TIME SERIES ANALYSIS ON THE NUMBER OF POS TRANSACTIONS IN NIGERIA

BY

FAMUYIWA, OLUWASEYI AYOMIDE

17/56EG053

BEING PROJECT REPORT SUBMITTED TO THE DEPARTMENT OF STATISTICS IN PARTIAL FUFILMENT OF THE REQUIREMENTS FOR THE AWARD OF BACHELOR OF SCIENCE (HONOURS) DEGREE IN STATISTICS OF THE UNIVERSITY OF ILORIN, ILORIN, NIGERIA.

NOVEMBER, 2022

ATTESTATION

This is to certify that the project work entitled **time series analysis on the number of POS transactions in Nigeria** is an original work carried out by **Famuyiwa, Oluwaseyi Ayomide,** with the Matriculation number **17/56EG053** under my supervision.

. . . .

Famuyiwa, Oluwaseyi Ayomide
17/56EG053

. . . .

Dr R.B Afolayan
Supervisor

CERTIFICATION

This work has been read and approved as meeting the partial requirement for the award of Bachelor of Science Degree (B. s. c), in Statistics, University of Ilorin, Ilorin, Nigeria.

……………………………… …………………

DR R.B AFOLAYAN **Date**

SUPERVISOR

………………………………… ………………

PROF. G.M. OYEYEMI **Date**

HEAD OF DEPARTMENT

………………………… ………………

PROF. F.B ADEBOLA **Date**

EXTERNAL EXAMINER

DEDICATION

This project is dedicated to Almighty God for sparing my life throughout my stay in University of Ilorin and for making this project a success, and also to my lovely parents and siblings for the everlasting support financially, morally and spiritually. I pray that the Almighty God continue to be with you and grant you all your heart desires. (Amen).

ACKNOWLEDGEMENTS

I give all the glory and honour to Almighty God, the ruler of the universe for his unending mercies and special grace showered upon me throughout my days at University of Ilorin as an Undergraduate.

My profound gratitude goes to my supervisor Dr R.B. Afolayan for his immeasurable support, guidance, correction, tolerance and contribution throughout the period of this project, His wise advice and encouragement aided the writing of this project in countless ways.

I am grateful to the Head of Statistics Department, University of Ilorin, Prof G.M Oyeyemi, and all the lecturers in Statistics Department for the inerasable, distinctive and solid Foundation laid on me. My special regards goes to my zealous Level adviser Dr O.A Adeniyi for the guidance, counselling and scolding throughout the course of this study.

I appreciate my parents, Mr & Mrs Famuyiwa whom God used as a vessel to bring me out to the world and by whose support my schooling is made possible, I pray to Almighty God on your behalf for long life in good health to reap the fruits of your labour. I would also like to thank my siblings, Oluwaseun Morenike and Omolara Yetunde, and the family of Iweka and Omoleye for their love, support and advice.

I would also like to express my special thanks of gratitude to my friends and course mates who have supported me morally through the course of this study. I pray that Almighty God answers their heart desires.

ABSTRACT

The high level of difficulty of waiting in line to be served in bank has made the use of electronic means of transactions grow significantly. Point of Sales (POS) is an electronic means of payment that is widely used as a means of payment in different parts of the world. The major challenge to the deployment of POS terminal in Nigeria has remained poor network and high cost of operation. This research focuses on studying the adaptation of Nigerians to electronic means of payment through the use of Point of Sales as a means of transaction for the past few years with a view of predicting the future use of electronic means of payment in Nigeria. A time plot of the data was constructed to study the components of time series present in the data. The plot revealed that there is an underlying upward trend in the use of Point of Sales in Nigeria. Stationary time series data is considered the best type of data to use for prediction because of its ability to maintain constant mean and variance with a unit change in time. The augmented dickey fuller test for the stationary of the data was used and the results revealed that the time series data was non- stationary at level but stationary at first difference. Box Jenkins time series procedure was administered to determine a suitable model for prediction, the box-Jenkins time series model is simply known as the Autoregressive Integrated Moving average (ARIMA) model. The patterns of the Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) were studied to have an idea of the appropriate Box-Jenkins time series model to be used. The study revealed that Autoregressive (AR) model was appropriate for prediction. In order to evaluate the efficiency of our models, the data was divided into train and test data. The first five years of our data was used for the models and the remaining 12 data point (last year) was used to test or evaluate the models. The Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Mean Absolute Percentage Error (MAPE) were used to select the best models among the following competing models: ARIMA(1,1,0), ARIMA(2,1,0), ARIMA(3,1,0), ARIMA(4,1,0), ARIMA(5,1,0), and ARIMA(6,1,0).Using the model with minimum value of AIC, BIC and MAPE, ARIMA(6,1,0) was chosen to be the best model for forecasting POS data. Several tests to validate the results of our model were carried out, namely, Box ljung, and Normality test. ARIMA (6,1,0) was used to predict the number of Point of Sales transactions for the next two years. The results suggested that there would be a steady increase in the use of P0S in Nigeria. Consequently, the study recommends that the Federal government should make sure that there is an adequate supply of POS devices for Nigerians. This would ensure that there are enough POS machines to be used for the increasing number of POS transactions.

TABLE OF CONTENTS

CHAPTER TWO: LITERATURE REVIEW AND METHODOLOGY

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 Background of Study

The history of Point of Sale (POS) systems began in 1879, with the invention of the cash register. The first cash register was created by James Ritty, a bar owner from Ohio, and his brother in 1879. They called it Ritty's Incorruptible Cashier. Ritty eventually sold John H. Patterson his patent who would establish the National Cash Register (NCR) Company in 1884. NCR Corporation made additional changes to the cash register. Ultimately, these helped transform it into a digital machine that would be used widely for decades.

The location where a consumer completes the payment for goods or services is referred to as the "point of sale." (Hayes, 2020).It is an electronic payment system that was introduced to Nigerians by the Central bank of Nigeria (CBN) in 2012 to further drive home its cashless policy aimed at enhancing Nigeria's payment system. Since the initiative's launch, it has achieved exceptional success, with the total amount of transactions climbing to N99.6 million and the deployment of POS terminals increasing from 5000 to 153,167 in April 2014. POS system can significantly simplify daily business tasks, notably making it easier to manage retail business. A point of sales helps make payments faster, waiting time is greatly reduced for consumers with a point of Sale. Nigerians increasingly prefer cashless payments over holding and transacting in cash, and the country's e-payment industry is continuing to show strong growth commitment. Payment systems are tools that make monetary and other financial transactions easier to clear and settle. They are a key service offered by banks and other financial organizations and are used in place of tendering cash in transactions. They consist of the institutions, tools, regulations, practices, and standards that are put in place to facilitate the exchange of money between parties fulfilling one another's responsibilities. Payment systems can typically be divided into two categories: traditional (physical) and electronic payment systems (EPS). Documentary credits like letter of credits and negotiable instruments like checks are the two types of traditional payment methods. Conversely, with the use of EPS, payments for electronic transactions can be made without the use of cash or checks. It can be used in a variety of ways, including with credit and debit cards, virtual cards, electronic funds transfers, e-wallets, mobile payments, internet banking, and financial transactions (Oturu, 2020).

The basic goal of time series modelling is to meticulously gather and thoroughly analyse historical data from a time series in order to create a model that accurately captures the series' underlying structure. Data points collected over time may contain internal structures (such as autocorrelation, trends, or seasonal fluctuation) that need to be taken into account. Time series take this into account. The behaviour of time series variables such as means of payment statistics is not consistent and to forecast it is irrational. Despite these assertions, many multinational cooperation and speculators continue to make hedging decisions based on forecasted statistics using ex-post data(data available after that time) as their basis. These hedging decisions are made under the premise that patterns exists in the ex-post data and these patterns provide an indication of future movement of means of payment statistics. If such patterns exist, then it is possible in principle to apply modern mathematical tools and techniques to forecast the means of payment statistics (Hamilton, 1994).

Insecurity in Nigeria is a major cause for choosing to go with the e-payment rather than transacting in cash. In Nigeria, using an ATM or virtual card to make a payment is both simpler and safer, especially in light of the rise in robberies, bag snatching, and kidnappings, among other crimes. Since the heat of the COVID-19 pandemic, POS firms have grown dramatically because to the hardship of waiting in line to be served in the bank only to make withdrawals. In June 2021, there were 976,898 registered POS terminals in the nation, up from 523,488 at the same time in 2020, according to Nigeria Inter Bank Settlement System (NIBSS). Nigerians are increasingly adopting cashless purchases, as seen by the rise in e-payment transactions in the country. For Nigerians, especially young people, the rise in the number of POS enterprises in the nation has created a significant source of employment (Oyekanmi, 2021).

This study intends to use the data available on payment system statistics in Nigeria from 2014-2019 to understand the use of Number of POS systems transactions in Nigeria and to forecast future trend in the usage of POS system transactions in Nigeria using an appropriate time series model.

1.2 Statement of Problem

Poor network and high cost of operation have remained major challenges to the deployment of POS terminals in Nigeria. Time series analysis involves a wide range of data variances, therefore analysts occasionally need to create intricate models. Analysts, however, are unable to take into account all variations and cannot apply a particular model to all samples.

1.3 Aim and Objectives of Study

The aim of the study is to Provide an appropriate time series model to forecast (predict) the Number of POS transactions in Nigeria and the specific objectives of this study are to:

i) Build a time plot to investigate the trends in Nigeria's POS transaction volume;

ii) Fit an appropriate model for the number of POS transactions over the years; and

iii) Predict the number of POS transactions for the next two years.

1.4 Scope of Study

This study utilizes the time series technique to analyse the Number of POS systems statistics in Nigeria and fit the specific change process of Number of POS transactions statistics to predict the trend of change. This study covers a period of 6 years (2014-2019). It takes the monthly Number of POS transactions in Nigeria into consideration.

1.5 Significance of Study

This study investigates the impact of POS transactions on the growth of Nigeria electronic payment systems as well as the adaptation of Nigerians to electronic payment systems rather than traditional payment systems through the number of POS transactions in the country. Electronic means of payment systems has over the years proven to have certain advantages over traditional means of payment, Since the advent of POS terminals in Nigeria, study has made it known that POS can serve as a means of income to entrepreneurs and has made people more self-employed, it is also important to note that Electronic means of payment like POS transactions also faces certain challenges that hinders its functionality and effectiveness occasionally, therefore through the course of this study, time series analysis is carried out to examine whether POS transactions is a type of payment system that would continue to be incorporated by Nigerians in the future or is a type of payment system that would eventually lose its relevance in the financial sector.

1.6 Organisation of the Project

This project investigates the use of Times Series model to examine the trend of payment systems in Nigeria. The project is made up of four chapters. The first chapter contains the introduction of this study, statement of problems, aim and objectives of study, scope of study, significance of the study and the organisation of the study. The second chapter will entail the review of literature and methodology (the method of analysis) which include identification of the appropriate time series model structure as well as the order of the AR, MA, and ARMA terms, specification of the criterion used for the model selection, model testing /validation and prediction. The third chapter will contain the analysis of the data that has been collected and the discussion of results obtained. The fourth chapter will be made up of summary and conclusion.

CHAPTER TWO

LITERATURE REVIEW AND METHODOLOGY

2.1 Introduction

This chapter talks about the literature review of time series Analysis, components of time series analysis, time plot as well as stationarity and non-stationarity of time series. This chapter would also discuss extensively about the methodology used to fit a time series model that can be used to predict future values of the number of POS transactions in Nigeria.

2.2 Literature Review

2.2.1 Time series

A time series is a sequence of data points that occur in successive order over some period of time (Hayes, 2021).Time series is a set of data point collected at equal or non-equal time intervals. Examples of time series data are weather data, Rainfall measurements, Temperature readings, Heart rate monitoring, quarterly sales, Stock prices, automated stock trading, industry forecasts, Daily sales etc. In time series analysis, analysts record data points at consistent intervals over a set of period of time rather than just recording the data points randomly. The difference between time series data and other data is that time series analysis can show how variables change overtime. Time series analysis typically requires a large number of data points to ensure consistency and reliability. Time series analysis helps organizations understand the underlying causes of trends or systematic patterns over time. When organizations analyse data over consistent intervals, they can also use time series forecasting to predict the likelihood of future events. Time series analysis is used for non-stationary data (things that are constantly fluctuating over time or are affected by time. Industries like Finance, retail, and economics frequently use time series analysis because currency and sales are always changing. The goal of time series analysis is to find patterns in the data and use the data for predictions. For example, if your data is affected by past data, one way to model that behaviour is through the AR process (Glen, 2013).

2.2.2 Classification of time series data

Stock time series data simply refers to measuring characteristics at a specific moment, similar to a static picture of the data as it was. Flow time series data simply means measuring the activity of the attributes over a certain period.

2.2.3 Components of time series

The forces which affect the values of an observation in a time series are the components of a time series. The components of time series can be explained in four different ways.

i. Secular trend

This is also known as a long term movement. The trend shows the general tendency of the data to increase or decrease during a long period of time. A trend is an average, smooth, long-term tendency. The growth or decline does not necessarily have to be going in the same direction over the course of the specified length of time. The overall trend must be stable, upward, or downward. (Chatfield, 2013)

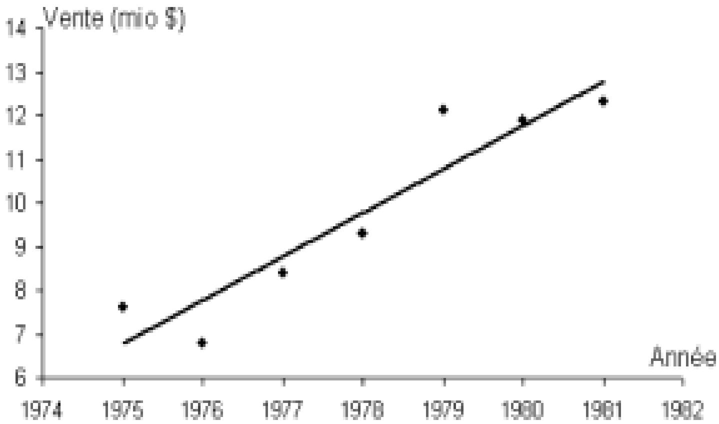

Fig 2.1: Secular trend

If we plot the time series values on a graph in accordance with time t, the pattern of the data clustering shows the type of trend. If the data are more or less clustered around a straight line, the trend is linear; otherwise, the trend is described as non-linear.

ii. Seasonal variations

They are short term movements. These are the rhythmic forces that function consistently and periodically during a period of less than a year.. They have the same or almost the same pattern during a period of 12 months, This variation may exist in data recorded hourly, weekly, quarterly, or monthly, but not in data recorded yearly. Natural forces or customs created by humans both contribute to these variances. Seasonal fluctuations are significantly influenced by the various seasons or climatic conditions. Production of crops depends on seasons, the sale of umbrella and raincoats in the rainy season, and the sale of electric fans and the A.C shoots up in summer seasons. The effects of man-made conventions such as some festivals, customs, habits, fashions and some occasions like marriages is easily noticeable. They recur themselves year after year. A seasonal rise should not be taken as a sign of improved business circumstances.

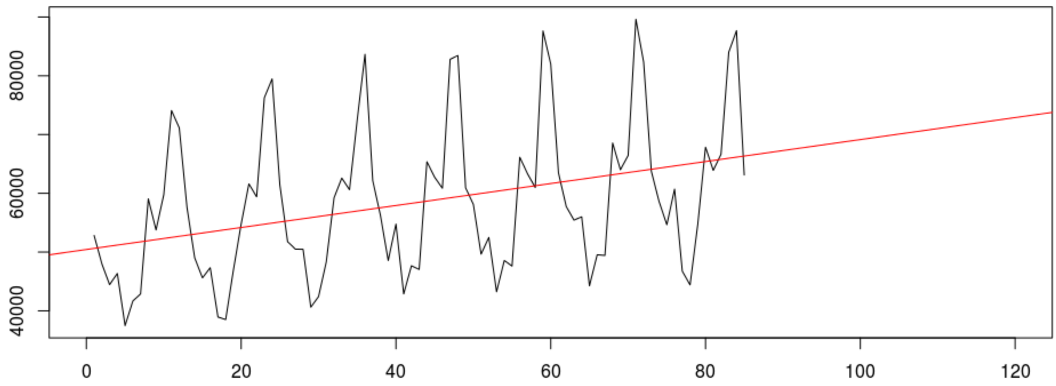

Fig 2.2: Seasonal variation

iii. Cyclic variations

Additionally, they have short-term motions. They are the fluctuations in a time series that persist for a period longer than a year. One complete period is a cycle. This cycle movement is sometimes called the ‘business cycle’. It is a four-phase cycle comprising of the phases of prosperity, recession, depression and recovery.

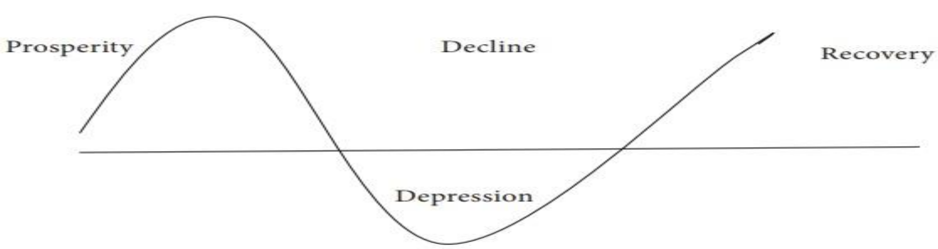

Fig 2.3: Cyclic variation

iv. Random or irregular movements

Another element contributes to the fluctuation in the variable under investigation. These variations are chaotic, unanticipated, unmanageable, and unforeseeable. These forces are earthquakes, wars, flood, famines, and any other disasters.

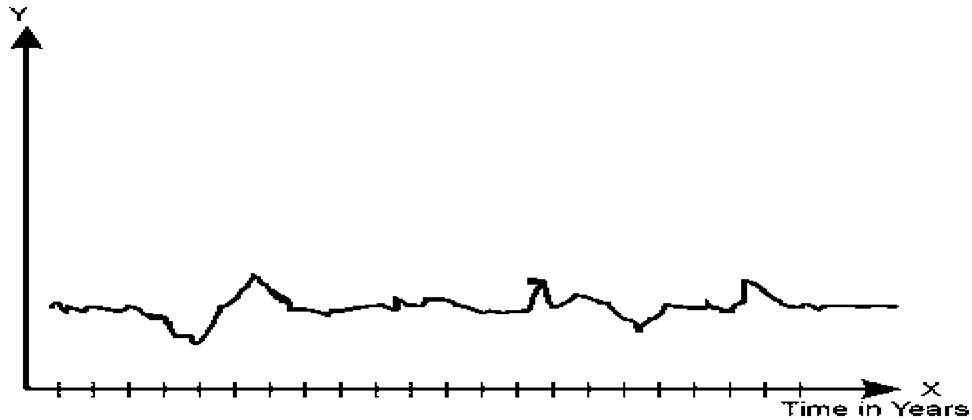

Fig 2.4: Irregular variation

2.2.4 Mathematical model for time series analysis

Mathematically, a time series function is given as

$$y_t = f(t) \quad (2.1)$$

Y_t, here is the value of the variable under study at time t. If the population is the variable under study at the various time period $t_1, t_2, t_3 \ldots t_n$, then the time series is t: $t_1, t_2, \ldots.. t_n; y_t$; $y_1, y_2, \ldots y_n$

2.2.5 Additive model for time series analysis

If y_t is the time series value at time t. T_t, S_t, C_t , and R_t are the trend value, seasonal value, cyclic and random fluctuations at time t respectively. According to the Additive Model, a time series can be expressed as

$$y_t = T_t + S_t + C_t + R_t \tag{2.2}$$

This model assumes that all four components of the time series act independently of each other.

2.2.6 Multiplicative model for time series analysis

The multiplicative model assumes that the various components in a time series operate proportionately to each other according to this model

$$y_t =. T_t \times S_t \times C_t \times R_t \tag{2.3}$$

2.2.7 Mixed models

This is combination of both additive and multiplicative model. The time series can be done using the model

$$y_t = T_t + S_t \times C_t \times R_t \text{ or } y_t = T_t + S_t + C_t R_t \qquad (2.4)$$

2.2.8 Time plot

A time plot, sometimes called a time series graph, charts values across time. They are similar to Cartesian plane x-y graphs, but while an x-y graph can plot a variety of “x” variables(for example, height, weight, age),time plots can only display time on the x axis. Unlike pie charts and bar charts, these plots do not have categories. Time plots are good for showing how data changes over time. Time graph is a revealer for a method you can use to analyse your data (Glen, 2013).

An example of time plot is shown below.

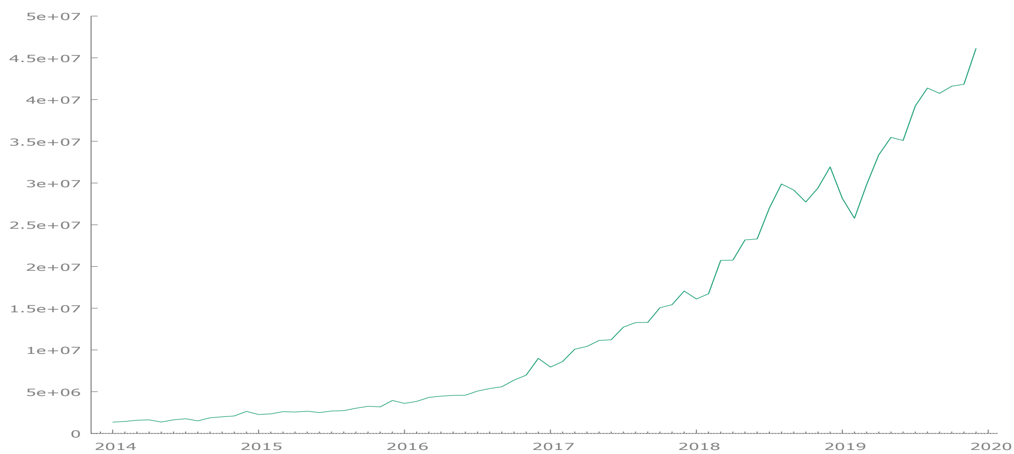

Fig 2.5: Time plot of monthly number of POS transactions from 2014-2019

2.2.9 Stationarity of time series

A stationary time series are series whose statistical properties like mean, variance, co variance do not vary with time or this statistical properties are not function of time. Time series that are stationary also lack trends or seasonal elements (Vijay, 2021). Real-world data sets are rarely stationary. If you ‘ve got a real-life data set, in most cases you won’t be able to run any processes on the data set directly, and you won’t be able to make useful predictions from it. One solution is to make the data stationary by transforming it. Stable time series are simpler for statistical models to accurately and reliably forecast. A stationary time series data set will not experience a change in distribution shape when there is a shift in time; Basic properties of the distribution like the mean, variance and co variance remain constant. Thus makes the model better at predictions (Glen, 2013).

2.2.10 Types of stationary series

i. Strict stationary: Satisfies the mathematical definition of a stationary process and mean, variance and co variance are not function of time.
ii. Seasonal Stationary: Series exhibiting seasonality.
iii. Trend stationary: Series exhibiting trend.

The most basic methods for stationarity detection rely on plotting the data, and visually checking for trend and seasonal components. Trying to determine whether a series was generated by a stationary process just by looking at its plot is a dubious task. However, there are some basic statistical tests that can be used to check for stationarity in a time series. The results of statistical tests are strongly inferred. They can only be used to provide information about how strongly a null hypothesis can be supported or disproven. For a particular problem, the result must be understood in order to have any value. However, they provide a quick check and confirmatory evidence that the time series is stationary or non-stationary (Vijay, 2021).

2.2.11 Test for stationarity in time series

There are various statistical tests and plots that can be used to detect stationarity in time series. In this study we would look into the use of ADF (Augmented Dickey fuller test) to determine stationarity in the time series data.

2.2.12 Augmented dickey fuller test for stationarity

A statistical test known as a unit root test is what is used for the augmented Dickey fuller test.

2.2.13 Unit root

A unit root is a characteristic of some stochastic processes (such a random walk) in probability theory and statistics that can lead to issues with statistical inference when using time series models (Vijay, 2021). While they share many characteristics with trend-stationary processes, unit root processes and those processes differ in a number of ways. A time series could be trend-stationary while also being non-stationary and lacking a unit root. In both unit root and trend-stationary processes, the mean can be growing or decreasing over time, however, in the presence of a shock, trend – stationary processes are mean-reverting (i.e. the time series will converge again towards the growing mean, which was not affected by the shock, an example of a trend stationary time series data is a seasonal time series data), while unit-root processes have a permanent impact on the mean. i.e. no convergence over time.

Consider a discrete-time stochastic process (y_t, t= 1, 2, 3…), and suppose that it can be written as an auto regressive process of order p:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + a_3 y_{t-3} + \ldots + a_p y_{t-p} + \varepsilon_t \qquad (2.5)$$

Here, (ε_t , $t = 0,1,2,3, \ldots$) is a serially uncorrelated, zero-mean stochastic process with constant variance σ^2. If m=1 is a root of the characteristic equation,

$$m^p - m^{p-1} a_1 - m^{p-2} a_2 - \cdots - a_p = 0 \qquad (2.6)$$

A unit root is then present in the stochastic process. The expanded Dickey broader test equation widens the model to accommodate high order regressive processes.

$$\Delta y_t = \alpha + \beta t + \gamma_{y_{t-1}} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} \ldots + \delta_p \Delta y_{t-p} + \varepsilon_t \qquad (2.7)$$

From the above equation, more differencing terms has been added and the rest of the equation remains the same as that of dickey fuller test. This adds more thoroughness to the test. The assumptions guiding the conduct of ADF test is how ever the same as that of dickey fuller test.

Augmented Dickey Fuller (ADF) test is conducted with the following assumptions;

Null hypothesis (H_0): Series is non- stationary or series has a unit root.

Alternative Hypothesis (H_1): Series is stationary or series has no unit root.

If the null hypothesis is failed to be rejected, then the test provides evidence that the series is non-stationary.

Conditions to reject Null hypothesis (H_0)

If Test statistics $<$ Critical Value and p-value < 0.05, then reject Null hypothesis (H_0), i.e., time series does not have a unit root, meaning it is stationary. It does not have a time dependent structure. (Vijay, 2021).

For more complex and extensive sets of time series models, the enhanced Dickey-Fuller test is used. The result of the augmented Dickey-Fuller test is zero. The more negative it is, the stronger the rejections of the hypothesis that there is a unit root at some level of confidence.

2.2.14 Testing procedure for augmented dickey fuller test

The ADF test is conducted using the same testing methodology as the dickey fuller test, but the model below is used instead.

$$\Delta y_t = \alpha + \beta t + \gamma_{y_{t-1}} + \delta_1 \Delta y_{t-1} + \cdots + \delta_p \Delta y_{t-p} + \varepsilon_t \qquad (2.8)$$

Where α is a constant, β is the coefficient on a time trend and p is the lag order of the auto regressive process. By including lags of order p the ADF allows for higher order auto regressive processes. This means the lag length p has to be determined when applying the test.

The unit root test is carried out under the null hypothesis $\boldsymbol{\gamma} = 0$ versus the alternative hypothesis of $\boldsymbol{\gamma} < 0$. once a value for the test statistic

$DF_T = \frac{\hat{\gamma}}{SE(\hat{\gamma})}$ Is computed, it can be compared to the relevant critical value for the dickey fuller test. We are only concerned with the negative values of our test statistics DF_T. If the calculated test statistics is less (more negative) than the critical value, then the null hypothesis of $\Upsilon = 0$ is rejected and no unit root is present, otherwise the null hypothesis is not rejected.

The logic behind the ADF test is that if the series is characterized by a unit root process then the lagged level of the series (y_{t-1}) will provide no relevant information in predicting the change in

y_t Besides the one obtained in the lagged changes (Δy_{t-k}).In this case the $\Upsilon = 0$ is not rejected. In contrast, when the process has no unit root, it is stationary and hence exhibits reversion to the mean. So the lagged level will provide relevant information in predicting the change in the series and the null hypothesis of a unit root will be rejected.

2.2.15 Stationarity and differencing

Most statistical modelling methods work on the assumption that the data is stationary. Time series that exhibit trends or seasonality are not stationary; they alter the time series' value at various points over time. Calculating the differences between successive observations is one method of making a non-stationary time series stationary. Differencing can help stabilise the mean of a time series by removing changes in the level of a time series, and therefore eliminating (or reducing) trend and seasonality (Kwiatkowski, 1992).

2.2.16 Order of differencing

The order of difference of a non-stationary time series data is the number of times the data has to be transformed such that the trend or seasonal components of the data has been completely removed. In most time series data set, we hardly go beyond the second order difference before the data becomes stationary. After differencing must have been carried out on the original time series, a new time series is generated called the differenced time series.

1. First order difference

This is simply the change between consecutive observations in the original series; it is also called lag-1 difference and can be written as

$$y'_t = y_t - y_{t-1}. \qquad (2.9)$$

Where;

$y'_t =$ Differenced observation

$y_t =$ Observation at the current time e.g. number of POS transactions in the month of February.

$y_{t-1} =$Observation at the previous time e.g. number of POS transactions in the month of January.

The differenced series will have only T-1 values (T= total number of observations), since it is not possible to calculate a difference y'_1 for the first observation.

2. Second order difference

Occasionally some temporal structure may still exist after performing a differencing operation therefore, the data will not appear to be stationary and it may be necessary to difference the data a second time to obtain a stationary series: it can be written as

$$y''_t = y'_t - y'_{t-1}$$

$$y''_t = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2})$$

$$y''_t = y_t - 2y_{t-1} + y_{t-2} \quad (2.10)$$

In this case, y''_t will have T-2 values. In practice, it is almost never necessary to go beyond second – order differences, if the temporal structures are not completely removed after second order difference, then other forms of transformation (e.g. standardization) may be necessary.

2.3 Methodology

2.3.1 Box- jenkins time series models

The ARIMA (Auto regressive integrated moving average) model was firstly proposed by box and Jenkins in the early 1970s(Jenkins, 1976) and they are very popularly used in hydrology, especially for modelling monthly flows, seasonal flows as well as generation of some models(data generation) used for real time forecasting models. ARIMA is made up of two processes; the Autoregressive AR and the moving average MA.AR model believes that data values at current time spots are built over data values at previous time spots. If we use AR model to forecast, our job is to look at how many days do we have to look back to in order to forecast current value. Also MA model believes that the current data value is results of previous expected events. For example, Current stock price can be as a result of an election result, Covid 19 pandemic, and better sales revenue etc.

These Univarate models are used to better understand a single time-dependent variable, such as temperature over time, and to predict future data points of variables. These models also work on the assumption that the data is stationary. Analysts have to account for and remove as many differences and seasonality in past data points as they can. Luckily, the ARIMA model includes terms to account for moving averages, seasonal difference operators, and auto regressive terms within the model. This method is adopted in this study because of the nature of the data we wish to analyse. The number of POS transactions in Nigeria is a single time-dependent variable and hence can be analysed using box-Jenkins model.

According to (Petris, 2009), the ARMA model is a combination of AR (p) and MA (q) model, where p represents the order of terms of the AR model and q represents the order of terms of the MA model, this model believes that combining AR and

MA model is best for predicting a set of data. If no differencing is involved in the ARIMA model, then it becomes simply an ARMA. Majorly, ARIMA models intend to describe the current behaviour of variables in terms of linear relationships with their past values.

2.3.2 Auto regressive models

An AR model is simply a linear regression of the current value of the series against one or more prior values of the series (Glen, 2015). When there is some correlation between the values in a time series and the values that come before and after them, it is useful for predicting. In an AR model, you can only use the past data to model the behaviour, hence the name autoregressive (the Greek prefix-means "self"). The value of p is called the order of the AR model.

The Autoregressive (p) model is defined by the equation:

$$Y_t = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \ldots + \phi_p Y_{t-p} + \varepsilon_t \qquad (2.11)$$

Where:

i. $Y_{t-1}, Y_{t-2} \ldots Y_{t-p}$ are the past series values(lags),

ii. ε_t is white noise and is independently and identically distributed with a mean of zero and a constant variance.

iii. C is defined by the following equation:

$$C = \left(1 - \sum_{i=1}^{p} \phi_i\right) \mu, \quad (2.12)$$

Where μ is the process mean.

The AR process is an illustration of a stochastic process, which includes elements of randomness or uncertainty. Given the randomness, it's possible to fairly accurately anticipate future patterns using historical data. The process typically gets close enough for it to be helpful in most circumstances. Conditional models, Markov models, and transition models are other names for AR models.

2.3.3 Operator b (back shift operator)

Operator B is a very handy tool in expressing AR models in elegant forms, the effect of Operator B is to shift the argument to one step behind. For example, in an AR(1) model below

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t \quad (2.13)$$

Y_{t-1} Can be replaced with $\mathrm{B}Y_t$, the use of B on Y_t shifts Y_t to Y_{t-1}, we can then express AR(1) model as;

$$Y_t = \phi_1 \mathrm{B} Y_t + \varepsilon_t$$

$$Y_t(1 - \phi_1 \mathrm{B}) = \varepsilon_t$$

Consequently, $\mathrm{B}Y_{t-1}$ shifts Y_{t-1} to Y_{t-2} in an AR (2) model below;

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

$$Y_t = \phi_1 \mathrm{B} Y_t + \phi_2 \mathrm{B} Y_{t-1} + \varepsilon_t$$

$$Y_t = \phi_1 \mathrm{B} Y_t + \phi_2 \mathrm{B}(\mathrm{B} Y_t) + \varepsilon_t$$

$$Y_t = \phi_1 \mathrm{B} Y_t + \phi_2 B^2 Y_t + \varepsilon_t$$

$$Y_t - \phi_1 \mathrm{B} Y_t + \phi_2 B^2 Y_t = \varepsilon_t$$

$$Y_t(1 - \phi_1 \mathrm{B} - \phi_2 B^2) = \varepsilon_t$$

In general, AR(P) model is written as;

$$Y_t(1 - \sum_{i=1}^{p} \phi_i B^i) = \varepsilon_t \tag{2.14}$$

2.3.4 Moving average models

Moving average is a technique to get the overall idea of the trends in a data set. Conceptually, a moving average model is a linear regression of the series' most recent value against random shocks or white noise from one or more previous values. The MA model is a specific case and essential element of the more general ARMA and ARIMA models of time series, together with the AR model.

The MA (q) model is given by the equation:

$$Y_t = \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \theta_3\varepsilon_{t-3} + \cdots + \theta_q\varepsilon_{t-q} \quad (2.15)$$

Where:

i. $\theta_1, \theta_2, \ldots, \theta_q$ are the parameters of the model.

ii. $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots, \varepsilon_{t-q}$ are white noise error terms.

The moving average model should not be confused with the moving average, together with the AR model. The moving average model is a key component of the more general ARMA and ARIMA models of time series, while the moving average is commonly used with time series data to smooth out short-term fluctuations in time series data. If we use the moving average(model) to Analyse time series data, then our job is to identify how many days of our expected events are we looking back to forecast the current value. However the major difference between these two processes is that AR models depends on the lagged values of the data you are modelling to make forecasts while MA models depends on the errors(residuals) of the previous forecasts you made to make a current forecasts, hence error(noise) quickly vanishes with time in MA model.

2.3.5 Autoregressive moving average model

An ARMA (p, q) model, is a combination of AR (p) & MA (q) model, it is used to describe weakly stationary stochastic time series in terms of two polynomials. The first of these polynomials is for auto regression, the second is for the moving average.

The ARMA (p, q) equation is given by

$$Y_t = C + \varepsilon_t + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{i=1}^{q} \Theta_i \varepsilon_{t-i}. \qquad (2.16)$$

2.3.6 Steps in box jenkins model selection

In practise, when given a time series data set, the first step is to draw a time series plot chart, If it follows stationary time series data, the time series should be flat, the data value should change around a central value. If the time series in the time series plot chart is not flat, then the time series data follows non-stationary time series data which would need to be transformed to stationary before it can be used.

2.3.7 Model identification

To determine a proper model for a given time series data, it is necessary to carry out the Autocorrelation function (ACF) and Partial autocorrelation function (PACF) analysis. These statistical measurements show the relationships between the data in a time series. Plotting the ACF and PACF against subsequent time lags is frequently helpful for modelling and forecasting purposes. These plots help in determining the order of MA and AR terms (Jenkins, 1970).

2.3.8 Auto correlation function

Auto correlation is a mathematical representation of the degree of similarity between given time series and a lagged version of itself over the successive time intervals. Autocorrelation measures the relationship between a variable's current value and its past value. The first technique we use to determine values for p, d, and q is to compute the sample ACF from the data and compare this to known properties of the ACF for ARIMA models (Orlaith, 2011).

The sample autocorrelation is given as

$$r_k = \frac{\sum_{t=i}^{n-k}(Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}. \tag{2.17}$$

The sample autocorrelation r_k provides us with an estimate for the true autocorrelation ϱ_k. The ACF is very useful in identifying the MA process. This is because the first q terms in the ACF of an MA (q) processes are non-zero and the remaining terms are all zero. In non-stationary time series, the terms of the ACF do not decay to zero very quickly as they do in a stationary time series, but the ACF of a stationary series decays exponentially to zero.

2.3.9 The partial auto correlation function

It is the correlation between two variables, presuming that we are aware of and take into account the values of some other set of variables. This correlation function enables us to determine the sequence in which an AR process occurs. The PACF computes the correlation between two variables Y_t and Y_{t-k} after removing the effect of all intervening variables $Y_{t-1,}Y_{t-2,\ldots,}Y_{t-k+1}$. we can think of the PACF as a conditional correlation.

$$\phi_{kk} = corr\left(Y_t, Y_{t-k} \mid Y_{t-1,}Y_{t-2,\ldots,}Y_{t-k+1} \right). \tag{2.18}$$

There are more general patterns in the combination of the ACF and PACF plots that can be used to identify the order of ARIMA models:

Table 2.1: general patterns in the combination of the ACF and PACF plots

	ACF	PACF
AR model	Geometric decline	Sharp loss of significance at p lags
MA model	Sharp loss of significance at q lags.	Geometric decline.
ARMA model	Geometric decline	Geometric decline

Table 2.1 gives us an idea of the order of the time series model to be used, then a series of candidate models are fitted with a view of selecting the candidate model that best fits the data at hand. The best model is selected with the help of statistical model selection tools like the akaike information criterion (AIC) and the Bayesian information criterion (BIC).

2.3.10 The Akaike Information Criterion (AIC)

AIC is a tool for estimating prediction error. AIC calculates the quality of each model in relation to the other models given a set of models for the data. AIC deals with both the risk of over fitting and the risk of under fitting.

Suppose we have a statistical model of some data. Let k be the number of estimated parameters in the model. Let $\hat{L}$ be the maximum likelihood function for the model. Then the AIC value of the model is given as;

$$\text{AIC} = 2K - 2\ln(\hat{L}) \tag{2.19}$$

Where

K= number of estimated parameters in the model

$\hat{L}$= maximum value of the likelihood function of the model.

The model with the lowest AIC value among a group of potential ones for the data is the one that should be used. It is crucial to remember that AIC only assesses a model's quality in relation to other models, not its inherent quality. Thus if a model fits poorly, AIC will not give any warning of that. Hence, after selecting a model through AIC, it is usually a good practice to validate the absolute quality of the model. Such validation includes check of the model’s residuals to determine whether the residuals seem like random and tests of the model’s prediction.

2.3.11 Bayesian Information Criterion (BIC)

In statistics, the BIC is a criterion for model selection among a finite set of models; models with the lower BIC are generally preferred. It is based partly on the likelihood function and it is closely related to the AIC.

The BIC is formally defined as

$$\text{BIC} = k\ln(n) - 2\ln(\hat{L}) \tag{2.20}$$

Where

i. $\hat{L}$ = the maximized value of the likelihood function of the model.
ii. $x =$ The observed data.
iii. n= the number of data points in x, the number of observations, or the sample size.
iv. K= the number of parameters estimated by the model.

2.3.12 Limitations of the Bayesian Information Criterion

i. The BIC is only valid for sample size n much larger than the number k of parameters in the model.
ii. The BIC cannot handle complex collection s of models.

2.3.13 Parameters estimation

The Yule-Walker method for ARMA models (walker, 1931) may be one of the best known estimation method in time series analysis. It uses the r_1 estimator to estimate the lag 1 autocorrelation for the AR parameter

$$\hat{\phi}_{r_1} = \frac{\sum_{t=1}^{T-1}(y_t - \bar{y})\,(y_{t+1} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2}, \tag{2.21}$$

Where y_t is the observed score at time t,(t=1,2,3,…,T) and $\bar{y}$ is the mean score over the T observations.

The standard error of the $\hat{\phi}_{r_1}$ is calculated as:

$$SE_{r_1} = \sqrt{\frac{\hat{\sigma}_e^{\,2}}{(T-1)\hat{\sigma}_y^{\,2}}}, \tag{2.22}$$

Where $\hat{\sigma}_y^{\,2}$ is the estimated variance of y_t and $\hat{\sigma}_e^{\,2}$ is the estimated variance of the error.

The estimation of initial values of MA parameters is given as :

$$\varrho_k = \frac{-\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q}{1+{\theta_1}^2 + {\theta_2}^2 + \cdots + {\theta_q}^2} = 0$$

(2.23)

K=1,2,3,…,q

For k=1

$$\varrho_1 = \frac{-\theta_1}{1+{\theta_1}^2} \tag{2.24}$$

2.3.14 Mean Absolute Percentage Error (MAPE)

The MAPE is a measure of how accurate a forecast system is. It measures this accuracy as a percentage. The working principle of MAPE is based on splitting the observed data set into two groups namely; the train data and the test data, the various time series models are each gotten using the train data set and the effectiveness or accuracy of the time series models are determined by forecasting the test data. The model with the lowest MAPE is chosen as the best model.

$$\text{MAPE} = \frac{1}{n} \times \sum_{t=1}^{n} \left| \frac{Y_o - Y_p}{Y_o} \right| \tag{2.25}$$

2.3.15 Model diagnostics

The aim of model diagnostics is to validate the model that must have been identified as the best for predicting a particular time series data. The process of verifying that a statistical model's outputs are appropriate in relation to the actual data generation process is known as model validation. Analysis of the residuals is part of residual diagnostics, which is done to see if they appear to be truly random., i.e. to determine if the residuals are independently and identically distributed with a mean of zero and a constant variance (white noise).the use of statistical tests such as the Box-ljung test can be used to determine the randomness of the residuals of a time series model.

2.3.16 Box-ljung test

The Ljung Box test is a way to test for the absence of serial autocorrelation, up to a specified lag k (Glen, 2014). In essence, it is a test for lack of fit: if the residual autocorrelations are relatively modest, we conclude that the model does not exhibit appreciable lack of fit. This test rejects the null hypothesis (model does not show lack of fit) if the p-value is significant .To run the Ljung Box test, we must calculate the statistic Q. For a time series Y of length n:

$$Q(m) = n(n+2)\sum_{j=1}^{m}\frac{{r_j}^2}{n-j}, \tag{2.26}$$

Where:

i. r_j = The accumulated sample autocorrelations.

ii. m= the time lag.

We reject the null hypothesis and say that the model shows lack of fit if

$$Q > X^2{}_{1-\alpha,h}$$

Where:

$X^2{}_{1-\alpha,h}$= the value that is found on the chi squared distribution table for significance level α and degrees of freedom h.

The condition attached to the use of the box-Jung test on ARIMA models is that the degrees of freedom h must be equal to m-p-q, where p and q are the number of parameters in the ARIMA (p, d, q) model.

2.3.17 Shapiro-Wilks Test

The Shapiro-Wilks test can be used to evaluate if a random sample is taken from a normal distribution. You receive a W value from the test, and lower values mean your sample is not normally distributed (you can reject the null hypothesis that your population is normally distributed if your values are under a certain threshold). The W value formula is as follows.

$$W = \frac{\left(\sum_{i=i}^{n} a_i x_{(i)}\right)^2}{\sum_{i}^{n}(x_i - \bar{x})^2} \tag{2.27}$$

Where:

i. The values of the ordered random sample are xi.

ii. a_i are constants generated from the co-variances, variances and means of the sample (size n) from a normally distributed sample.

2.3.18 Forecasting

Forecasting is using the knowledge we have at one time to estimate what will happen at some future moment of time. (Lewis, 1969).Forecasting uses historical data as inputs to make informed estimates that are predictive in determining the direction of future trends (Alicia, 2020).

Forecasting is a scientific process which aims at reducing the uncertainty of the future state of business and trade, not dependent merely on guess work, but with a sound scientific footing for the decision on the future course of action.

Several areas, including business and industry, government, economics, environmental sciences, medicine, social science, politics, and finance use forecasting as a key tool. Short-term, mid-term, and long-term forecasting issues are frequently separated into these categories. Problems with short-term forecasting require predicting occurrences for a short amount of time (days, weeks, or months). Medium-term forecast extends from one to two years into the future.

Long-term forecasting problems can extend beyond that by many years.

2.3.19 Categories of forecasting methods

1. Qualitative forecasting technique

They are appropriate when there are no historical data available. They are typically used while making long- or intermediate-range choices. Market research, the Delphi method, informed opinion and judgment, and historical life cycle analogies are a few examples of qualitative forecasting techniques.

2. Quantitative forecasting technique

They are models that predict future data based on historical data. When it is acceptable to believe that some of the data's patterns are anticipated to continue into the future, they should be used. These methods are usually applied to short or intermediate range decisions. Examples of quantitative forecasting methods are simple exponential smoothing, Poisson process model based forecasting.

CHAPTER THREE

DATA PRESENTATION AND ANALYSIS

3.1 Introduction

In this chapter,the time series data would be presented and analyzed using the Box-Jenkins time series technique with a view of obtaining a candidate time series model that would be used to predict future values of the Number of POS transactions in Nigeria.A time plot would be constructed to study the components of time series present in the data.The stationarity of the time series data would be tested using the Augmented dickey fuller test,the candidate model would be selected to forecast future values of the Number of POS transactions in Nigeria using several tests such as,the Partial Autcorrelation Function(PACF),Auto-Correlation Function(ACF),Aqaike Information Criterion(AIC) and Bayesian Information Criterion(BIC),and the model would be validated using the Box-ljung test , Mean Absolute Percentage Error (MAPE),and Shapiro-wilks' test. A forecast would be made to future patterns of the time series data.This study used gretl 4.0 and microsoft excel 2010 for data analysis.

3.2 Data Presentation

The data in table 3.1 is a secondary data obtained from the website of Nigeria central Bank (CBN) statistics database from 2014-2019. The data is a monthly record of number of POS transaction from 2014-2019.

Table 3.1:monthly record of number of POS transaction from 2014-2019.

MONTHS						
YEARS	JAN	FEB	MAR	APR	…	DEC
2014	1,349,561	1,433,005	1,576,671	1,624,564	…	2,636,865
2015	2,267,257	2,338,113	2,605,160	2,560,973	…	3,946,721
2016	3,599,828	3,833,884	4,311,844	4,463,599	…	8,994,764
2017	7,946,772	8,606,875	10,093,335	10,420,611	…	17,057,465
2018	16,102,962	16,731,362	20,728,441	20,751,162	…	31,926,618
2019	28,162,746	25,778,644	29,820,754	33,368,778	…	46,138,228

Source of data : http://statistics.cbn.gov.ng/cbn-onlinestats/QueryResultWizard.aspx

3.3 Data Analysis

3.3.1 Time plot

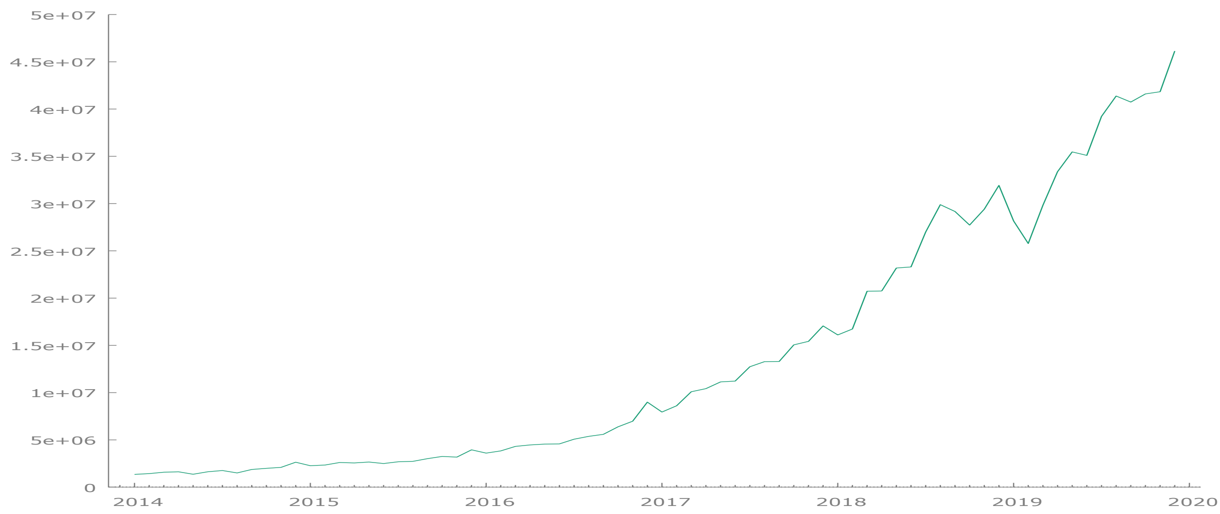

Fig 3.1: time plot of the monthly Number of POS transactions in Nigeria between the period of 2014-2019.

The time series plot in fig 3.1 of the number of POS transactions on the y axis and the corresponding months of the transactions on the x axis recorded at regular time intervals exhibits a clear long increasing pattern over the entire time span.This is sometimes referered to as time trend.There are no obvious outliers in the time series,it can also be seen from the time plot above that the data exhibits a unique feature of a non stationary data which is a non mean reverting nature of the time series data caused by the presence of a unit root process,a unit process is such that the mean of the time series data would continue to grow and fail to converge, consequently,it can be seen from the time plot above that the data set experiences a change in distribution shape when there is a shift in time hence affecting the future behaviour of the data.These properties that the time series possesses would affect the validity of predictions that would be made and hence needs to be dealt with.

3.3.2 Stationarity test

Augmented dickey fuller test for stationarity

The augmented dickey fuller test is a statistical test used to test for stationarity of a time series data and is applied below on the Number of POS transactions to further support the information provided about the time series data by the time plot above.

Hypothesis

H_0: Number of POS transactions is non-stationary

H_1: Number of POS transactions is stationary

Table 3.2: Augmented dickey fuller test statistics value and p-value

TEST	TEST STATISTICS	P-VALUE
ADF	-0.630656	0.9768

It can be seen from table 3.2 that the augmented dickey fuller test statistics (-0.630656) has a p-value (0.9768) that is greater than a 0.05 level of significance (alpha), therefore there is not enough evidence to reject H_0 and the monthly number of POS transactions is said to be non-stationary. The monthly number of POS transactions data needs to be stationary in order to apply Box-Jenkins time series model and hence the data must be differenced and tested again using the augmented dickey fuller test

3.3.3 Differenced series

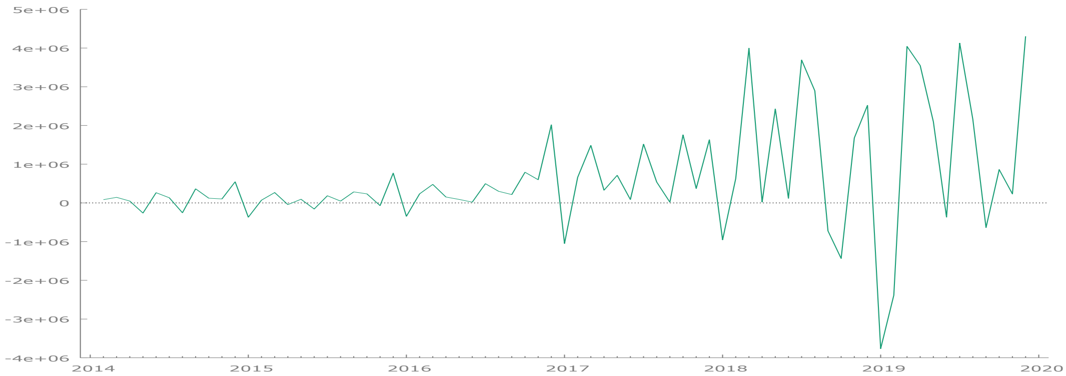

Fig 3.2: time plot of the monthly Number of POS transactions in Nigeria between the period of 2014-2019 after first order difference.

Intuitively from the time plot above, after the first difference has been carried out, it can be seen that there has been a change in the behavior of the trend of the time series data, it is observed that there is a presence of a stationary trend in the time series data, it can be seen from the time plot above that the data set does not experience a change in distribution shape when there is a shift in time but rather the time series data reverts to the mean ,therefore which means that the time series data is now stationary but a statistical test is carried out below to prove the claim.

3.3.4 Stationarity of time series after first order difference

Augmented dickey fuller test for stationarity after first order difference

Hypothesis

H_0: Number of POS transactions is non-stationary

H_1: Number of POS transactions is stationary

Table 3.3: Augmented dickey fuller test statistics value and p-value after first order difference

TEST	TEST STATISTICS	P-VALUE
ADF	-9.08016	2.693e-16

Table 3.3 shows that the augmented dickey fuller test statistics(-9.08016) has a p-value (2.693e-16) that is less than a 0.05 level of significance, therefore the null hypothesis is rejected and the monthly number of POS transactions in Nigeria is said to be stationary after first order difference.

3.3.5 Model identification

The next step after making sure the number of POS transactions data is transformed to a stationary time series data is to identity the proper model that would be used to forecast future values of the data.

3.3.6 The PACF and ACF Functions

The auto correlation function and partial auto correlation functions provide guidance for the order that is appropriate for the time series model.

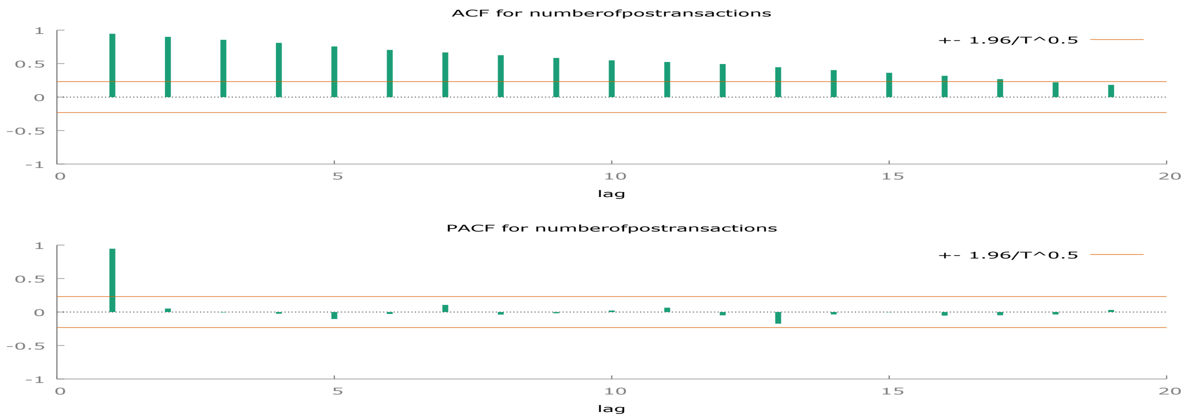

Fig 3.3 : ACF and PACF for number of POS transactions

The ACF and PACF clearly show significant values. There are few notable patterns in the ACF and PACF which can help choose preliminary orders of the ARIMA model.

1. The ACF function shows a slower, gentle decline. This is sometimes referred to as a geometric pattern.
2. The PACF shows a sharp cut off after the first lag.

The combination of the two features above is indicative of an AR model.

3.3.7 The Akaike Information Criterion and Bayesian Information Criterion

Table 3.4: AIC and BIC values for candidate models

MODEL	AIC	BIC
ARIMA(1,1,0)	2212.333	28.43
ARIMA(2,1,0)	2203.229	28.346
ARIMA(3,1,0)	2204.241	28.407
ARIMA(4,1,0)	2203.940	28.448
ARIMA(5,1,0)	2204.636	28.504
ARIMA(6,1,0)	2172.326	28.054

According to AIC and BIC from table 3.4, the model with the lowest AIC (2172.326) and lowest BIC (28.054) is ARIMA (6,1,0).

AR (6) MODEL

$$Y_t = C + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \phi_4 Y_{t-4} + \phi_5 Y_{t-5} + \phi_6 Y_{t-6} + \varepsilon_t$$

3.3.8 Mean Absolute Percentage Error

The mean absolute percentage error is a measure of how accurate a forecast system is. It measures this accuracy as a percentage. In this study, the first five years of the time series data was used as the train data and the sixth year observations were used as the test data.

Table 3.5: MAPE values for candidate models

MODELS	MAPE
ARIMA(1,1,0)	12.157
ARIMA(2,1,0)	10.896
ARIMA(3,1,0)	11.61
ARIMA(4,1,0)	11.982
ARIMA(5,1,0)	8.7109
ARIMA(6,1,0)	7.914

From the table above, the model with the lowest Mean absolute percentage error is ARIMA(6,1,0) and is chosen as the best model.

3.3.9 Time plot of forecast using ARIMA(6,1,0) model on the train data to predict the test data

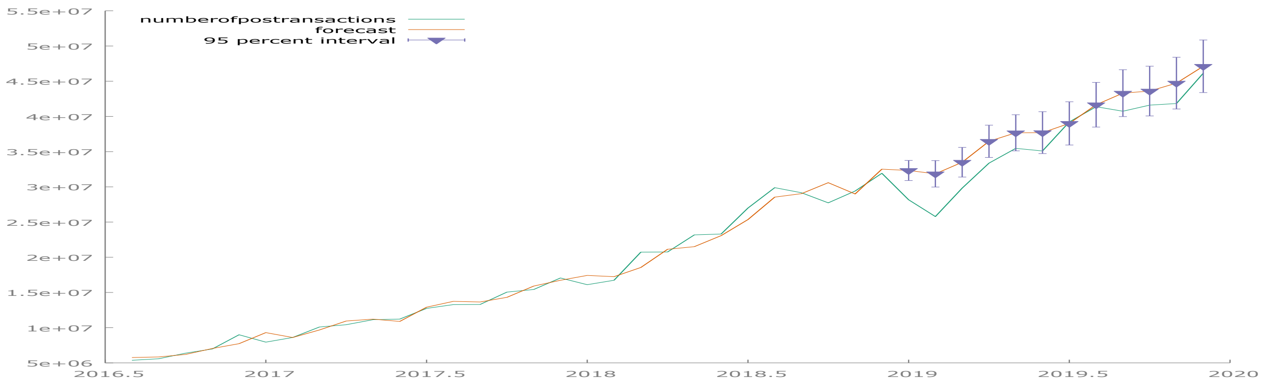

Fig 3.4: Time plot of forecast using part of the data as train data to predict the remaining part of the data.

Fig 3.4 is a time plot that displays a graphical representation of the forecast (red line) against the observed values (green line), the error of this model is 7.914% which makes the model approximately 92% accurate.

3.3.10 Model Estimation of ARIMA (6, 1, 0)

Table 3.6: Estimation of parameters for ARIMA (6, 1, 0) model

	COEFFICIENTS	STANDARD ERROR	P VALUE
Constant	−238671	131558	0.0696
phi_1	−0.0710295	0.0901191	0.4306
phi_2	−0.161207	0.0926271	0.0818
phi_3	−0.0606610	0.0973172	0.5331
phi_4	0.0130881	0.0946294	0.8900
phi_5	0.0756058	0.0906147	0.4041
phi_6	−0.641975	0.0900298	9.99e-013

A Study of table 3.6 at 5% level of significance shows that at least one of the lag coefficients of the AR (6) model was significant hence which implies that the general model is a significant model. The fitted model is obtained below.

$$Y_t = -238671 - 0.0710295\, Y_{t-1} - 0.161207 Y_{t-2} - 0.0606610 Y_{t-3} + 0.0130881 Y_{t-4} + 0.0756058 Y_{t-5} - 0.641975 Y_{t-6}$$

3.3.11 Model diagnostics

The AR model is further investigated for its significance using the residuals that arise from the data by applying this particular model. The test determines whether or not the autocorrelations for the errors (residuals) are non-zero. This test is called the Box-ljung test.

The box-ljung test

Hypothesis

H_0 = no correlation among the residuals of the fitted model

H_1= there is presence of correlation among the residuals of the fitted model.

Table 3.7: Box-Ljung test statistics value and p-value

TEST	TEST STATISTICS	P VALUE
Ljung test	7.48746	0.2781

The Box-Ljung test statistics (7.48746) has a p-value (0.2781) which is greater than a 0.05 level of significance, therefore H_0 is not rejected. According to the box-ljung test, the residuals of the fitted AR(6) model have no auto correlation which implies the AR(6) model does not show significant lack of fit.

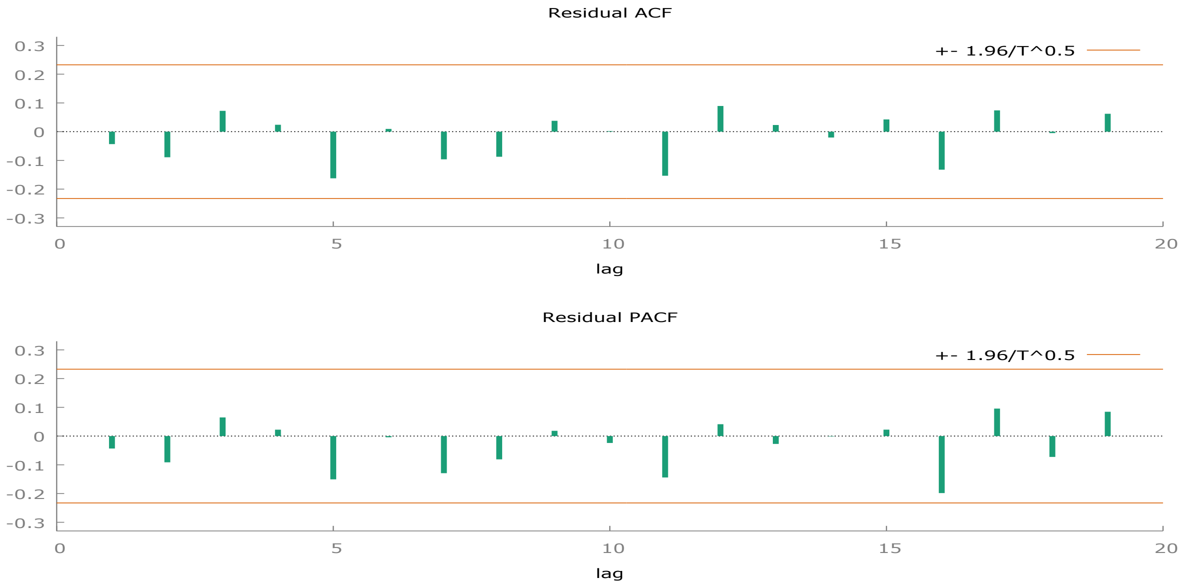

Fig 3.5: ACF and PACF for the residuals of the fitted AR(6) model

The ACF and PACF of the residuals from the AR model falls below the significance line at all lag values (the red line) in these graphs. This implies that there are no significant serial correlations between the residuals of the chosen AR model, and hence can be used to predict future values of the number of POS transactions in Nigeria.

3.3.12 Forecasts

Table 3.8: Predicted monthly number of POS transactions for 2020 and 2021

MONTHS	PREDICTION	L.C.L	U.C.L
JANUARY 2020	45790022	43851593	47728450
FEBUARY 2020	46270434	43624738	48916129
MARCH 2020	49137280	46097008	52177552
APRIL 2020	51067424	47710180	54424667
MAY 2020	53322817	49640849	57004786
JUNE 2020	52646452	48598957	56693947
JULY 2020	55307186	51239990	59374382
AUGUST 2020	57862010	53761331	61962688
SEPTEMBER 2020	58506923	54308532	62705315
OCTOBER 2020	59732137	55419103	64045171
NOVEMBER 2020	60885539	56489002	65282075
DECEMBER 2020	64242596	59802346	68682846
JANUARY 2021	65285196	60601157	69968236
FEBUARY 2021	66114526	61218308	71010744
MARCH 2021	68510272	63499243	73521300
APRIL 2021	70662344	65556277	755768411
MAY 2021	72816813	67587376	78046249
JUNE 2021	73364211	67958849	78769573
JULY 2021	75573218	70114170	81032266
AUGUST 2021	78217568	72704056	83731079
SEPTEMBER 2021	79678217	74059643	85296791
OCTOBER 2021	81229746	75491812	86967680
NOVEMBER 2021	82879985	77051613	88708357
DECEMBER 2021	85786110	79902832	91669388

3.3.13 Time plot of forecasts

.

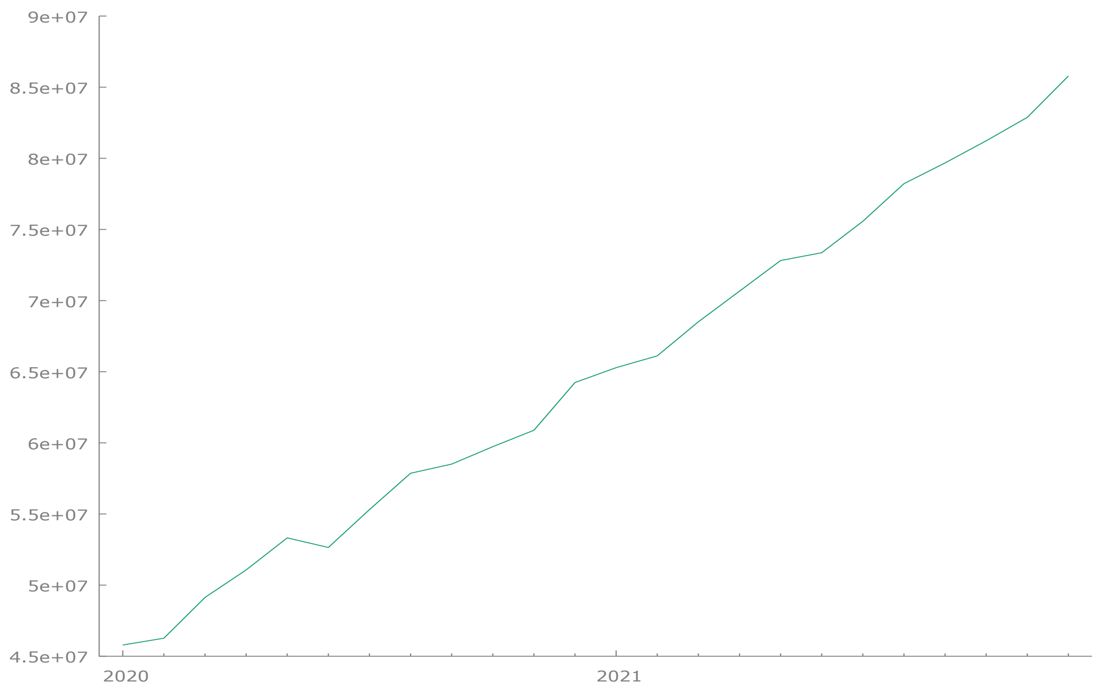

Fig 3.6 :Time plot of forecasted monthly number of POS transactions in Nigeria

From the time plot above, it can be seen that the number of POS transactions is expected to continue to increase for the next few years to come, this behavior is quite expected because it has been discovered through recent studies that Nigerians are starting to pay attention to electronic means of payment rather than the corresponding traditional payment systems.

3.3.14: Shapiro-Wilks' Test For Normality

Hypothesis

H_0: Forecast data is normally distributed

H_1: Forecast data is not normally distributed

Table 3.9: Shapiro-wilks' test statistics and p-value

TEST	TEST STATISTICS	P VALUE
Shapiro-wilks' test	0.958858	0.415911

Table 3.9 shows that the test statistics (0.9588) has a p- value (0.4159) which is greater than the level of significance (0.05).Therefore, the null hypothesis is not rejected and we conclude that the forecast data is normally distributed.

CHAPTER FOUR

SUMMARY AND CONCLUSION

4.1 Introduction

This chapter entails the summary of the previous chapters and conclusion based on the results obtained from the analysis of the data.

4.2 Summary

Before the development of electronic means of transactions, people often have to go to the bank before making transactions such as withdrawal of money, deposits, sending money to another person etc, these processes sometimes are inconvenient, slow, or stressful. The introduction of electronic means of payments such as the Point of Sale (POS) enables users to easily carry out daily transactions at their own comfortable time, users can even pay for goods they buy at the supermarket through the use of POS digital machine. POS is an electronic means of payment system that is used to carry out multiple transactions. The history of POS systems began with the invention of cash register. James Ritty invented the first cash register in 1879 and eventually sold it to john H. Patterson ,who established the National cash register in 1884 and ultimately helped transformed it into a digital machine. POS was introduced to Nigerians by Nigeria Central Bank (CBN) in 2012, the initiative has recorded a remarkable success since its introduction, the deployment of POS terminals rose from 5000 to 153,167 in April 2014.The main focus of this project was to study the pattern of usage of POS as a means of payment system by Nigerians, provide an appropriate time series model to forecast the number of POS transactions in Nigeria and predict the number of POS transactions for the next two years. The data used for this project is a secondary data from the website of CBN statistics data base from 2014-2019. A time series data is simply a sequence of data points that occur in successive order over a period of time, A time series data can either be stationary or non-stationary. Stationary time series data is considered the best type of data to use for prediction because of its ability to maintain constant mean and variance with a unit change in time; however the components of time series such as the trend, seasonality, cyclical variation and irregular variation are factors that restrict a time series data from being stationary. The time plot of the number of POS transactions was constructed and examined for the possible component(s) of time series. An upward trend was discovered in the data using the time plot, the Augmented Dickey Fuller (ADF) test was carried out on the time series data to test for stationarity of the data .The ADF test is a statistical test used to determine the stationarity of a time series data, the ADF test returned a test statistics value of -0.630656 and a p-value (0.9768) which is greater than the level of significance(0.05),therefore, there is not enough evidence to reject the general claim of the ADF test (time series data are non-stationary) and hence it is concluded that the time series data collected for the purpose of this study is non-stationary. A stationary data is considered a better option to perform a time series analysis on therefore differencing technique was used on the non – stationary data. Differencing technique is a method of data transformation that is usually applied to non-stationary time series data to make the data stationary. First order difference was carried out on the number of POS transaction data and it was again tested for stationarity using the ADF test. The ADF test statistics value (-9.08016) of the differenced series had a p-value (2.693e-16) which is less than the level of significance (0.05),therefore, there is enough evidence to reject the general claim of the ADF test and hence it was concluded that the number of POS transactions was made stationary after first difference. The Box Jenkins ARIMA time series model was developed to predict the future number of POS transactions. ARIMA (Auto-regressive integrated moving average) is made up of two processes; the Autoregressive AR and the moving average MA.AR model believes that data values at current time spots are built over data values at previous time spots while MA model believes that the current data value is results of previous expected events. The ARIMA model can sometimes be an AR, MA or ARMA model depending on the pattern shown by the ACF and PACF plot. Autocorrelation function(ACF) and the partial auto correlation function(PACF) was plotted to help determine the order of an appropriate time series model, the ACF plot showed a steady decline between successive lags while the PACF plot showed a sharp cut off after the first lag, the pattern of the two plots suggested an

Autoregressive (AR) model. The Akaike information criterion (AIC) and Bayesian Information Criterion (BIC) was used to filter one of many possible AR (Auto regressive) model to be used to predict future values of the monthly number of POS transactions and the model chosen to fit the number of POS transactions was the ARIMA (6,1,0) model having the lowest AIC (2172.326) and the lowest BIC (28.054). Before using any time series model to forecast future values, it is important to test the accuracy of the time series model on the current values of the series. This project adopted the Mean Absolute Percentage Error (MAPE) to test for the accuracy of various AR model using the first 5 years of the data as train data to predict the last year of the data. ARIMA (6,1, 0) model was the model with the lowest MAPE (7.914) and was selected as the model with the highest accuracy to predict number of POS transactions for the next two years. The model was validated by the Box ljung test and Shapiro – Wilks' test and was used to forecast for 2020 and 2021.

4.3 Conclusion

POS businesses in the country have formed a major source of employment for Nigerians, especially youths. It can be concluded based on AIC, BIC and MAPE used in this project that a fitted ARIMA(6,1,0)= $-238671 - 0.0710295\,Y_{t-1} - 0.161207Y_{t-2} - 0.0606610Y_{t-3} + 0.0130881Y_{t-4} + 0.0756058Y_{t-5} - 0.641975Y_{t-6}$ is the most appropriate model that to predict the number of POS transactions in Nigeria. It is expected that there would be increased use of POS as means of payment system in Nigeria for 2020 and 2021 based on the time plot of the forecast model.

REFERENCES

Adam, H. (2020), “Point of sale”. Retrieved December 2021 from https://www.investopedia.com/terms/p/point-of-sale.asp.

Agresti A. (1990), “Categorical Data Analysis. John Wiley and Sons, New York”.

Beyer, W.H. CRC. (2002), “Standard Mathematical Tables, 31st ed. Boca Raton, FL:CRC Press,pp.536 and 571”.

Brockwell, P. J. and Davis, A.R. (1996), “Introduction to Time series and Forecasting, Second Edition”.

Brownlee, J. (2017), How to remove trends and stationarity with a Difference Transform in python”. Retrieved December 2021 from https://machinelearningmastery.com/remove-trends-seasonality-difference-transform-python/

Chatfield, C. (1995), “The Analysis of Time series.(4th edition) .chapman &Hall”. Retrieved December 2021 from https://www.statisticshowto.com/timeplot/

Elliot, G., Rothenberg, T.J., and Stock, J.H.(1996), “Efficient Tests for an Autoregressive Unit Root”.

Eric, C. (2019), “Introduction to the Fundamentals of Time Series Data and Analysis”. Retrieved December 2021 from https://www.aptech.com/blog/introduction-to-the-fundamentals-of-time-series-data-and-analysis/

Everitt, B.S., and Skrondal, A. (2010), “The Cambridge Dictionary of Statistics, Cambridge University Press”.

Fuller, W.A.(1976), “Introduction to statistical Time series. New york: John Wiley and Sons”. Retrieved December 2021 from http://www.ru.ac.bd/stat/wp-content/uploads/sites/25/2019/03/504_08_Fuller_introduction-to-Statistical-Time-Series-Second-Edition.pdf

G.E.P. Box, and G. Jenkins. (1970), “Time series Analysis, Forecasting and control, San Francisco, CA”.

Glen, S. (2013), “Autoregressive Model: Definition & The AR process” From StatisticsHowTo.com: Elementary Statistics for the rest of us. https://www.statisticshowto.com/autoregressive-model/

Glen, **S. (2013)**, "Shapiro-Wilk Test: Definition, How to Run it in SPSS" From StatisticsHowTo.com: Elementary Statistics for the rest of us. Retrieved January 2022 from https://www.statisticshowto.com/shapiro-wilk-test/

Hamilton, D.J. (1994), “Time series analysis by Hamilton”. Retrieved January 2022 from http://www.ru.ac.bd/stat/wp-content/uploads/sites/25/2019/03/504_02_Hamilton_Time-Series-Analysis.pdf

Hyndman, R. J. and Athanasopoulous, **G. (2013),** “Forecasting: principles and practice”.

Jingbo, S., Yanan, F., and Ousseni, B. (2016), “Compensation Mechanism for Early Termination of Highway BOT projects based on ARMA model. International Journal of Architecture, Engineering and construction”.

Krone, T.A, and Timmerman, M.EA (2017), “Comparative simulation study of AR(1) estimators in short time series”.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., and Shin, Y. (1992), “Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of econometrics,54(1-3),159-178”. Retrieved December 2021 from https://doi.org/10.1016/0304-4076(92)90104-Y

Markoidakis, Wheelwright and Willey, J. (1978), “Forecasting methods and applications”.

Oturu, D. (2020), “Categorization of Nigerian payment systems”.

Petris, G.et.al.(2009), “Dynamic Linear Models with R. Springer Science & Business Media”. Retrieved December 2021 from https://www.statisticshowto.com/arma-model/

Samuel, O. (2021), “Christmas spending pushes POS transactions in Nigeria to all-time high of N6.4 trillion in 2021”.

Vijay, K. G. (2021). “Statistical tests to check stationarity in Time Series part 1”.Retrieved January 2022 from https://www.analyticsvidhya.com/blog/2021/06/statistical-tests-to-check-stationarity-in-time-series-part-1/