturers of the department of Computer Science, Prof. (Mrs.) V. A. Akwukwuma, Prof. (Mrs.) F. A. Egbokhare, Prof. A. A. Imianvan, Prof. G. O. Ekuobase, Prof. (Mrs.) A. O. Egwali, Prof. F. I. Amadin, Prof. K. C. Ukaoha, Prof. (Mrs.) S. Konyeha, Engr. Dr. F. A. U. Imouokhome, Dr. (Mrs.) V.I. Osubor, Dr. F. O. Chete, Mr. P.E.B. Imiefoh, Mr. E.E. Obasohan, Dr. (Mrs.) G. Aziken, Mr. E. Nwelih, Mr. S.O.P. Oliomogbe, Mrs. A. R. Usiobaifo, Mr. E. C. Igodan, Dr. F. O. Oliha, Dr J. C. Obi, Dr. E. P. Ebietomere, Dr. (Mrs.) R. O. Osaseri, Mr. K. O. Otokiti, Miss I.O. Usiosofe, Mrs. T. Agenmomen, Mr. F. Osagie, Mr. 1. E. Obayagbona. for their relentless service to the students of the Department.

ABSTRACT

The research project aim to take advantage of all available information for analyzing data as a critical component for its success it will be looking at other business intelligence system, their strengths and weaknesses, errors, strategies etc and build on it.

TABLE OF CONTENT

CHAPTER TWO: LITERATURE REVIEW

CHAPTER THREE: RESEARCH METHODOLOGY

LIST OF FIGURE

LIST OF TABLE

CHAPTER ONE

INTRODUCTION

1.0 BACKGROUND OF STUDY

The demand for Business Intelligence (BI) applications continues to grow even at a time when demand for most information technology (IT) products is soft (Soejarto, 2003; Whiting, 2003). Yet, information systems (IS) research in this field is, to put it charitably, sparse. While the term Business Intelligence is relatively new, computer-based business intelligence systems appeared, in one guise or other, close to forty years ago. BI as a term replaced decision support, executive information systems, and management information systems (Thomsen, 2003). With each new iteration, capabilities increased as enterprises grew ever-more sophisticated in their computational and analytical needs and as computer hardware and software matured. In this paper BI systems are defined as follows:

BI systems combine data gathering, data storage, and knowledge management with analytical tools to present complex internal and competitive information to planners and decision makers.

Implicit in this definition is the idea (perhaps the ideal) that business intelligence systems provide actionable information delivered at the right time, at the right location, and in the right form to assist decision makers. The objective is to improve the timeliness and quality of inputs to the decision process, hence facilitating managerial work. Sometimes business intelligence refers to on-line decision making, that is, instant response. Most of the time, it refers to shrinking the time frame so that the intelligence is still useful to the decision maker when the decision time comes. In all cases, use of business intelligence is viewed as being proactive. Essential components of

proactive BI are:

i. Real-time data warehousing

ii. Data mining

iii. Automated anomaly and exception detection

iv. Proactive alerting with automatic recipient determination

v. Seamless follow-through workflow

vi. Automatic learning and refinement

vii. Geographic information systems

viii. Data visualization

1.1 MOTIVATION OF STUDY

Present day association settings are very convoluted and continually evolving. Organizations, in both public and private sector, are under extraordinary pressure for offering an explanation to the top management about change condition and innovation. To do so, it requires an organization to possess strategic, operational and tactical decisions; however, they are complicated and are taken rapidly. The basic leadership requires a lot of data, information, and knowledge. One ought to process this data as required basic leadership and fast, on time and ongoing preparing is expected to end up modernized (Turban et al, 2010). Furthermore, nowadays business life cycle has turned out to be shorter. Henceforth, to gain the competitive advantage organization must have quick and proper decision making. Decision maker indeed needs good data, to make the right decision at the right time and place (Farjami, 2015).

The concept of BI existed during the 1950s and it grew out from a technology called decision

support. Decision support is still used by many companies to come up with decisions that would help them to gain competitive advantage amongst their competitors.

BI has grown strong during the recent years mainly due to increased data collection and better technology with greater storage capacity. Due to the improvement of technology, the company can use BI to store a large amount of data with cheaper rate. Companies have access to a lot of data in the form of smartphone, internet records, and social media activities and so on. BI can sift through these data to find patterns and trends (Raisinghani, 2004).

In any organization regardless its size, the business activities include the administration of extensive amounts of data from both inner and outside business conditions; all these data identified with interior operations, advertise, clients, providers, economic assets, and so forth., historically cumulated, on action times of the organization, shape the reason for some complex and greatly helpful economic and money related problems in the organization's administration decision making process (Mihaelia and Rozalia, 2012).

1.2 AIM

The aim of this project is to augment a business intelligence system using social media data.

1.3 OBJECTIVES

The objectives of this study are;

i. To support the process of decision-making.

ii. Designed to help present and manipulate vital data into information that will support the improvement of the decision-maker business and activities.

iii. To ensure smooth running and management of decisions taken by the decision-maker and organization.

1.4 SCOPE OF RESEARCH

This scope of study is designed to enlighten organizational decision makers of the following levels of organizational decision-making:

(a) Operational

(b) Strategic

(c) Tactical

These decision makers are IT professionals, Chief Information Officers and Chief Technology Officers alike who require the efficient and effective analysis of data "in order to better understand the situation of their business and improve the decision-making process.

1.5 SIGNIFICANCE OF STUDY

The significance of this study is to build on existing business intelligence systems that is being used by organizations to make informed decisions to satisfy customer needs with the aid of Extract, Transform And Load tools.

CHAPTER TWO

LITERATURE REVIEW

Business intelligence (BI) is a term used to describe a range of software applications and technologies that are used to collect, integrate, analyze, and present data. The goal of BI is to help organizations make better-informed business decisions by providing them with accurate and up-to-date information.

The concept of business intelligence dates back to the early 1900s, when companies first began using data to make strategic decisions. In the 1940s and 1950s, the development of electronic computers paved the way for more advanced data analysis techniques, such as statistical modeling and predictive analytics.

In the 1960s and 1970s, the emergence of new technologies, such as computer-aided software engineering (CASE) and data warehousing, enabled organizations to collect, store, and manage large amounts of data more effectively. The 1980s and 1990s saw the rise of executive information systems, which were designed to provide high-level executives with easy access to key performance indicators and other important data.

Today, business intelligence systems have evolved to include a wide range of tools and technologies, including data mining, online analytical processing (OLAP), and dashboards. These systems are used across a wide range of industries, from finance and healthcare to retail and manufacturing, to help organizations make more informed decisions.

2.0. Concepts of Business Intelligence Systems

1. **Data integration:** the process of combining data from multiple sources into a single, comprehensive view.

2. **Data warehousing:** a database specifically designed for storing and analyzing large amounts of data.
3. **Data mining:** the process of discovering patterns and relationships in large datasets.
4. **Analytics:** the use of statistical, mathematical, and computational techniques to analyze and interpret data.
5. **Visualization:** the use of graphs, charts, and other visual aids to help users understand and interpret data.
6. **Reporting and dashboards:** the generation of reports and summaries of data, often presented in graphical form on a dashboard.
7. **Collaboration and sharing:** the ability to share data and insights with others within the organization, enabling collaboration and decision making.
8. **Predictive modeling:** the use of algorithms and statistical models to forecast future trends and outcomes based on historical data.
9. **Data governance:** the policies, processes, and systems that are used to manage and control access to data within an organization.
10. **Business intelligence strategy:** the overall plan and approach for leveraging data and analytics to support decision making and drive business value.

BI is characterized as frameworks that gather, change, and present organized information from various sources lessening the required time to acquire significant business data and enable their efficiency use in management decision making process (Hamer, 2004), permitting dynamic enterprise information look, recovery, examination, and clarification of the necessities of administrative choices (Nofal and Yusof, 2013). As indicated by (Tyson, 1986), BI concentrates on gathering, process and present information concerning customers, contenders, the business sectors, technology, and products. Pirttimäki (2007) depicts BI as a procedure that incorporates a series of activities, being driven by the particular data needs of decision makers and the

objective of achieving competitive advantage.

BI is a framework that transforms information into data and afterward into learning, consequently enhancing company's basic decision-making process (Singh and Samalia, 2014). BI is characterized as a framework which gathers, changes and shows organized information from various sources. BI is a system and an answer that helps decision makers to comprehend the economic circumstance of the firm (Nofal et al., 2013). BI is termed to as a set of numerical and methodological models for examination utilized for extracting data and valuable information from raw information for utilizing confused basic leadership prepare (Vercellis, 2013). Similarly, Wixom and Watson (2010, p.14) mention that ―Business intelligence (BI) is a broad category of technologies, applications, and processes for gathering, storing, accessing, and analyzing data to help its users make better decisions.

BI helps administrators by breaking down information from various resources in better basic leadership at both tactical and strategic level, for customary utilization, conventional data frameworks farewell, yet for hierarchical and functional planning; new tools are required for business analysis (Rasoul and Mohammad, 2016).

2.1. Data, Information and Knowledge

In BI context, we always see the word data, information, and knowledge which could lead us getting confused on its use and implication. Carlo (2009) distinguishes their definition.

i. **Data**: It refers to a structured codification of single primary entities and as well as of transactions involving two or more primary entities (Carlo, 2009). BI is popular among companies mainly because of analysis of data that is of any form and formulate a strategy accordingly. Generally, data is classified into three types—structured data, semi-

structured data, and unstructured data.

Structured data are information that is fixed form, the data may be a collection of forms of websites, and detailed address that can be easily read by the computers since the data is already standardized. Unstructured data are information that cannot be easily read by computers, which may be text, documents, video tapes, websites, and pictures (Jermol et al. 2003), or any other type of information that cannot be clearly sorted or organized into rows and columns. Information is used many times to Company data are found across different locations and places in the form of Customer Relation Management (CRM) programs, marketing automation systems and social media platforms.

ii. **Information:** It refers to the result of extraction and processing activities carried out on data, and it appears meaningful for those who receive it in a specific domain.

iii. **Knowledge:** It is formed from information which is used to make decisions and develop the corresponding actions. Hence, we could say that knowledge consists of information that puts to work into a specific domain, and it is enhanced by the experience and competence of decision makers in tackling and solving complex problems.

2.2. BUSINESS INTELLIGENCE ARCHITECTURES

Carlo (2009) uses the following pyramid to describe how business intelligence system is constructed.

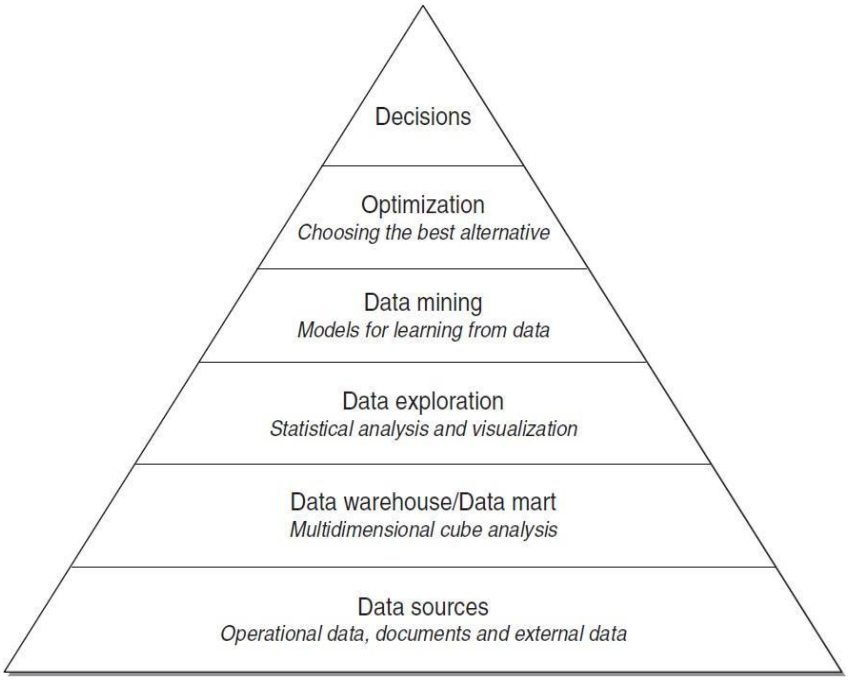


Figure 2.1: A Typical Business Intelligence System

i. **Data sources**: The sources mostly consist of data belonging to operationalize systems, but may also include unstructured data, such as emails, and data received from external providers.

ii. **Data warehouse/Data mart**: Data warehouses are used to consolidate different kinds of data into a central location using a process known as *extract, transform and load* (ETL) and standardize these results across systems that are allowed to be queried. Data marts are generally small warehouses that focus on information on a single department, instead of collecting data across a company. They limit the complexity of databases and are cheaper to implement than full warehouses.

iii. **Data exploration**: Data exploration is a passive BI analysis consisting of query and reporting systems, as well as statistical method.

iv. **Data mining**: Data mining is active BI methodologies with the purpose of information and knowledge extraction from data.

v. **Optimization**: Optimization model allows us to determine the best solution out of a set of alternative actions, which is usually fairly extensive and sometimes even infinite.

vi. **Decisions**: When business intelligence methodologies are available and successfully adopted, the choice of a decision pertains to the decision makers, who may also take advantage of informal and unstructured information available to adapt and modify the recommendations and the conclusions achieved through the use of mathematical models.

2.3. ENABLING FACTORS IN BUSINESS INTELLIGENCE PROJECTS

Some factors such as technologies, analytics and human resources that (Carlo, 2009) mentions are more critical than others to the success of a BI project.

2.3.1. Technologies

The crucial enabling factors that have facilitated the development of BI systems in the complex organization and enterprise are hardware and software technologies. This pattern has empowered the utilization of advanced processes which are required to utilize inductive learning strategies and enhancement models, keeping the processing times inside a sensible range. Additionally, it allows the appropriation of best-in-class graphical perception strategies, featuring real-time animations. A further important factor gets from the exponential increment in the limit of mass storage's, again at low costs, enabling any organization to store terabytes of information for business insight analysis. What's more, system network, as *Extranets or Intranets*, has played an essential part in the diffusion inside organizations of data and learning separated from BI.

2.3.2. Analytics

Mathematical model and analytical methodologies play an important role in information advancement and knowledge taking out from the accessible data inside most organizations. The mere visualization of the data according to timely and flexible logical views, plays a relevant role in facilitating the decision-making process, but still, represents a passive form of support. Hence, it is essential to apply more advanced models of inductive learning and optimization in order to achieve active forms of support for the decision-making process.

2.3.3. Human resources

The human resources of an organization are built up by the competencies of those who operate within its boundaries, whether as individuals or collectively. When employees possess the ability of knowledge that could acquire information and then translate it into the practical way, they will have a major influence on the quality of decision-making process. The organization must emphasize the personal skills of its knowledge workers to work out creative solutions and to devise effective action plan if it implements an advanced BI system. Every company could access to available analytical tools equally, but if a company wants to have the competitive advantage over its competitors, it should employ human resources endowed with a greater mental agility and willing to accept changes in decision-making style.

2.4. DISCUSSION ON HOW ECONOMIC SCHOOL OF THOUGHT RELATES TO BI

Jack (1971) states that information is considered to be a classic example of a ―collective good‖, the type of commodity for which private incentives is supposed to lead to under-provision rather than over-provision on the market. According to Richard et al. (1983), information plays two

crucial roles which the first role refers to the physical state of the world that can indicate the quality of the goods one considers acquiring.

The understanding that competitive markets facilitate the efficient production and allocation of resources in a decentralized manner, that is without a complete exchange of information among economic agents (Radner, 2011). In another way, it could be emphasized the premise that economic agents come to markets with diverse information that is not publicly available, or at least only at substantial cost. The mention of information implies the prior existence of uncertainty about something, whether that uncertainty is probabilistic or not. Likewise, Paul (1981) mentions that when each trader is able to access his/her own private source of information, or when traders can acquire information at a cost, the traders' strategic options may be drastically different than in the case where all information is public. The prices vary directly with underlying qualities. Higher prices indicate better quality.

Another perspective from the school of thought under the theme of Game Theory, decision theory is the primary framework that Radner (2011) extended to formalize the theory of teams. Decision theory refers to making choices under uncertainty. The core of the problem is for many players with a common payoff to make a choice under uncertainty with only partial knowledge, so the extension seems natural. Furthermore, according to Radner, a decentralized organization is defined as one with more than one decision maker, in which different decision makers are responsible for different decision variables and make those decisions on the basis of different information, and in which the outcome to the organization depends jointly on the several decisions and on some stochastic environmental variables. Similarly, the concept of BI is to make an effective decision making in the organization, only data or information from one or two departments is not adequate to optimize the overcome of the decision, but the joint data or

information from all related departments is very crucial to gather enough information to make the right decision.

The definition of business intelligence system used for this literature review is: A set of integrated tools, technologies and programmed products used to collect, integrate, analyze, and transform data into information (Koronios and Yeoh, 2010). This information is then used to enable effective business decision making.

2.5. DEFINITIONS OF THE FOUR MOST COMMON COMPONENTS OF A BUSINESS INTELLIGENCE SYSTEM

Business intelligence systems are used for intelligent exploration, integration, aggregation, and a multidimensional analysis of data originating from various information resources… data is treated as a highly valuable corporate resource (Kronos and Yeoh, 2010). And although definitions vary, and business needs dictate the necessity for different components and complexity for a business intelligence system, all business intelligence systems require, at a minimum, four specific components to produce business intelligence. These components are described throughout the larger literature to the degree that they are now taken-for-granted and they include (a) data warehouses, (b) ETL tools, (c) OLAP techniques and (d) data mining (Olszak and Ziemba, 2006). This section defines and details the functioning aspects of each of these four components.

2.5.1. Data warehouses

A data warehouse is a collection of relevant business data that is organized and validated (Cody et al. 2002) so that it can be analyzed to support business decision making. Data warehouses are populated with data that has been extracted from distributed databases, often heterogeneous and, in

some cases, external to the organization which is using it. Data warehouses are subject oriented databases that are integrated into an information system. They are time relevant, meaning that they are snapshots of a point of time within the information system and they are not updatable so as to maintain the integrity of the historical point in which the snapshot of data is taken.

Data warehouses are offline, meaning that they reside on a different system than that of the data of which they are storing a snapshot. These decisions are based on the history of the business that a data warehouse is holding (Schink, 2009). The data warehouse is considered the core component of a business intelligence system (Negash, 2004). In addition to providing the snapshot of historical data, a data warehouse also provides room for the thematic storing of aggregated information, data that has been analyzed by an ETL tool then loaded into the appropriate data warehouse (Olszak and Ziemba, 2003). A well implemented data warehouse is easy to use, allows for quick information recovery, stores more information, improves productivity, allow2s for better decisions, and increases an organization's competitive advantage (Hwang and Xu, 2007). Hevner and March (2005) conclude that the key role of a data warehouse is to provide an understanding of business problems, opportunities, and performance based on compelling business intelligence facilitating decision making.

2.5.2. Extract, Transform, Load (ETL) tools.

ETL tools and processes are responsible for the extraction of data from one or many source systems, as they transform data from many different formats into a common format and then load that data into a data warehouse (Schink, 2009). ETL tools are tasked with extracting information deemed central to the business. They manipulate and present that data into information that is then used for managerial decision making (Arnott et al., 2004).

ETL solutions are divided into three distinct stages that find and convert data from various sources, and insert the resulting product into a data warehouse. These three stages of ETL are;

1. **The extraction stage**: This stage involves obtaining access to data originating from different, often heterogeneous sources. These sources are often distributed across multiple platforms and can be part of a customer's information system (Schink, 2009).

2. **The transformation stage**: This stage transforms the extracted data and is considered the most complex stage of the ETL process. The transformation stage converts the data into the same schema of the data warehouse to which it is to be loaded. The transformation phase is usually performed by means of traditional programming languages, script languages or the SQL language (Olszak and Ziemba, 2006).

3. **The load stage**: The load stage *pushes* the transformed data and loads the data warehouses with data that are aggregated and filtered (Olszak and Ziemba, 2007).

The requirement of a business intelligence system to be able to extract data in different formats from disperse sources, transform them into like formats, and then load them into the appropriate data warehouse has traditionally made the ETL process the most expensive aspect of a business intelligence system (Hevner and March, 2005). The ETL process can do low level analysis and transformation in this data warehouse prior to loading it into the enterprise data warehouses (Castellanos et al., 2009). ETL tools can be written to have more emphasis on one particular aspect of the ETL process over the other.

2.5.3. On-line Analytical Processing (OLAP) Techniques

The origins of On-Line Analytical Processing are rooted in the difficulties encountered when performing data analysis on databases that are constantly being updated during transactions via

other information systems (Airinei and Homocianu, 2009). OLAP attempts to analyze complex data in real time on a database that is constantly updated with transactional data. The OLAP optimizes the searching of huge data files by means of automatic generation of SQL queries (Olszak and Ziemba, 2006).

OLAP allows user access, analysis and modeling of business problems and sharing of information that is stored in data warehouses (Olszak and Ziemba, 2007). As noted by Olszak and Ziemba (2007), OLAP offers techniques for data analysis and drilling data and the tools are mainly used for interactive report generations. Matei (2010) states that OLAP tools use data mining techniques and statistical methods to create readable, fast report generation that is used for forecasting that can further assist in strategic decision making. These reports are generated based on a manager's pre-defined criteria (dimensions).

2.6. THE SPECIFIC ROLE OF EACH COMPONENT IN A BUSINESS INTELLIGENCE SYSTEM

Airinei and Homocianu (2009) describe business intelligence systems as a means to exploit information in order to help managers solve their structured and unstructured problems. Each component of a business intelligence system can be used to exploit information in one or more of these selected managerial decision-making actions: (a) acquiring information; (b) searching/gathering information; (c) analyzing information; and (d) delivery of information

(Olszak and Ziemba, 2007).

By analyzing historical data, business intelligence systems strive to eliminate communication barriers that exist at the different organizational levels within a company. These barriers are

considered *noise* during the decision-making process. By allowing decisions to be made using consistent information (Matei, 2010), this method of analysis enables managers to evaluate former activities and direct future actions.

The managerial decision-making action a particular component of a business intelligence system can support varies based on many factors. These factors include the type of organization using the business intelligence system, the sector a business operates in, and the maturity of the business intelligence system (Rodrigues, 2002). Olszak and Ziemba (2003) present a framework for the actions a manager takes in order to make business decisions. The alignment among actions and their corresponding business intelligence components is summarized in Table 3. Although different components appear to exhibit crossover functions, it is the level of detail and outcome that are very different. A detailed discussion is provided below.

Table 2.1.: BI System Components Aligned with Managerial Decision-Making Actions

Business Intelligence System Component	Managerial Information Actions
ETL Tools	Acquiring/gathering, Searching
Data Warehouse	Acquiring/gathering
OLAP Techniques	Analyzing, Delivery
Data Mining	Analyzing, Delivery

i.

ii.

Acquiring/gathering information.

Acquiring information has become increasingly more difficult as modern organizations adopt more distributed information systems in which to store their business-critical data (Hevner and March, 2007). This action is used to find the business issue. As Olszak and Ziemba (2006) point out, this action utilizes ETL tools, directing the processes to find what information are needed and into which data warehouse to deposit that information (Shi et al., 2006).

Searching information.

After the data are extracted from operational databases (Castellanos et al., 2009), the newly loaded high-quality data are mined using data mining techniques and processes. This action is performed at different levels of data quality. Lower quality data are searched by utilizing ETL tools. The more refined or mature an ETL tool, the higher the data quality of the data being loaded into a data warehouse (Schink, 2009).

Analyzing information.

Managers need to create data models to understand and address business issues. Through data preprocessing and applying OLAP and data mining techniques managers can analyze information from multiple dimensions at varying degrees of granularity, and tasked with a different level of analysis (Shi et al., 2006). For example, information derived through analysis directly affects decisions related to promotional campaigns, forecasting sales and financial results and, in some cases, can be used in fraud detection (Olszak and Ziemba, 2007).

OLAP summarizes data and makes forecasts based on historical data. Data mining discovers hidden patterns in data. Data mining operates at a detail level instead of a summary level. In other words, data mining predicts, while OLAP forecasts.

Data mining and OLAP can be used to analyze:

(a) **Financial Data:** analyzing and reporting on costs, revenue and profitability

(b) **Marketing Data:** analyzing sales receipts, sales profitability, sales target, actions taken by competitors

(c) **Customer Data:** analyzing time of contact, customer profitability, customer behavior, customer satisfaction, and customer loyalty

(d) **Production Data:** analyzing production *bottle necks,* delayed orders, in-process materials, tool up-time

(e) **Logistical Data:** analyzing relationships in a supply chain and delivery partnerships

(f) **Wage Related Data:** analyzing wage types, payroll surcharges, payroll collections, employee contributions, and average wages

(g) **Personal Data:** analyzing employee turnover, employee type, presentation of information related to individual data

Delivery of information.

Data mining is also used in the delivery of information within an organization. In business intelligence systems, data mining can not only interpret, and evaluate results generated from the analysis performed on data stored in a data warehouse, but it can also display reports enabling decision makers to discover various patterns, generalizations, and regularities (Olszak and Ziemba, 2007). In the same way, OLAP creates ad hoc report generation using simpler data mining techniques by summarizing data without the pattern matching that is unique to the data mining process (Matei, 2010). As Olszak and Ziemba (2003) point out, data mining provides a detail-oriented report while OLAP provides a generic summary of information. Without well-defined delivery, management may get extensive reports that are not only inappropriate for the decisions being made at that time but the reports may contain too much information that may cause managers to overlook critical data (Jaklic et al., 2010).

2.7. HOW BUSINESS INTELLIGENCE SYSTEMS CAN BE USED TO BETTER FACILITATE BUSINESS DECISION MAKING AT EACH LEVEL OF MANAGEMENT

By utilizing business intelligence systems organizations are collecting, treating and diffusing information with the objective of reducing uncertainty in the making of decisions Shi et al. (2006). These decisions are often made under pressure, almost always at critical times in which businesses need real-time data.

A business intelligence system allows managers to make decisions using real time data by monitoring competition, carrying out constant analysis of numerous data and considering different variants of organization performance (Olszak and Ziemba, 2007). As figure 2 shows, data is extracted from operational databases, customer databases, and from data collected pertaining to the competition. The business intelligence system extracts this data from these various data sources, transforms it into specified formats, and then loads the newly formatted data into specially designated data warehouses that are available to all three levels of decision making within the organization: operational, strategic, and tactical (Negash, 2004). Each level of the organization will utilize different OLAP techniques and data mining process to analyze data and report information that is most relevant to them. The information generated from the business intelligence system will be used in all decision-making processes. At the strategic level, decisions set objectives and push the decision direction to the tactical level of the organization. At the tactical level information is mined from the business intelligence system to develop tactics to realize the strategic objectives and, in-turn, will push a decision down to the operational level of the organization. Both the tactical and operational levels of management are reactive to the strategic decisions of the organization (Cella et al., 2004). Figure 2 shows how data and decisions flow in an organization:

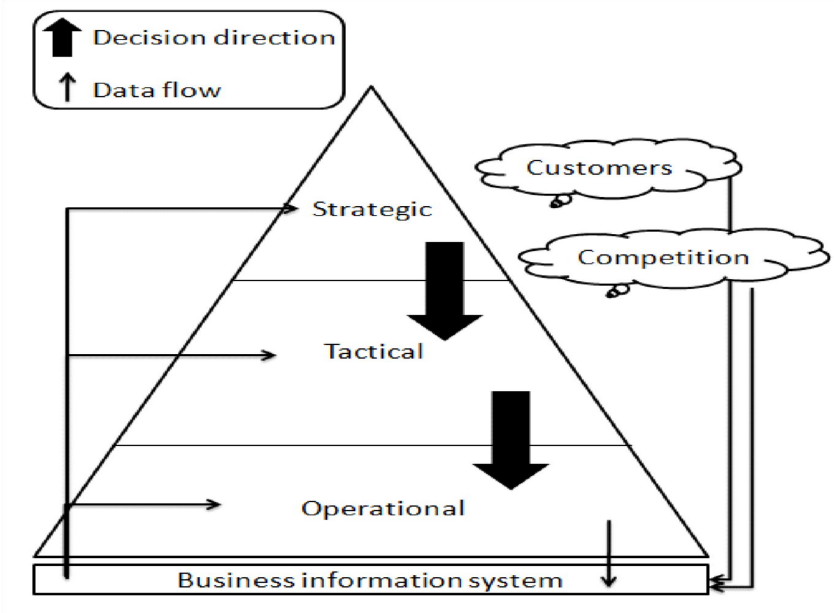


Figure 2.2: Organizational Decision Flow Overview

Even with a shared objective, different levels of the organization will utilize information for different purposes. At strategic and tactical levels, information provides input to senior managers; at the operational levels, information provides input to lower-level managers (Negash, 2003).

Operational level decisions.

At the operational level, decisions affect or are related to the ongoing operations of an organization. These decisions are generally based on up-to-date financial data, sales and co-operation with suppliers and customers (Olszak and Ziemba, 2007). Data are the life blood of daily operations in an organization and business intelligence takes that data and presents it to decision makers in the form of information (Barone, Jiang, Mylopoulus, Won, and Yu, 2010). Business intelligence systems provide information used at the operational level of an organization to address the following specific actions (Olszak and Ziemba, 2006):

1. Identify problems and ‘bottlenecks’

2. Provide analysis of “the best” and “the worst”

3. Provide analysis of products

4. Provide analysis of employees

5. Provide analysis of regions (using measurable metrics such as sales, costs or quantifiable results)

6. Perform ad-hoc analysis and answer questions related to department’s ongoing operations, up to date financial standing and sales.

Operational level decisions are noted as being the decisions that allow an organization to run its day-to-day activities (Esat et al., 2007). The information provided by the business intelligence system is at a summary level and the data feed into the business intelligence system from the operational level of an organization is analyzed and combined with other external information to create direction and allow for strategic planning to occur.

i. Tactical level decisions.

Decisions made at the tactical level are related to planning and rely on real-time data and forecasting to direct the future actions of marketing, sales, finance and capital management. Tactical decisions are often used to support strategic decisions (Olszak and Ziemba, 2007). The literature details these related tactical decision-making activities as being supported by business intelligence systems:

1. Analyses of deviations from the realization of plans for particular organizational units, individuals or indicators

2. Decisions related to the direction of marketing, sales, finance and capital management

3. Forecasting of demand for a given product or service

The information derived through these activities allows for optimizing future actions and for modifying organizational aspects of the company's performance.

ii. **Strategic level decisions.**

Strategic level decisions set objectives as well as ensure that those objectives are realized. Business intelligence systems provide information in support of strategic decision related to the development of future results based on historical results, profitability of offers (made or received) and the effectiveness of distribution channels (Olszak and Ziemba, 2007). Negash (2004) asserts strategic decisions use business information systems to create forecasts based on historical data from the past, combining it with current performance and then to estimate how conditions will play out in the future. Based on the literature, information provided by business intelligence systems inform these kinds of decisions made at the strategic level:

1. Whether to enter new markets

2. The possibility of changing a company's orientation from product-centric to customer centric
3. The launch of a new product (Watson & Wixom, 2007)

4. What objectives to set and to follow through on the realization of such established objectives (Olszak and Ziemba, 2007)

CHAPTER THREE

RESEARCH METHODOLOGY

This chapter explains in detail the method used in extracting data from social media sites to be able to build on existing business intelligence systems.

3.0 Conceptual Framework

In today's modern world of technology, social networks have made an enormous impact. People will openly express and debate their opinions on a topic, making it a valuable source of knowledge. Sentiment analysis, also known as opinion mining, is a technique for identifying people's opinions or reactions to goods, programs, organizations, persons, and incidents.

Figure 1 shows that the system necessarily requires completing a series of critical steps to cover the life cycle of data accumulated from multiple social media sources (Paik et al; 2013) This process aims to analyze this data according to several important components, such as data collection, classification of this data, content analysis, data storage, etc.

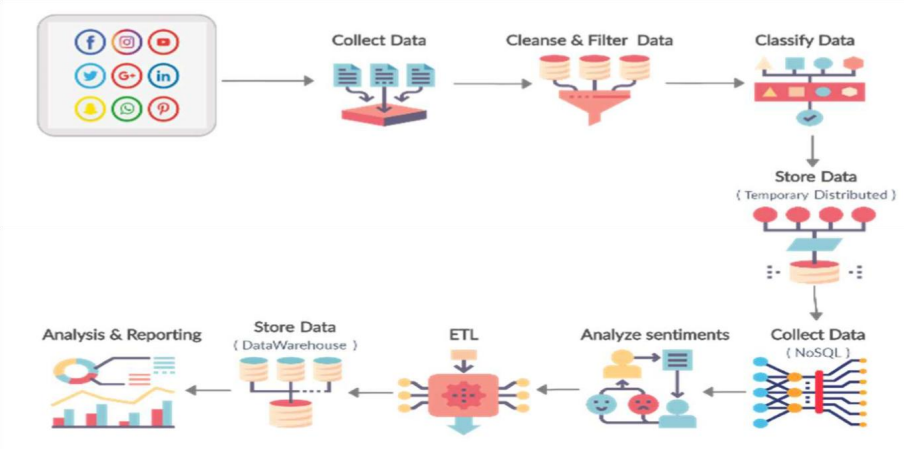


Fig. 3.1. Conceptual Framework.

3.1. Data Collection

Data collection is the first phase in this process because it is a complicated one that requires the use of technologies to gather information from various sources. Organizations may optimize workflow procedures, acquire a clear picture of performance, solve complicated challenges, and comprehend customer behavior with the help of data management. Through many social media sites, like Facebook, Twitter, Instagram, and others, organizations can gather data. For this project, I'll be scraping my data from Twitter. using the Twint Extraction tool

3.2 Twint Extraction Process

Twint is the data collection tool utilized in the research. It is a sophisticated Python-written Twitter scraping tool that enables Twitter profiles to be scraped for Tweets without using Twitter's API.

The Data collected is multiple tweets on twitter about **Airtel Network Ltd**. Airtel Network Limited is a leading telecommunications services provider in Nigeria. It has its headquarters in Lagos, It has been around for over 20 years. This research is conducted to check customer reviews on their service as it will in a long way help to improve their service.

scrape.py

```python
import twint

# Configure
c = twint.Config()
c.Search = "airtel"
c.Lang = "en"
c.Since = "2020-01-01"
c.Until = "2022-12-30"
c.Geo = "9.077751, 8.6774567, 105km"
c.Limit = 300
c.Output = "./test.json"
c.Store_json = True

# Run
twint.run.Search(c)
```

Figure 3.2 Twint Extraction Process

After the extraction process, the data collected were unstructured and saved as **test.json** as shown below in **Figure 3.3**

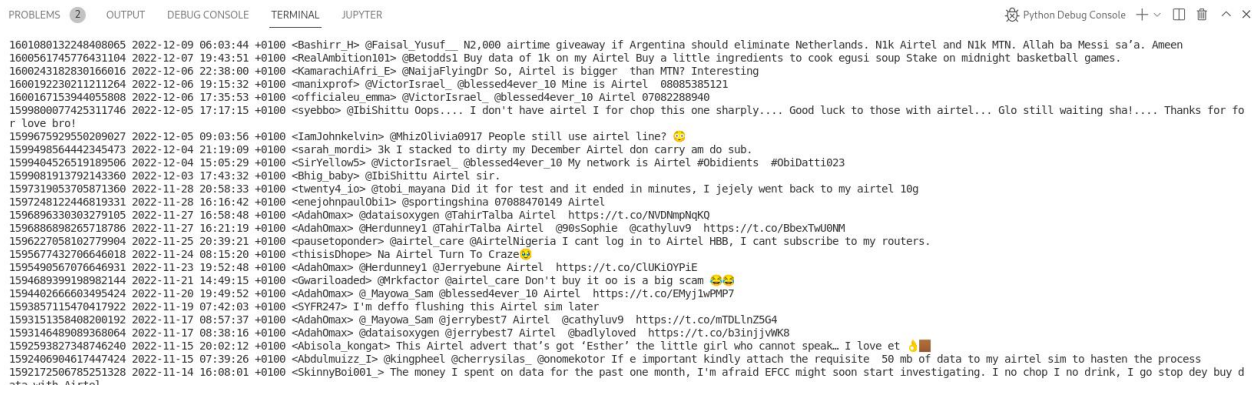


Figure 3.3 Tweets Extracted

3.3. Data Reading and Filtering

The result of the data fetched by **Twint** was an exported json file. The json file was read as a data Frame using a python library called **Pandas.** This read data comes with several columns of both wanted and unwanted data. so therefore there is a need for data filtering .

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from textblob import TextBlob
%matplotlib inline
```

```
# Data Imports
data = pd.read_json('test.json',lines=True)
data.head()
```

	id	conversation_id	created_at	date	time	timezone	user_id	username	name	place	...	geo	source
0	1601953515403894786	1601953515403894784	2022-12-11 15:54:15+01:00	2022-12-11	15:54:15	100	542302673	aruwa_ab	Abubakar Ibrahim Aruwa		...	9.077751,8.6774567,5km	
1	1604235232907038721	1604230125633306624	2022-12-17 23:00:59+01:00	2022-12-17	23:00:59	100	1356882521414381570	mspikin3	MS Store		...	9.077751,8.6774567,105km	
2	1604232976551878657	1604181784551366656	2022-12-17 22:52:01+01:00	2022-12-17	22:52:01	100	1479877399202443269	odeeubana	Cirilla of UI		...	9.077751,8.6774567,105km	
3	1604230295586164736	1604212610890711040	2022-12-17 22:41:22+01:00	2022-12-17	22:41:22	100	155314755	pilotsmt	Michael		...	9.077751,8.6774567,105km	
4	1604224892789989377	1604212610890711040	2022-12-17 22:19:54+01:00	2022-12-17	22:19:54	100	1568918275752345603	adekunl46032443	Adekunle		...	9.077751,8.6774567,105km	

5 rows × 36 columns

Figure 3.4 Data filtering and dropped columns

The image above depicts Pandas reading data with several unwanted columns and irrelevant data. The data contains several relevant columns; it is therefore necessary to drop the unwanted columns, leaving the only needed columns.

3.4. Data Cleaning

After Data filtering, we are left with a data frame containing just one column named tweet. This column contains all the tweets of several individuals gathered using our Twint tool. Even though we have filtered out our data by simply dropping some irrelevant column, It appears the data is still very messy and cannot be used in its current condition. The data currently contains special characters merged with words and between sentences, contains links, underscores e.t.c The presence of these characters between and before the texts makes it impossible for our tool to read and analyze words. The tool we used in solving this problem is known as **regular expression (re)** module. The **re** module also known as regular expression is a python module used to write specific codes that target particular text patterns, and special characters. We therefore write a regex program that helps remove all unwanted characters in our tweets so our Ai tool can read, analyze and properly categorize the tweets.

```
# Cean The Text
def cleanText(text):
          text = re.sub(r'@[A-Za-z0-9]+','',text) #remove @ mentions
          text = re.sub(r'_[A-za-z0-9]+','',text)
          text = re.sub(r'@[\sA-za-z0-9]+','',text)
          text = re.sub(r'#','',text) # remving the # symbol
          text = re.sub(r'RT[\s]+','',text) # removing RT
          text = re.sub(r'https?:\/\/\S+','',text) #remve the hyperlink
          text = re.sub(r'mtn','',text)

          return text


data = data['tweet'].apply(cleanText)

data.head()
```

Figure 3.5 Data cleaned using regular expression

3.5. Data Classification

This stage includes evaluating the quality of the social media data after it has been filtered, and using text classification. Designating a collection of computerized methods for extracting, and quantifying information from textual documents, is a process known as text classification, categorization, Summarization, clustering (Han et al; 2014), etc. This phase would allow the use not only of online data recovery and text cleaning tools but also of computational tools for constructive data use. There is a set of text classification algorithms used like SVM, Decision Tree (Meier et al; 2019), etc. Many text mining tools have been developed to analyze the performance of social media platforms. These allow keeping track and interpret online texts from news, blogs, emails, and other outlets. Text mining tools can also help understanding how people are reacting to your brand and content on social media by measuring the number of messages, likes, and supporters.

3.6. Sentiment Analysis

Sentiment analysis is a considerable phase since it focuses on the processing and analysis of emerging functionalities related to customer sentiment. Consequently, it will be possible to process the customer's opinion that is collected via its publications on social networks to identify the customer's opinion on a product or service. This phase is about distinguishing the feelings and emotions of the customer and giving a clear vision of the voice of the customer. In the sentiment analysis phase, the detection system refers to the detection of unusual expressions. The emotional state of gaining customers through social media reveals the value of organizational awareness and industry for decision-making (Laghaei et al; 2017). The impact of ignoring or disregarding these patterns may be intense. To achieve this, we have to look into two important topics in sentiment analysis. which are sentiment polarity and sentiment subjectivity.

Polarity simply refers to the strength of a decision, whether positive or negative while **Subjectivity** measures the proportion of factual and personal information in the text. Defining functions that will calculate negative and positive polarity is necessary and further visualizing the tweet data using different types of plots and graphs helped to give better understanding of the data.

A Word cloud helps to visualize textual data and aids the discovery of trends and patterns. It enlarges the words that appear more frequently in the data.

3.7. Extract, Transform, Load

Since user opinions are so important in decision-making processes, many data warehouse solutions use the ETL (Extract, Transform, and Load) phase to incorporate opinions into a cleaning and integration process (Lu et al, 2018). The ETL system seeks to make data more relevant so that users' views shared on social networks can be analyzed. The ETL aims to retrieve data first, then convert it using a series of processes such as error deletion, correction, data manipulation, counting of posts, analysis of comments, labeling of sentiments, etc. Finally, we will go through the ETL's final process, which is responsible for loading the transformed data into a data warehouse. The ETL phase is critical since it allows the data to be structured so that it can be used by the tools in the following steps.

3.8. Analysis & Reporting

Finally, reports are created as part of this pre-process to help the end-users understand the results. These end-users would be able to get a better understanding of consumer behavior, allowing them to interpret the data and making it understandable (Sebastian,2002). Reporting

is about transforming data into information, while analysis is the process of transforming information into knowledge.

CHAPTER FOUR

IMPLEMENTATION OF THE SENTIMENT ANALYSIS

The steps we will take to execute sentiment analysis, utilizing some of the Natural Language Processing strategies outlined, will now be described. The use of a very popular library called TextBlob is introduced.

4.0 Data Importation

The first step is to import all required dependencies, followed by an import of our data from our local directory. As mentioned earlier in the previous chapter, the scraped data is stored as a JSON file. Over here, we are reading the JSON file through Pandas and making it a dataframe.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from textblob import TextBlob
%matplotlib inline
```

```
[ ] # Data Imports
    data = pd.read_json('test.json',lines=True)
    data.head()
```

Figure 4.1 Reading in the JSON data

4.1 Text Cleaning Using Regular Expressions

```
[ ]  # Cean The Text
     def cleanText(text):
               text = re.sub(r'@[A-Za-z0-9]+','',text) #remove @ mentions
               text = re.sub(r'_[A-za-z0-9]+','',text)
               text = re.sub(r'@[\sA-za-z0-9]+','',text)
               text = re.sub(r'#','',text) # remving the # symbol
               text = re.sub(r'RT[\s]+','',text) # removing RT
               text = re.sub(r'https?:\/\/\S+','',text) #remve the hyperlink
               text = re.sub(r'mtn','',text)

               return text


     data = data['tweet'].apply(cleanText)

     data.head()
```

Figure 4.2 cleaning the text

The result of the above code is found below

```
0    Join Aruwaab9ja VTU Official WhatsApp 👇 and bu...
1                                                  MTN
2               Zenith coupled with MTN network,tweh!!
3                                                 Mtn
4                                                  Mtn
Name: tweet, dtype: object
```

Figure 4.3 Data cleaning result

4.2 Getting Subjectivity and Polarity.

As earlier mentioned in the previous chapter, **Polarity** refers to an extraction of either positive or negative sentiments of a particular text data while **Subjectivity** measures the proportion of factual and personal information in the text. To get our subjectivity and polarity, we create functions and make use of the sentiment. Subjectivity as well as the sentiment. Polarity function in the TextBlob module. We then create two new columns to hold both polarity and subjectivity.

These two functions created earlier are to be applied to the newly created columns of the main data.

```
# Subjctivity function
def getSubjectivity(text):
          return TextBlob(text).sentiment.subjectivity

# Polarity function
def getPolarity(text):
          return TextBlob(text).sentiment.polarity


# create 2 new columns

data['Subjectivity'] = data['tweet'].apply(getSubjectivity)
data['Polarity'] = data['tweet'].apply(getPolarity)

#
#
#  show the new columns

data
```

Figure 4.4 Get subjectivity and polarity

	tweet	Subjectivity	Polarity
0	Join Aruwaab9ja VTU Official WhatsApp 👇 and bu...	0.125000	0.000000
1	MTN	0.000000	0.000000
2	Zenith coupled with MTN network,tweh!!	0.000000	0.000000
3	Mtn	0.000000	0.000000
4	Mtn	0.000000	0.000000
...	...	...	...
306		0.000000	0.000000
307	MTN. Wow, I've been anticipating a worship...	0.833333	0.333333
308		0.000000	0.000000
309	MTN	0.000000	0.000000
310	What is wrong with MTN??!!!😫	0.900000	-0.500000

311 rows × 3 columns

Figure 4.5 Result for Subjectivity and Polarity function

4.3 Visualization Using Word Cloud

A word cloud is a graphic representation of text data in which the larger the letters are, the more frequently they are used. The visualization of unstructured text data and the discovery of trends and patterns are both made possible by-word clouds.

```
# plot Word cloud

allWords = ' '. join(twts for twts in data['tweet'])
wordcloud = WordCloud(width=500,height=300,random_state = 21,max_font_size = 119).generate(allWords)
plt.imshow(wordcloud,interpolation='bilinear')
plt.axis('off')
plt.show()
```

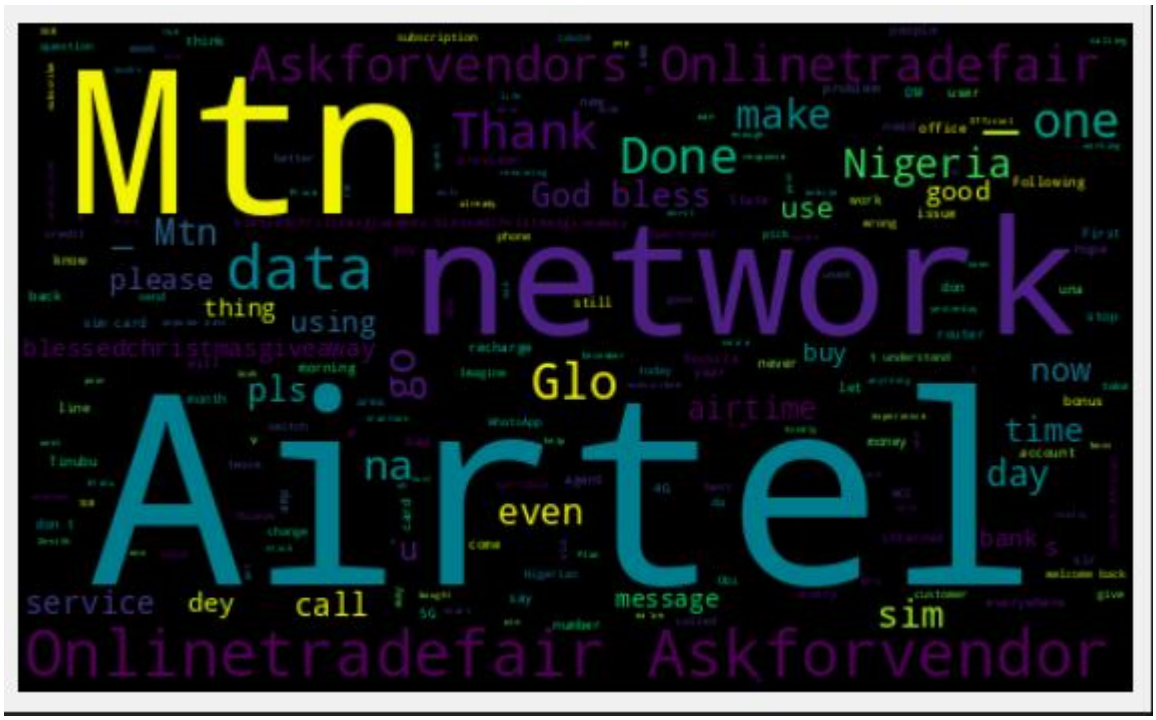


Figure 4.6 Plotting word cloud

From the above word cloud plot, I see three major words that appear larger than normal (**Mtn, Network, and Airtel**). Without further analysis, this suggests that the major problem with this data or the main focus of the data revolves around these three large words. By interpretation, it suggests the fact that customers or end users are complaining about the Airtel network, hence the comparison made.

Because these are all still assumptions and It's not a good practice to generalize on the output of a Word cloud Plot, there is therefore a need to further engage data analysis and other visualization methods.

4.4 Analysis

In this phase, we begin the analysis of the textual data and define functions that help give context to the data analysis.

```
# Function to Compute Negative, Neutral and Positive Analysis
def getAnalysis(score):
        if score < 0:
                return 'Negative'
        elif score == 0:
                return 'Neutral'
        else:
                return 'Positive'

data['Analysis'] = data['Polarity'].apply(getAnalysis)

# show the Data Frame
data
```

✓ 0.5s

Figure 4.7a get analysis

A get-analysis function is created in the above code. It calculates and returns the polarity of the data. To implement and view the functions' results, a new column on the main data is created and named "Analysis." It is in this column that the result of the polarity calculation is applied.

The result of the above code is below

	tweet	Subjectivity	Polarity	Analysis
0	Join Aruwaab9ja VTU Official WhatsApp 👇 and bu...	0.1250	0.0000	Neutral
1		0.0000	0.0000	Neutral
2	Zenith coupled with network,tweh!!	0.0000	0.0000	Neutral
3	Mtn	0.0000	0.0000	Neutral
4	Mtn	0.0000	0.0000	Neutral
...	...	...	...	...
470	Hello Airtel, I just subscribed for 1.5 GB da...	0.4750	0.3750	Positive
471	💯 AIRTEL	0.0000	0.0000	Neutral
472	Airtel	0.0000	0.0000	Neutral
473	Airtel	0.0000	0.0000	Neutral
474	Enough is enough already changed my sim card t...	0.4375	0.0625	Positive

475 rows × 4 columns

Figure 4.7b result of analysis

4.5 Visualize Positive Polarity Tweets

```
#Print all the psitive Tweets

j = 1
sortedData= data.sort_values(by=['Polarity'])
for i in range(0,sortedData.shape[0]):
        if(sortedData['Analysis'][i] == 'Positive'):
                print(str(j)+')'+sortedData['tweet'][i])
                print()
                j+=1
```

✓ 0.1s

Figure 4.8a

The above program helps to print out all positive polarity tweets. The result of the code can be found below:

```
6) Thank You o  I be think say na only me. Be like  don high

7)I use 3 major internet providers; Fiberone,  and Spectranet and I can categorically state that Spectranet and  are thieves when it comes to this data thing especially !!

8)How is traction app sponsoring my data?? Data that I bought with my own money oh.. !!!!

9)  Drop Mine too  09034994259 LOYAL FOLLOWER😂✌

10)   I hope to win this time...Airtel, Mtn, 9mobile.

11)  I don't support what he is saying, but this is not true. Before  came to Nigeria, Vodafone wanted coming but were prevented.  just gave OBJ a good deal

12) No brainer.   You don't want to know how much I spend on data every month.  Just make the network .

13) , earpiece, eating and watching series at the same time (pretty much)
...
92) Hello Airtel, I just subscribed for 1.5 GB daily subscription (6k) and I was debated but the subscription isn't successful 09019569582

93)Enough is enough already changed my sim card to Airtelng I have been using mtnng for the past 7 years but the network problem is getting out of hand, lets's experience better networking with Airtel✨
```

Figure 4.8b

4.6 Visualize Positive Polarity Tweets

```
j = 1
sortedData= data.sort_values(by=['Polarity'])
for i in range(0,sortedData.shape[0]):
        if(sortedData['Analysis'][i] == 'Negative'):
                print(str(j)+')'+sortedData['tweet'][i])
                print()
                j+=1
```

✓ 0.8s

Figure 4.9a

```
1)  Active

2)Making calls on  network has been really frustrating. So bad

3)So  can make you feel you're doing the wrong thing for using their network 😣😣😣

4)It is  or , all the other network providers can go to hell with their shitty network

5),  clearly has no explanation for this message that pops up, this is absolutely crazy

6)The new mtn app has terrible user experience. The whole login experience is terrible  please fix it.
```

Figure 4.9b

4.7 Visualize using Scatter Plots

A scatter plot is a particular kind of plot or mathematical diagram that displays values for typically two variables for a collection of data. It uses Cartesian coordinates. When the points are coded, a further variable can be seen. In this case we are plotting our Polarity(x-axis) against the Subjectivity (on the y-axis).

```
#lot polarity
plt.figure(figsize=(8,6))
for i in range(0,data.shape[0]):
          plt.scatter(data['Polarity'][i], data['Subjectivity'][i],color='Blue')

plt.title('Sentiment Analysis')
plt.xlabel('Polarity')
plt.ylabel('Subjectivity')
plt.show()
```

✓ 5.4s

Figure 4.9c

4.8 Calculate Percentage of Negative and Positive Polarity

Figure

4.9.1

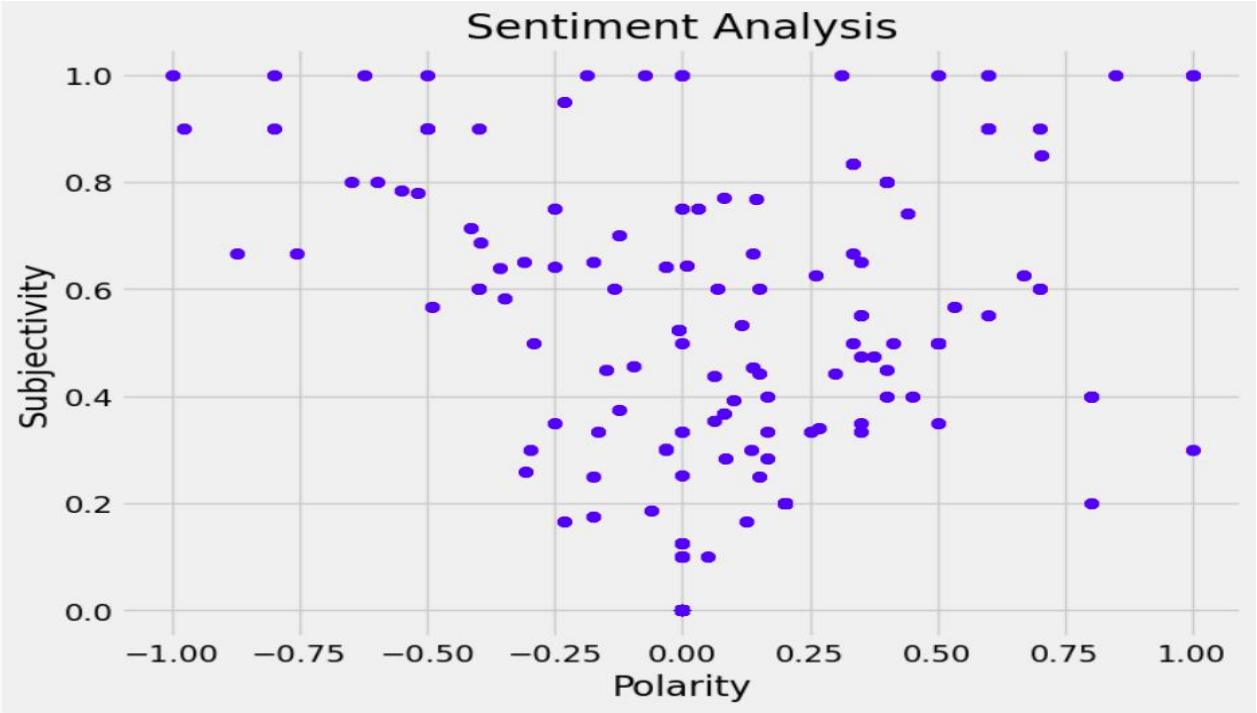


Figure 4.9.2

```
#Get percentage of Positive Tweets
ptweets = data[data.Analysis == 'Positive']
ptweets = ptweets['tweet']
round((ptweets.shape[0]/data.shape[0])*100,1)
```

✓ 0.1s

```
19.6
```

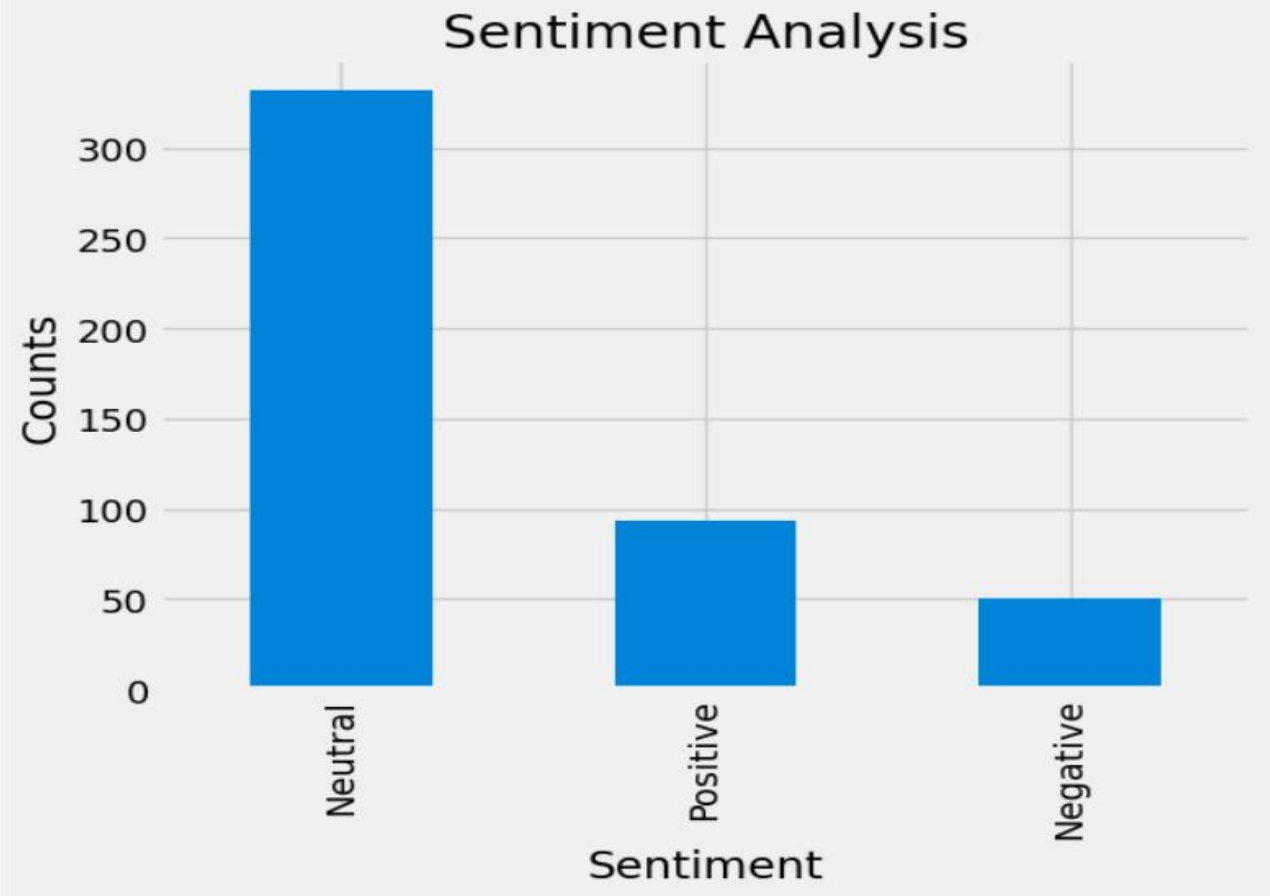
```
Ntweets = data[data.Analysis == 'Nagative']
Ntweets = Ntweets['tweet']
round((Ntweets.shape[0]/data.shape[0])*100,1)
```

✓ 0.4s

```
0.0
```

4.9 Visualization using Bar chart

```
#Show Value Counts

data['Analysis'].value_counts()

#Plot and visualize the counts

plt.title('Sentiment Analysis')

plt.xlabel('Sentiment')

plt.ylabel('Counts')

data['Analysis'].value_counts().plot(kind='bar')

plt.show()
```



The figure above displays the data with their respective lengths based on the proportionality of their values. It appears the positive polarity exceeds that of the negative.

4.9.1 Development tools

I. Visual Studio

Visual Studio is an integrated development environment from Microsoft. Microsoft Visual Studio is an IDE that is used to develop various kinds of software, including computer programs, websites, online applications, web services, and mobile applications. Compilers, completion tools, and other features are included to make the software development process easier.

II. Jupyter Notebook

A web-based interactive computing platform is the Jupyter Notebook. To create and share documents with live code, equations, visualizations, and text, you can use the free and open-source Jupyter Notebook web tool.

III. Python

Python is an interpreted, high-level, open-source programming language created by Van Guildo Rossum that is used for creating various applications. Python uses garbage collection and dynamic typing and enhances code readability.

IV. Pandas

Pandas is a free software software package created in python for manipulating and analyzing data. It includes specific data structures and procedures for working with time series and mathematical tables.

V. Text Blob

Python's TextBlob package is used to process textual data. It offers a straightforward API for getting started with typical natural language processing (NLP) activities like part-of-

speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and others.

VI. MatPlotLib

Python's Matplotlib toolkit provides a complete tool for building static, animated, and interactive visualizations. Matplotlib makes difficult things possible and simple things easy. With the help of general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK, it offers an object-oriented API for embedding plots into applications.

VII. Twint

Twint is the data collection tool utilized in the research. It is a sophisticated Python-written Twitter scraping tool that enables Twitter profiles to be scraped for Tweets without using Twitter's API.

CHAPTER FIVE

5.0 Conclusion

BI is not just an IT system, but a corporate asset which enables organizations to manage insights of their data, businesses, strategies, profits (past, present, and forecast) and performance. Thus, it helps to make better decisions.

Business IT Alignment plays an important role in an organization by enabling the business and IT management to work cohesively to support the organization's goals and objectives by implementing technology solutions such as BI-enabled business strategy.

SMEs requires a new approach in order to manage their insights of the businesses by implementing a BI solution, as various reasons discovered in the paper such as vendors targeted large enterprises, the capability of complex resources and finance. Hence some technology solutions were overviewed, and solutions such as cloud based or SQL Server BI solution is most promising to meet SMEs' requirements.

Overall, it is the challenge for organizations, SME or Large, to discover new ideas and processes, define strategy and implement the findings into BI capability as the competition grows, and market and customer demands increase. Most of all, BI is the ultimate asset and tool, which will enable organizations to manage insights of their business and make better decisions.

5.1 Recommendation

From the analysis of Airtel data, we discovered that Business Intelligence solution plays a very important role.

The use of modern like artificial intelligence helps in providing better business solutions.

After evaluation, the results obtained shows that the model would have done better if it understood the Nigerian way of communicating; slangs, pidgin. To enhance the model's accuracy, handling the peculiarity of such indigenous languages should be done by building Natural Language Processing tools for them.

APPENDIX

{"id": 1605142329756909570, "conversation_id": "1605119428517142530", "created_at": "2022-12-20 11:05:28 WAT", "date": "2022-12-20", "time": "11:05:28", "timezone": "+0100", "user_id":_