Bindura University of Science Education

FACULTY OF SCIENCE

DEPARTMENT OF STATISTICS AND MATHEMATICS

A TIME SERIES ANALYSIS OF ONLINE FOOD ORDERING SALES (DIAL A DELIVERY) IN ZIMBABWE (2020 – 2023).

BY

GLORIA TANAKA SHOKO (B201095B)

A DISSERTATION SUBMITTED IN PARTIAL FUFILMENT OF THE REQUIREMENTS FOR BSC.HONOURS IN STATISTICS AND MATHEMATICS

SUPERVISOR: DR T.W. MAPUWEI

JUNE 2024

APPROVAL FORM

The undersigned certify that they have read and recommended to the Bindura University of Science Education for acceptance of a dissertation entitled "A TIME SERIES ANALYSIS OF ONLINE FOOD ORDERING SALES (DIAL A DELIVERY) IN ZIMBABWE. Submitted by GLORIA TANAKA SHOKO, Registration Number B201095B in partial fulfillment of the requirements for the Bachelor of Science Honor's degree in Statistics and Financial Mathematics.

GLORIA TANAKA SHOKO …………...  10/06/2024

B201095B Signature Date

Certified by

DR T.W. MAPUWEI ……………  10/06/2024

SUPERVISOR Signature Date

Certified

DR. M. MAGODORA ……………  10/06/2024

Chairman of Department Signature Date

DECLARATION

I, GLORIA TANAKA SHOKO hereby declare that this submission is my own work apart from the references of other people's work which has duly been acknowledged. I hereby declare that; this work has neither been presented in whole nor in part for any degree at this university or elsewhere.

Author: GLORIA TANAKA SHOKO

Registration Number: B201095B

Signature: ………Shoko.

Date: 10 June 2024

DEDICATION

I dedicate this dissertation to my lovely mother, grandmother and father who have made sacrifices towards my personal and professional endeavors.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor, Dr. T. W. Mapuwei for his unwavering support, insightful guidance, and constructive feedback throughout the duration of this project. Your expertise and dedication have been invaluable in shaping this piece of work. I am also profoundly grateful to my family for their constant encouragement and understanding. Your support has been a pillar of strength for me, and could not have completed this project without your love and patience. Special thanks to my colleagues and friends for their encouragement, and to everyone who contributed their time, knowledge, and resources to make this project possible. Your assistance and collaboration have been greatly appreciated.

ABSTRACT

The food sector has changed dramatically as a result of the internet meal delivery services' explosive expansion, especially in developing nations like Zimbabwe. The goal of this research project is to provide a thorough time series analysis of online food ordering sales for Dial a Delivery, one of Zimbabwe's top service providers. Finding sales trends, seasonal patterns, and possible forecasting models is the main goal in order to enhance corporate strategy and gain a better understanding of consumer behavior. The study employs a variety of time series analysis approaches, such as decomposition, autocorrelation analysis, and ARIMA (Autoregressive Integrated Moving Average) modeling, using monthly sales data from January 2018 to December 2023. The information was taken from the sales records of Dial a Delivery, offering a substantial dataset for careful examination. The results reveal significant variations in sales, with peak periods corresponding to major holidays and festive seasons. Additionally, the analysis identifies a positive trend in sales growth over the study period, reflecting the increasing adoption of online food ordering services in Zimbabwe. The ARIMA (3,1,0) model developed in this study demonstrates high accuracy in forecasting future sales, with a mean absolute percentage error (MAPE) of 10.2%. The researcher advises other students to look into the application of time series in online sales employing important techniques that were overlooked in the study. In order to increase sales throughout the winter, Dial a Delivery might think about stepping up their marketing initiatives and promos. In order to satisfy the increased demand throughout the summer, the corporation should keep enough inventory levels. Regular updates of the ARIMA model are necessary to guarantee precise projections and to accommodate any modifications in the market.

Table of Contents

APPROVAL FORM ii

DECLARATION iii

DEDICATION iv

ACKNOWLEDGEMENTS v

ABSTRACT vi

LIST OF FIGURES ix

LIST OF TABLES x

CHAPTER ONE 1

1.0 Introduction 1

1.1 Background of the Study 1

1.2 Statement of the Problem 2

1.3 Research Objectives 2

1.4 Research Questions 3

1.5 Scope of the Study 3

1.6 Significance of the Study 3

1.7 Limitations of the Study 3

CHAPTER TWO 5

2.0 Introduction 5

2.1 Theoretical Literature 5

2.2 Empirical Literature 6

2.3 Research Gap 7

2.4 Proposed Conceptual Model 8

2.5 Conclusions 9

CHAPTER 3: REASERCH METHODOLOGY 11

3.0 Introduction 11
3.1 Research Design 11
3.2 Data Sources 11
3.3 Target Population and Sampling Procedures 11
3.4 Research Instruments 12
3.5 Methods of Data Collection 12
3.6 Description of Variables and Expected Relationships 12
3.7 Data Analysis Procedures 12
3.72 Analytical Model 13
3.73 Model Validation (Fitness) Tests 13
3.8 Ethical Considerations 13
CHAPTER FOUR: DATA PRESENTATION, ANALYSIS AND INTERPRETATION 14
4.0 Introduction 14
4.1 Descriptive Statistics/Summary Statistics 14
4.2 Pre-tests /Diagnostic tests 16
4.3 Model output /Results 20
4.6 Conclusion 25
CHAPTER FIVE 26
5.0 Introduction 26
5.1 Summary of Findings 26
5.2 Conclusions 26
5.3: Recommendations 26
5.4 Areas for further research 26
5.5 Conclusion 27
References 28

LIST OF FIGURES

Figure 4. 1 Online Sales Trend 16
Figure 4. 2 ACF Plot for stationarity Test 17
Figure 4. 3 PACF Plot for stationarity Test 17
Figure 4. 4 Achieving Stationary 18
Figure 4. 5 ACF Plot for differenced data 19
Figure 4. 6 PACF for differenced data.. 19
Figure 4. 7 Model residuals 21
Figure 4. 8 ACF of Residuals 22
Figure 4. 9 PACF of Residuals 22
Figure 4. 10 Histogram of Residuals 23
Figure 4. 11 Histogram with density curve... 23
Figure 4. 12 Q-Q plot.. 24
Figure 4. 13 Sales forecast.. 25

LIST OF TABLES

Table 4. 1 Descriptive Statistics.. 14
Table 4. 2 ADF Test.. 18
Table 4. 3 ADF for Differenced data 20

CHAPTER ONE

1.0 Introduction

An overview of the main ideas, background information, and importance of time series analysis in online food ordering sales (dial a delivery) for Zimbabwe are given in this chapter. Its purpose is to lay the groundwork for the following chapters by providing an overview of the main goals, parameters, and organization of this project. This chapter opens with a thorough analysis of the literature, highlighting the most important discoveries and influential figures in the study. The project's motive is then covered, with a focus on its applicability and possible impact.

1.1 Background of the Study

The growth of the internet and mobile devices has led to a significant increase in online ordering sales in Zimbabwe (Chase, 1996). Online sales have spiked in Zimbabwe and their potential impact on the business environment is significant. Online sales have expanded customer reach, job creation and data driven decision making. Time series is a series of numerical values obtained from consecutive periods, usually with equal intervals of time between them (Ruey, 2010). Time series analysis allows businesses to examine historical data and identify patterns, such as seasonal fluctuations, long-term trends, and cyclic behavior. Understanding these patterns, businesses can adjust their strategies to optimize operations, inventory management, marketing efforts, and resource allocation.

Time series analysis enables businesses to develop accurate forecasts for future trends, such as sales, demand, and customer behavior. Analyzing sales data, businesses can apply forecasting techniques like ARIMA models, exponential smoothing, to predict future sales volumes, identify peak periods, and anticipate demand fluctuations. This helps improve inventory management, staffing decisions, order fulfillment, overall operational efficiency. The aim of this study is to conduct a comprehensive time series analysis of the online ordering sales for Dial a Delivery in Zimbabwe, with a specific focus on the sales performance of Sambisa Brands. As the e-commerce sector continues to grow globally, understanding the dynamics of online sales becomes crucial for businesses to make informed decisions and optimize their strategies. By examining the historical sales data, this study seeks to identify trends, patterns, and potential factors that influence online ordering sales in Zimbabwe and provide valuable insights for the Simbisa Brand.

Online ordering has gained significant popularity in Zimbabwe, particularly in the food and beverage industry, due to its convenience and accessibility (Bhatnagar, Misra & Rao, 2000). Simbisa Brands, as a prominent player in the fast-food market, has invested in Dial a Delivery as their online sales platform, catering to a wide range of customers seeking quick and reliable food delivery services. Analyzing the time series data for online ordering sales can offer valuable insights into the performance of Dial a Delivery and contribute to enhancing Sambisa Brands' overall business strategy.

Online orders provide customers with the convenience of ordering food from the comfort of their homes (Statista, 2021). Ordering online has eliminated the need for physical travel to restaurants saving time. Online sales have revolutionized the food ordering industry, providing a seamless user experience. As technology continues to advance and consumer demand for convenience increases, the significance of online sales in the food business is expected to continue expanding, (McKinsey & Company, (2019). However, there has been little research on the relationship between online ordering sales and other variables.

1.2 Statement of the Problem

To understand the trends, seasonality and forecasted future sales patterns of Sambisa brands online ordering sales in Zimbabwe. This research seeks to understand and analyze the historical sales patterns and forecast future sales for Dial a Delivery, the online ordering service provided by Simbisa Brands in Zimbabwe, in order to improve sales forecasting accuracy and identify factors that impact sales performance.

1.3 Research Objectives

The main objectives of this study are:

1. To analyze historical trends in Simbisa Brands' online ordering sales in Zimbabwe.
2. To investigate the presence seasonality in the online ordering sales data.
3. To develop accurate time series forecasting models for predicting future sales pattern.

1.4 Research Questions

The research questions for this study are as follows:

1. What are the historical trends in Simbisa Brands' online ordering sales in Zimbabwe?
2. Is there any seasonality present in the online ordering sales data?
3. How can accurate time series forecasting models be developed to predict future sales patterns?

1.5 Scope of the Study

The scope of the study will be to analyze the trends and patterns of online sales in Zimbabwe over a period of time. This will be done using a time series analysis approach, which involves examining data that is collected over time.

1.6 Significance of the Study

The significance of conducting a time series analysis of online ordering in Zimbabwe lies in its ability to provide insights into consumer behavior, forecast demand, identify growth opportunities, evaluate external factors, and ultimately improve decision-making for businesses in this field**.** The study will focus on online sales.

1.7 Limitations of the Study

In this investigation, the researcher encountered certain challenges. The availability of reliable and comprehensive data on online sales in Zimbabwe is limited. This is due to the fact that the online sales industry is still relatively new in Zimbabwe and there is no central repository for data. The quality of the available data is also limited. This is due to the fact that many online retailers in Zimbabwe do not track their sales data in a standardized way. The time period of the study will be limited to the last five years. This is due to the fact that online sales data for Zimbabwe is not available for a longer period of time.

1.8 Definition of Terms

Time series

A time series is a collection of data acquired via the systematic measurement of a single variable over an extended period of time (IBM, 2013). A sequence of values observed throughout time is called a time series (NCSS, 2013). Mathematically, it is described as a set of vectors $x(t)$, $t = 0, 1, 2\ldots$ where t denotes the time passed. It is a sequential sequence of data points, usually measured across successive times (Adhikari, 2014).

CHAPTER TWO: LITERARURE REVIEW

2.0 Introduction

This analysis offers an in-depth overview of the current research landscape in time series analysis of online food sales. The introduction of e-commerce has significantly changed the food retailing environment and resulted in a rise in sales of online meal ordering. It's becoming more and more important to comprehend and predict the dynamics of online food sales as customer behavior changes in response to new technology and evolving tastes. This chapter presents a theoretical framework and conceptual model for time series analysis of online food ordering sales, aiming to integrate exogenous factors and advanced machine learning techniques to enhance predictive accuracy and derive actionable insights for e-commerce food retailers.

2.1 Theoretical Literature

Online Food Ordering

The rapid growth of the online food ordering industry has sparked significant academic interest in understanding the various factors that drive consumer adoption and usage of these platforms. Researchers have drawn upon a range of theoretical frameworks to offer a thorough comprehension of this phenomenon. Technology Acceptance Model (TAM), put out by Davis (1989), is one of the ideas that is most frequently used in this field. According to TAM, perceived utility and perceived ease of use are the primary factors that determine a person's intention to use a technology, which in turn influences their actual usage behavior. TAM describes how consumers come to embrace and use a particular technology. This model has been widely applied to comprehend the variables that affect customers' approval of and inclination to use online meal ordering services.

The Diffusion of Innovations idea, created by Rogers in 1962, is another well-known idea. This theory explains how novel concepts, innovations, or goods gradually permeate a social structure. It names a number of variables that affect the rate of adoption, including trialability, observability, complexity, compatibility, and relative advantage. This framework can be utilized to comprehend the aspects that lead to the proliferation of online meal ordering services within the industry, as well as their adoption and growth. The context of online meal ordering has also been examined using Ajzen's (1991) Theory of Planned Behavior (TPB). According to TPB, a person's attitude, subjective norms, and perceived behavioral control all have an impact on their purpose, which in turn determines their conduct. In the context of online meal ordering, this theory has been applied to investigate the variables that affect customers' intentions and real behavior, including their perceptions of the practicality and advantages of these platforms.

Mehrabian and Russell (1974) created the Stimulus-Organism-Response (S-O-R) paradigm, which has been used to analyze how environmental cues affect consumer behavior in the online food ordering market. According to this theory, an individual's internal state (O), which in turn influences their behavioral reactions (R), can be influenced by environmental stimuli (S). Using this approach, researchers have investigated how customers' emotional and cognitive states might be affected by the characteristics and design of online food ordering platforms (as environmental stimuli), which in turn can affect consumers' ordering behavior.

Additionally, frameworks focused on service quality and customer satisfaction, such as SERVQUAL The Kano Model (Kano et al., 1984) and Parasuraman et al. (1988), have been utilized to assess the quality of service provided by online food ordering platforms and its impact on customer satisfaction and loyalty. Finally, theories on value co-creation (Prahalad & Ramaswamy, 2004) and platform ecosystems (Gawer & Cusumano, 2014) have been applied to understand the collaborative and interactive nature of online food ordering platforms. These theories provide insights into how these platforms can facilitate value co-creation between customers, restaurants, and other ecosystem participants, leading to enhanced customer experiences and business performance.

2.2 Empirical Literature

Due to shifting consumer desires for convenience and on-demand services, as well as rising internet and smartphone usage, online meal ordering has become a global phenomenon. Numerous empirical research has looked at the patterns and factors that influence sales of online meal ordering. A study by Bhatnagar, Misra, and Rao (2000) analyzed the factors influencing consumer adoption of online grocery shopping in the United States. They found that perceived convenience, time savings, and lack of in-store effort were significant drivers of online grocery purchases. Consumers valued the ability to order from the comfort of their homes and avoid the hassle of visiting physical stores. The researchers also highlighted the importance of website design and user-friendliness in facilitating a positive online shopping experience.

Similarly, subjective norms, perceived enjoyment, and perceived ease of use were found to be important predictors of intention to use online meal ordering platforms in a study conducted on Chinese consumers by Jiang, Yang, and Jun (2013). Because consumers who felt confidence in their abilities to navigate the online ordering procedure were more inclined to use such services, the researchers highlighted the importance of perceived behavioral control.

Extending the research to a developing country context, Ryu and Jang (2016) investigated the determinants of online food delivery service usage in South Korea. Their findings indicated that food quality, delivery service quality, and user interface design were important factors influencing consumer adoption and continued usage of online food ordering platforms. Customers valued the ability to access a wide range of food options, receive timely and reliable delivery, and interact with a user-friendly digital platform.

In Zimbabwe, limited empirical research has been conducted on online food ordering sales. A study by Chidovi and Muzvidziwa (2020) explored the adoption of online food delivery services in Harare, Zimbabwe. They found that perceived convenience, perceived usefulness, and social influence were significant predictors of consumer intentions to use online food delivery platforms. The researchers also highlighted the importance of trust in the online service provider and the availability of payment options as key factors influencing consumer behavior.

2.3 Research Gap

The time series analysis of online food ordering has become increasingly relevant due to the rising prominence of e-commerce in the food industry. While existing literature has made notable contributions to understanding and forecasting online food sales dynamics, there are several potential research gaps that warrant further investigation. The incorporation of outside variables, such the weather, social gatherings, and public holidays, into time series models for online meal ordering is an area of unmet research need. Comprehending the ways in which exogenous variables influence the temporal patterns of online food purchasing behavior may improve the precision and dependability of forecasting models, therefore offering e-commerce food sellers useful information for optimizing their sales tactics.

Investigating the use of cutting-edge machine learning methods for time series analysis of online food ordering, such as recurrent neural networks and deep learning models. Examining how well these cutting-edge methods capture the complex dynamics of online food sales data may provide new insights into predictive modeling and improve the forecast accuracy of sales in the e-commerce food industry.

Examining the spatiotemporal aspects of online food ordering behavior. Investigating how geographical variations and regional factors influence the temporal patterns of online food sales could provide valuable insights for e-commerce platforms and food retailers targeting specific locations, thereby facilitating targeted marketing and sales strategies.

The impact of consumer sentiment and reviews on online food ordering. Analyzing how consumer feedback and sentiment data contribute to the temporal fluctuations in online food sales could offer valuable insights into the interplay between customer perceptions and purchasing behavior in the digital marketplace.

2.4 Proposed Conceptual Model

The increasing prevalence of e-commerce in the food industry has prompted a growing interest in understanding and forecasting the dynamics of online food ordering sales. This paper proposes a conceptual model for time series analysis of online food ordering sales, aiming to integrate exogenous factors and advanced machine learning techniques to enhance predictive accuracy and provide valuable insights for e-commerce food retailers. By incorporating external variables, such as weather conditions, social events, and consumer sentiment, into the time series analysis, the proposed model seeks to capture the complex interactions that influence online food sales dynamics. Furthermore, the integration of advanced deep learning models and recurrent neural networks are two examples of machine learning methods that have the ability to increase forecasting accuracy and reveal complex temporal patterns in online food ordering behavior.

Proposed Conceptual Model:

The conceptual model for time series analysis of online food ordering sales comprises the following key components:

Data Collection and Preprocessing: The process involves collecting historical online food ordering sales data, as well as relevant exogenous variables, and preprocessing the data to ensure its suitability for time series analysis.

Integration of Exogenous Factors: The model integrates exogenous factors, such as weather data, public holidays, social events, and consumer sentiment, to explore their impact on online food sales dynamics.

Time Series Analysis Techniques: To extract and comprehend the temporal patterns in online food ordering sales data, the model makes use of sophisticated time series analysis techniques, such as wavelet analysis, spectral analysis, and ARIMA models.

Advanced Machine Learning methods: To capture complicated temporal correlations and improve forecasting accuracy, the model integrates cutting-edge machine learning methods, such as recurrent neural networks and deep learning models (e.g., LSTM networks).

Forecasting and Assessment: The model forecasts sales of online food ordering by utilizing machine learning and integrated time series analysis. It then assesses the predictive performance with the use of relevant metrics.

Insights and Implications: The model aims to provide actionable insights for e-commerce food retailers by uncovering the impact of exogenous factors and showcasing the potential of advanced machine learning techniques in optimizing sales strategies.

2.5 Conclusions

This chapter presented a comprehensive overview of the theoretical foundation and conceptual framework for the proposed model of time series analysis of online food ordering sales. The chapter began by highlighting the increasing significance of e-commerce in the food industry and the growing need for advanced analytical approaches to understand and forecast online food sales dynamics. Building upon the literature on time series analysis, machine learning, and e-commerce, the conceptual model integrates key components, including data collection and preprocessing, integration of exogenous factors, advanced time series analysis techniques, and machine learning algorithms. By synthesizing these components, the model seeks to provide actionable insights for e-commerce food retailers and offer a novel approach to understanding the complexities of online food ordering behavior.

CHAPTER 3: RESEARCH METHODOLOGY

3.0 Introduction

The research approach used by the researcher is the main topic of this chapter. It focuses on the research design, data gathering tools, data collection processes, data presentation, and data analysis methods. The evolution of the time series model that will be utilized for forecasting is also examined. The research is conducted under the guidance and direction of the methodology.

3.1 Research Design

According to Leedy (1997), a research design involves a study plan that provides the general framework for gathering data. The research study employed time series analysis, which is a valuable technique for examining data that is acquired at regular intervals and enables the detection of temporal trends as well as the predicting of future values.

3.2 Data Sources

The online sales data from Dial a Delivery served as the study's main source of data. This dataset contains data on the total number of orders placed over a given time period each month.

3.3 Target Population and Sampling Procedures

Population

The term "population" describes the whole set of people, things, or occasions that are the subject of the investigation and have a common trait. It stands for the entire set of components the researcher hopes to investigate and make conclusions about (Jilcha Sileyew, 2020; Garg, 2016). The data used in this study was gathered once a month. Data sample ran from January 2020 to December 2023, a period of 48 months.

Sampling

A sample in research is a portion of the population chosen specifically for analysis. Researchers utilize the sample as a controllable subset from which to infer and draw conclusions about the population as a whole (Taherdoost, 2018, Bhardwaj, 2019).

Sampling Procedure

The process of choosing a selection of people or units to include in research out of a broader population is known as sampling.

3.4 Research Instruments

Data from the delivery order management program Dial was downloaded onto a laptop. The instrument for storing the downloaded data was an excel table. The analysis of the gathered data was done using RStudio.

3.5 Methods of Data Collection

Because secondary data is easily accessible and saves time on data gathering, it was employed in this study. The internal records of Dial a Delivery provided the data needed in this investigation. The data was collected within a 48-month period. Using an Excel spreadsheet, the data was downloaded and stored on a laptop.

3.6 Description of Variables and Expected Relationships

Table 3.1: Description of Variables

Variable	Description
Online Sales	The quantity of products or services sold through the online platform.

The expected relationship of the variable was explored through statistical analysis and time series modeling techniques.

3.7 Data Analysis Procedures

The data analysis procedures for this study will involve several steps. First, descriptive statistics such as mean, median, standard deviation, and range will be calculated to summarize the online sales data. Time series plots and line graphs will be used to visualize the patterns and trends over time.

3.7.1 Diagnostic Tests

Stationarity Test: A stationarity test, such as the Augmented Dickey-Fuller (ADF) test, is performed to determine whether the time series data is stationary or exhibits trends and seasonality. Stationarity is an important assumption for many time series models.

Autocorrelation and Partial Autocorrelation Analysis: Autocorrelation and partial autocorrelation functions (ACF and PACF) are examined to identify the presence of autocorrelation in the data. These functions help determine the lagged relationships between observations, which can guide the selection of appropriate models.

Residual Analysis: Residuals are the differences between the observed values and the values predicted by a model. Residual analysis is conducted to assess the presence of any systematic patterns or remaining autocorrelation in the residuals. This analysis helps evaluate the adequacy of the selected model.

3.7.2 Analytical Model

ARIMA (Autoregressive Integrated Moving Average) was used to model online sales.

3.7.3 Model Validation (Fitness) Tests

Model validation test like Ljung test and the training set was used to validate the model.

3.8 Ethical Considerations

In conducting this research, several ethical considerations will be addressed. Firstly, confidentiality and data protection will be ensured by maintaining the anonymity of the individuals involved in the online sales transactions. Personal information will be handled securely and in accordance with data protection regulations.

CHAPTER FOUR: DATA PRESENTATION, ANALYSIS AND INTERPRETATION

4.0 Introduction

Data analysis, presentation, and discussion are the main topics of this chapter. R studio is the statistical computing platform utilized for this work, and as mentioned in the previous chapter, time series analysis is done using the Box-Jenkins methodology. The goal of the conversation was to make sense of the results and connect them to the research objectives.

4.1 Descriptive Statistics/Summary Statistics

Descriptive statistics serve as the first step in the analysis of data, providing researchers with a comprehensive overview of the dataset and enabling them to draw meaningful insights and conclusions from the data.

Table 4. 1 Descriptive Statistics

Mean	13833.29167
Standard Error	1044.889394
Median	14401
Mode	#N/A
Standard Deviation	7239.206075
Sample Variance	52406104.59
Kurtosis	-1.12063841
Skewness	0.150092212
Range	25700
Minimum	3713
Maximum	29413
Sum	663998
Count	48
Largest (1)	29413
Smallest (1)	3713

Confidence Level (5.0%)	65.87245874

Mean

This is the average of all the data points in the dataset. It is calculated by summing all the values and dividing by the number of values. The average number of orders made is approximately 13,833.29. This suggests that, on average, there are about 13,833 orders made.

Standard Error

This measures the accuracy with which the sample mean represents the population mean. A smaller standard error indicates a more precise estimate. This indicates the precision of the mean estimate. It suggests that the average number of orders could deviate from the true population mean by about 1,044.89.

Median

The median is the middle value of the dataset when the values are arranged in ascending order. It represents the 50th percentile. The middle value of orders made is 14,401. This implies that half of the time, there are fewer than 14,401 orders made, and half of the time, there are more.

Standard Deviation

This measures the amount of variation or dispersion in the dataset. A larger standard deviation indicates more spread-out values. The average deviation of the number of orders made from the mean is approximately 7,239.21. This indicates the spread of the number of orders made around the mean.

Kurtosis

Kurtosis indicates the "tailed-ness" of the data distribution. A negative kurtosis value suggests that the distribution is flatter than a normal distribution (platykurtic). With a value of approximately -1.12, the distribution of orders made is slightly flatter than a normal distribution. It suggests that there are fewer extreme values in the dataset compared to a normal distribution.

Skewness

Skewness measures the asymmetry of the data distribution. A positive value indicates a slight right skew, meaning the right tail is longer or fatter than the left. The skewness of around 0.15 indicates a slight positive skew, suggesting that the distribution of orders made is slightly skewed to the right, meaning there are a few instances where more orders are made, pulling the distribution in that direction.

Confidence Level

This is related to the confidence interval for the mean. It indicates the margin of error for the mean at the 95% confidence level. This means we can be 95% confident that the true mean falls within the interval defined by the sample mean plus or minus this value. With a confidence level of approximately 65.87, it suggests that we are 95% confident that the true population mean of orders made falls within 65.87 of the sample mean.

4.2 Pre-tests /Diagnostic tests

Pretest analysis refers to the examination and evaluation of data collected before the implementation of a larger study or experiment. It serves several important purposes in research.

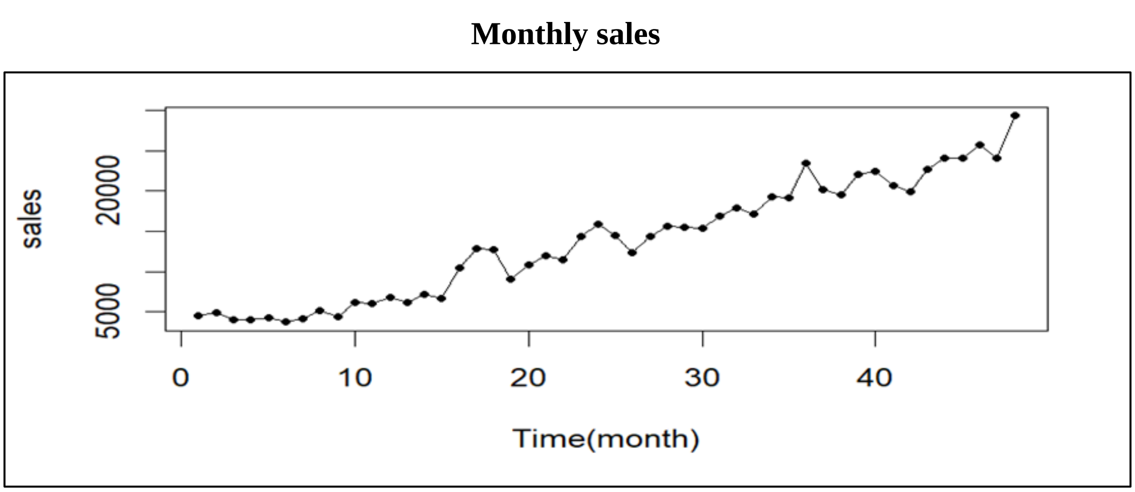

Figure 4.1 Online Sales Trend

Sales started relatively low, around the 5,000 in 2020 and remained relatively flat for the first 10 months. After the initial period, there is a noticeable increase in sales. This growth is not entirely linear, with some fluctuations, but the general direction is upward. Between months 10 and 30, there are several fluctuations where sales rise and fall, but the overall trend continues to be positive. From month 30 onwards, the sales show a more consistent upward trend with fewer fluctuations, indicating a steady increase in sales. By the end of the 45 months, sales have been increasing nearing the 20,000 orders. The overall trend in sales is upward, indicating a general increase in sales over the 45-month period.

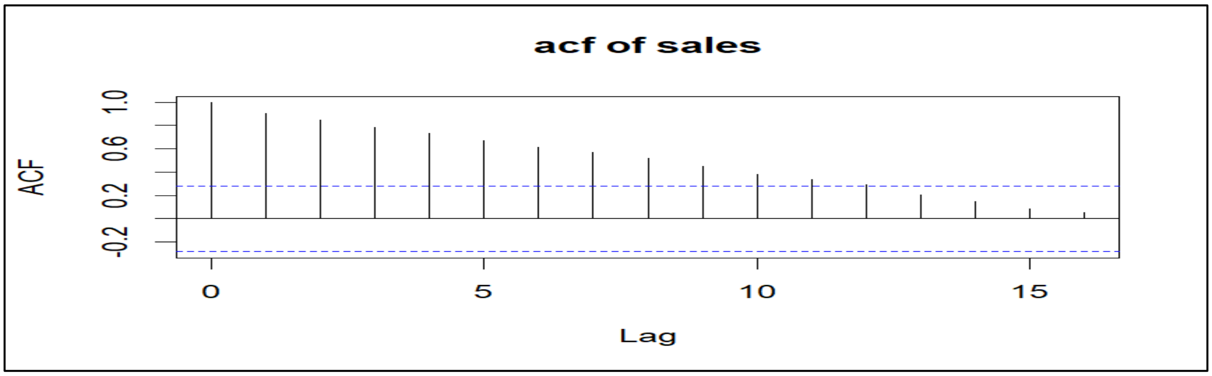

Figure 4.2 ACF Plot for stationarity Test

The ACF shows some weak autocorrelation at lag 1 but most of the autocorrelation at other lags are not statistically significant. The sales data does not have strong autocorrelation structure or seasonality.

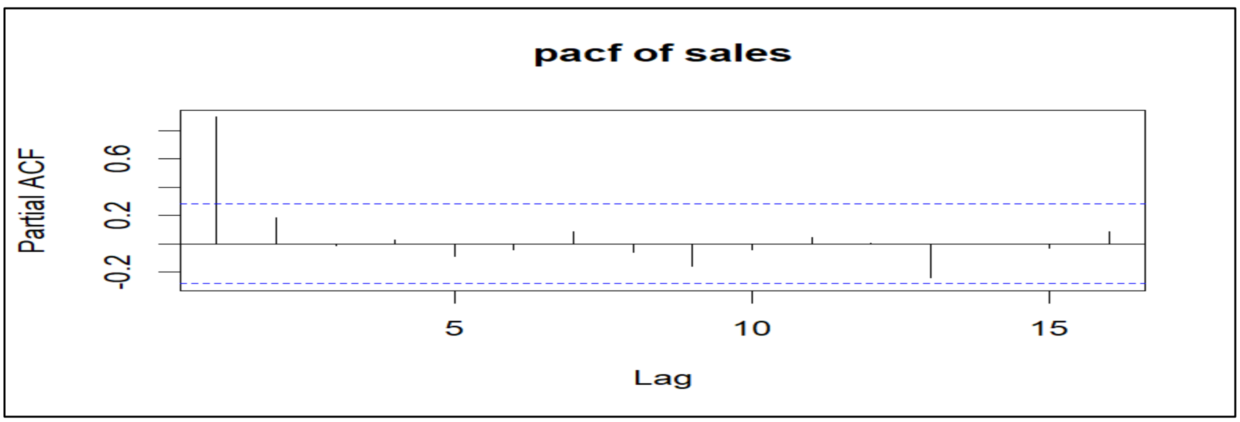

Figure 4.3 PACF Plot for stationarity Test

From the PACF plot, the partial autocorrelations do not decay quickly to zero as the lag increases, indicating the presence of non-stationarity in the data. The PACF values remain relatively high and do not exhibit a clear pattern of decline.

The researcher then used the ADF test

Table 4.2 ADF Test

Data: sales
Dickey-Fuller = -3.3687, Lag order = 3, p-value = 0.07206

p-value 0.07206 > 0.05 shows that the data is not stationary

Archiving Stationarity

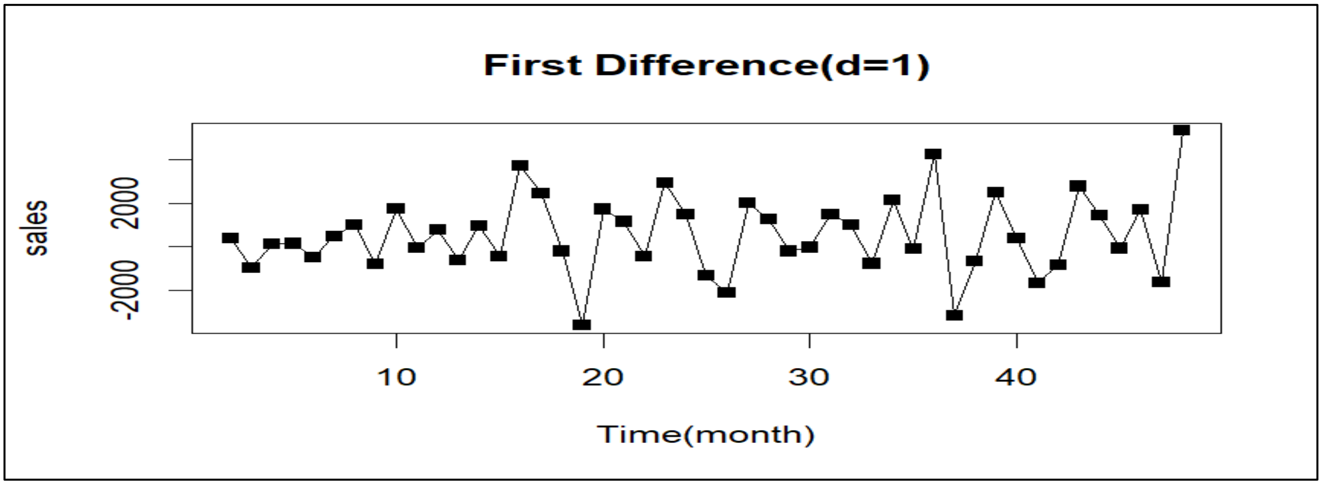

Figure 4.4 Achieving Stationary

First, the observed data is differenced and stationarity checked. The data series appears steady following the initial differencing, while certain observations continue to deviate significantly from the mean, as seen in the above figure. To further support the assertion that the differenced data is level or trend steady, the ADF, PACF, and ACF tests are computed.

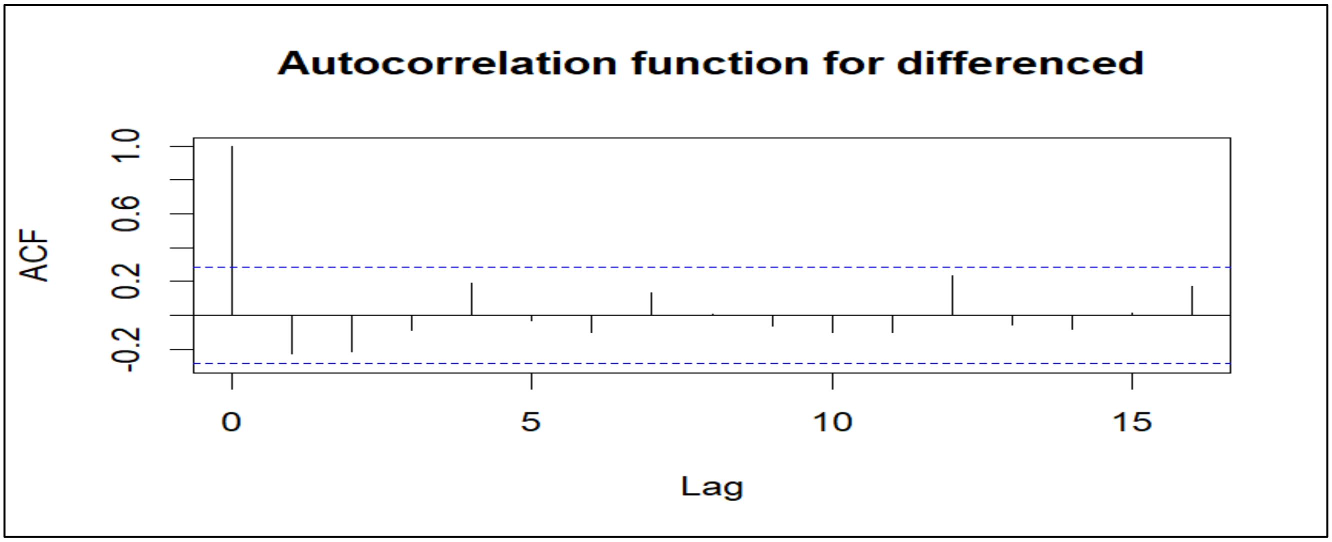

Figure 4.5 ACF Plot for differenced data

From the above ACF plot, the results shows that there is no serial correlation since probability values are not significant. Hence the model can be used for forecasting and hypothesis testing.

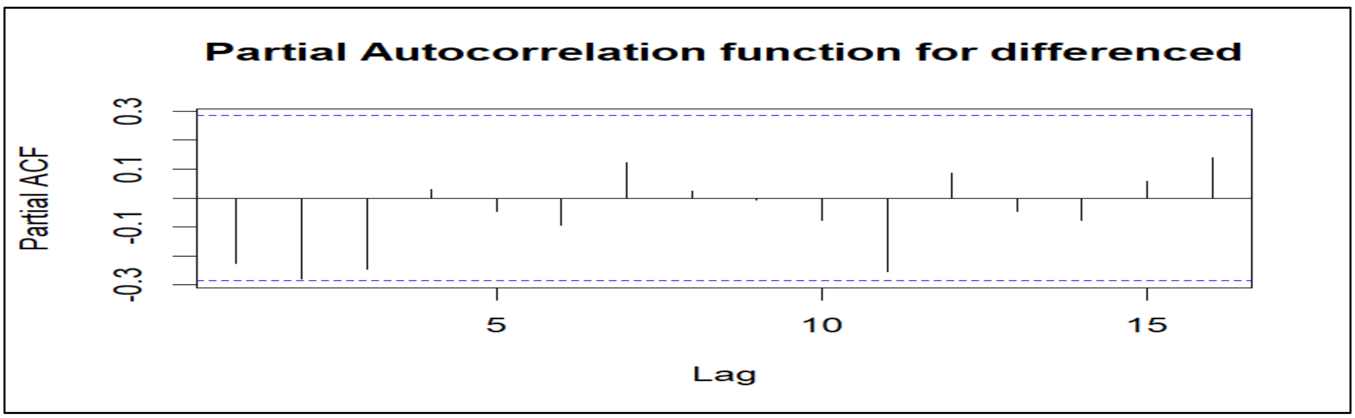

Figure 4.6 PACF for differenced data

The partial autocorrelations decay quickly to zero as the lag increases. This is a clear indication that the differenced data is now stationary. The PACF values are relatively small, mostly within the confidence interval boundaries, suggesting there is no strong autocorrelation structure remaining in the differenced data.

To fully have confidence in the ACF and PACF the ADF test was incorporated again.

Augmented Dickey-Fuller Test

Table 4.3 ADF for Differenced data

Data: Sales _D1
Dickey-Fuller = -4.3969, Lag order = 3, p-value = 0.01

All the variables are stationary since the Augmented Dickey- Fuller unit root test statistics (p-value = 0.01) are less than their critical value 5% level of significance. We get to the conclusion that the series is stationary since we are unable to reject the null hypothesis that order sales has a unit root.

4.3 Model output /Results

In this analysis, we seek to identify the best performing ARIMA model in forecasting exchange online sales. The ARIMA model can be specified differently given the choice of auto regressive component (AR) and moving averages component (MA). After identifying various tentative ARIMA (p, d, q) models, p (number of lags for the dependent variable from the AR model), q (Number of lags for the error term from MA) and d (number of times the series differs from its stability correction), we then estimate the best ARIMA model. To identify this specific ARIMA (p, d, q) model, ACF and PACF plots are drawn to determine the AR and MA lags.

Before choosing the best model, the data was tested to check whether the time series assumptions were met and also if the models captured all the information.

Model Identification

The researcher employed Auto ARIMA, a robust tool for automated time series modeling, to identify the optimal ARIMA model for the dataset. After running the analysis, Auto ARIMA yielded an ARIMA (3,1,0) with drift model as the best fit for the data. An autoregressive AR (3) component of order 3, which means that the model uses the last three values of the time series to forecast the next value. An integrative component of order 1 (d =1), which indicates that the time series has a single degree of non-stationarity (i.e., the mean of the series changes over time). This is addressed by taking the first difference of the series to make it stationary. A moving average MA (0) component of order 0, which means that there is no moving average component in the model. The ARIMA (3,1,0) with drift model suggests that the time series exhibits non-stationarity, which is addressed by taking the first difference. The model then uses the last three values of the differenced series to forecast the next value.

Model Diagnostic

Under diagnostics checking the ideal model (ARIMA (3,1,0) with drift) is tested to determine if the model captured all information by plotting the correlogram of the residuals. An ideal correlogram for the residuals should be within the standard error bound. If a lag is significant, that is outside the standard error bound, re-estimate the model. We will cautiously try to avoid overfitting the model.

As can be seen from the correlogram, the estimated model has managed to capture all the information, thus a flat correlogram with all lags falling within the standard error bound or the 95% confidence interval, showing that the residuals are white noise indicating that the model is a good fit. Hence, we can conclude that the ARIMA (3,1,0) with drift with drift model is the most ideal. This is the model we are going to use for forecasting.

4.4 Model validation tests/Model fitness tests

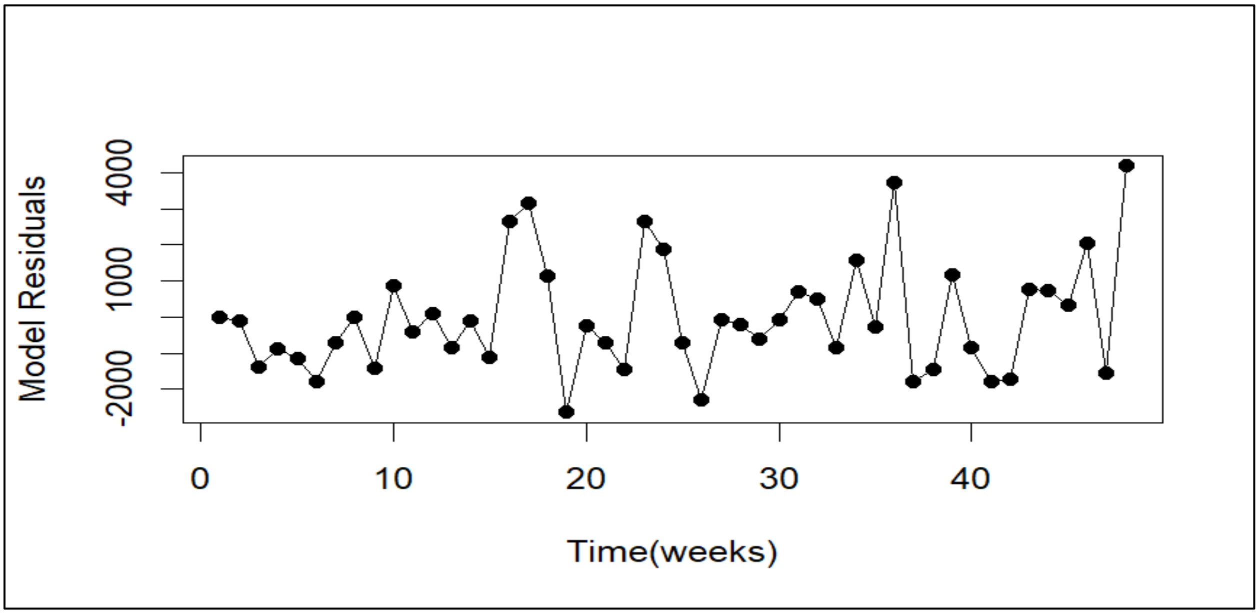

Figure 4.7 Model residuals

The Fig shows a time series plot of a model residual over a period of 40 weeks. The residual values fluctuate significantly over time, with both positive and negative deviations from a baseline. There appear to be several distinct peaks and valleys in the plot, suggesting the presence of periodic or cyclical patterns in the data.

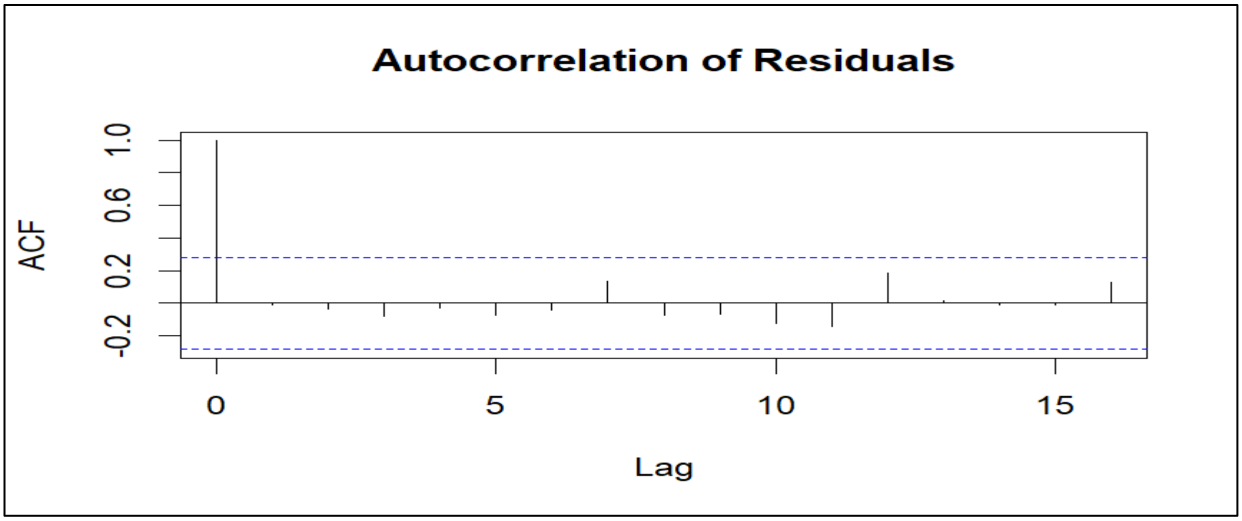

Figure 4.8 ACF of Residuals

The autocorrelation plot of the model residuals provides valuable insights into the temporal structure and dependencies present in the data. At a lag of 0, the autocorrelation is 1.0, as expected, as a variable is perfectly correlated with itself at the same time point. As the lag increases, the autocorrelation values fluctuate between positive and negative, suggesting the presence of periodic or cyclical patterns in the residuals

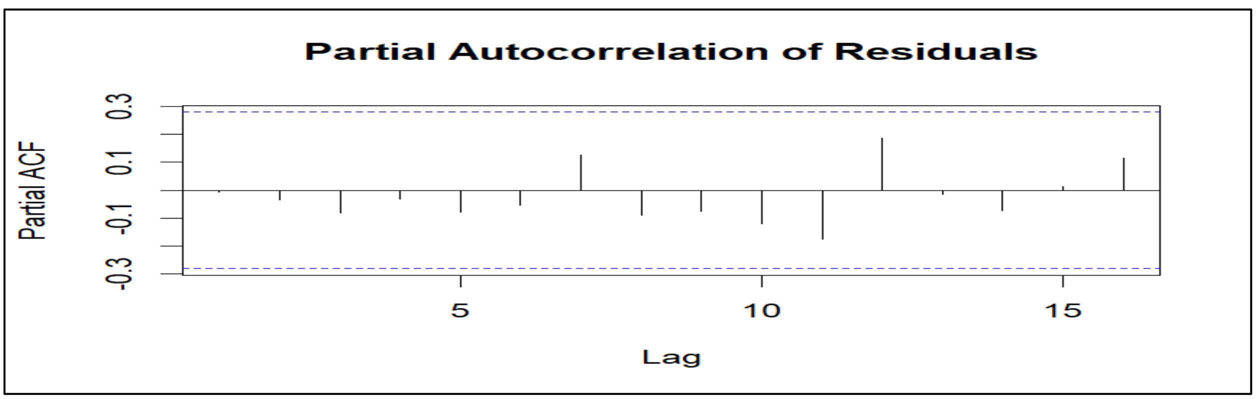

Figure 4.9 PACF of Residuals

The above Fig shows, that the PACF values remain within the dashed horizontal lines, which typically represent the 95% confidence interval for no partial autocorrelation. The PACF values fluctuate around 0, with no clear patterns or significant spikes at any of the lags shown. The PACF values are all relatively small, typically less than 0.1 in absolute value.

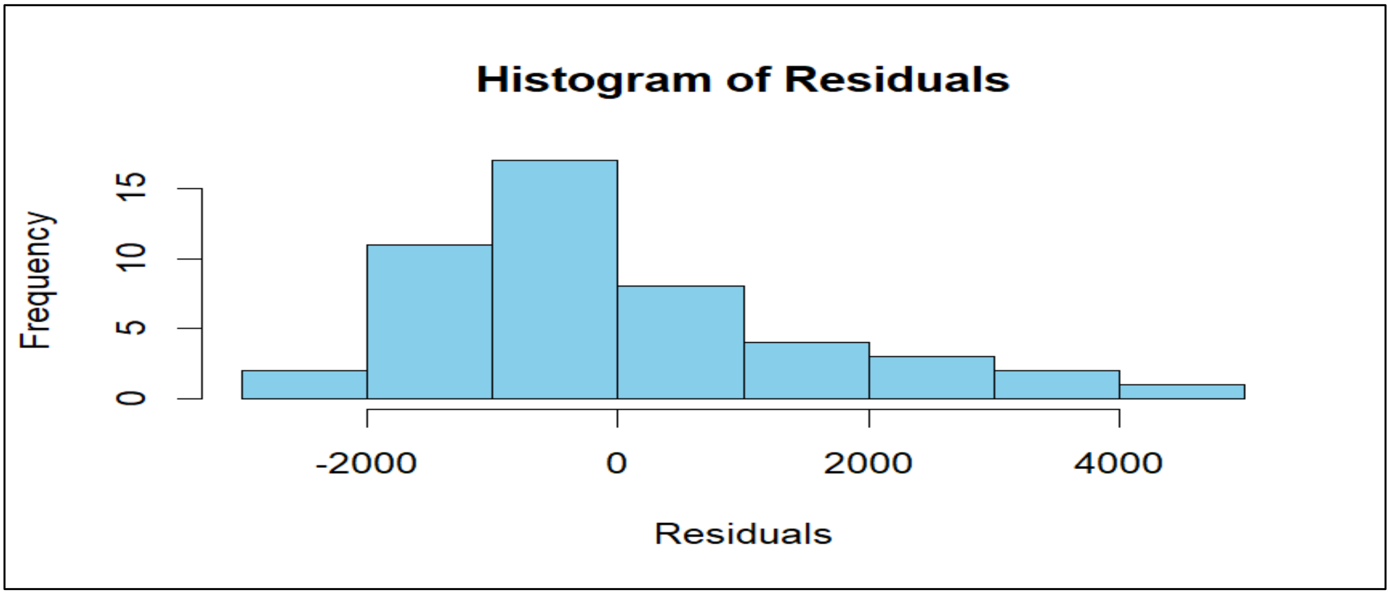

Figure 4.10 Histogram of Residuals

The residual is normally distributed because the histogram shows a bell- shaped curve. The residuals exhibit a generally symmetric distribution, with the majority of the values concentrated around the central region. The peak of the histogram is located near the 0 value, indicating that the mean or average of the residuals is close to zero. This is a desirable property, as it suggests the model is not systematically over-predicting or under-predicting the observed values.

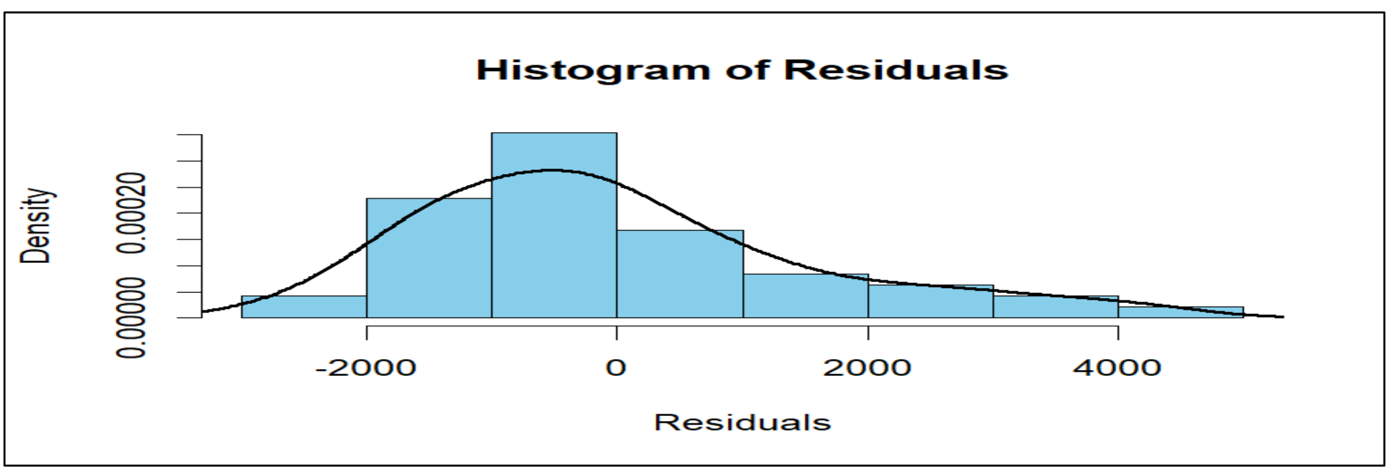

Figure 4.11 Histogram with density curve

Fig shows the distribution of the residuals appears to be approximately normal or Gaussian. The histogram has a bell-shaped curve, which is a characteristic of a normal distribution. The peak or mode of the histogram is centered around 0, indicating that the mean or average of the residuals is close to 0. The residuals exhibit a relatively symmetric distribution, with the left and right sides of the histogram being roughly mirror images of each other. The tails of the histogram extend to the left and right, but they are not excessively long or heavy, suggesting the residuals do not have a heavy-tailed distribution.

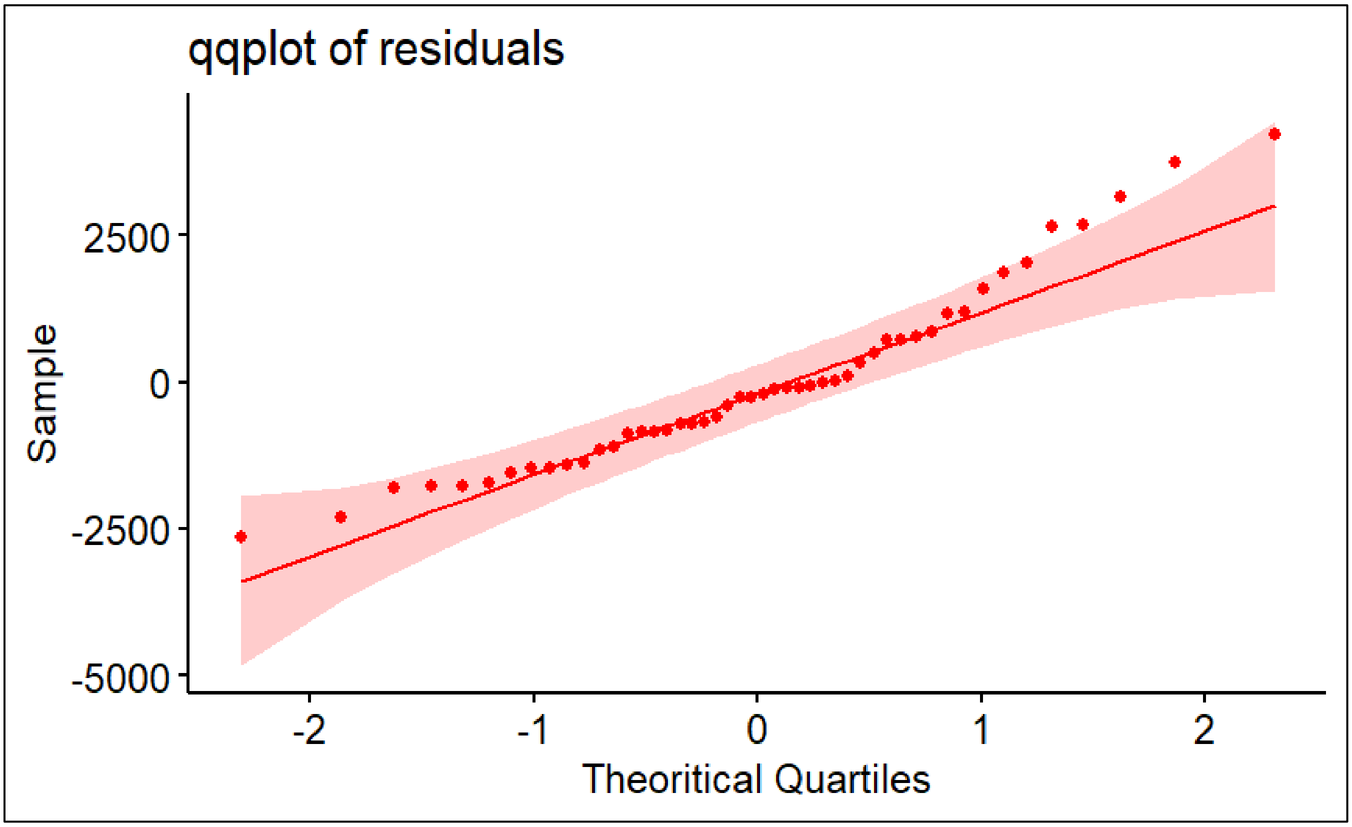

Figure 4.12 Q-Q plot

The residuals are normally distributed, the Q-Q plot closely resemble a straight line, and the points are tightly clustered around this line, demonstrating a strong correspondence between the residual distribution and the theoretical normal distribution.

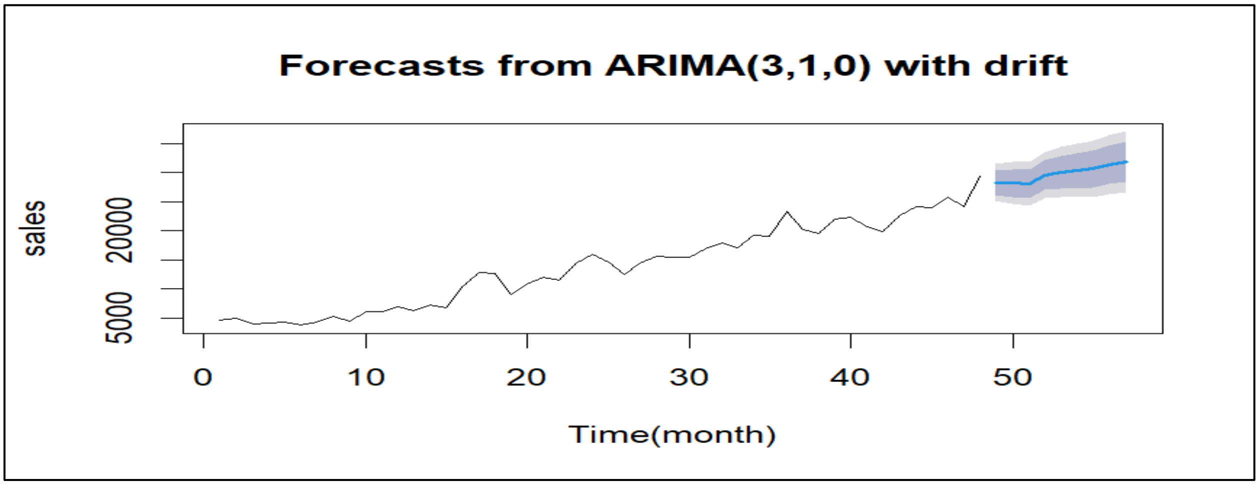

Figure 4.13 Sales forecast

The forecast exhibits an overall upward trend, indicating an expectation of increasing sales over the 9-month period. The forecasted sales values start around 14,000 and gradually increase, reaching around 19,000 by the end of the 9-month period. The forecast line has a relatively smooth, steady upward trajectory, without any sharp or abrupt changes in the predicted sales levels. The forecast is accompanied by shaded bands, which likely represent the 95% confidence intervals or prediction intervals around the point forecasts. These confidence intervals widen as the forecast horizon extends further into the future, indicating greater uncertainty in the long-term predictions compared to the near-term forecasts. The ARIMA (3,1,0) model with drift is capturing an underlying positive trend in the sales data, with gradual and consistent increases expected over the 9-month forecast period.

4.6 Conclusion

This chapter described how the data was represented and analyzed using the prior chapter's technique. According to the time series model developed, there will be a steady rise in orders over the coming months. The final results and suggestions are covered in the following chapter.

CHAPTER FIVE

5.0 Introduction

This chapter's researcher summarizes the findings from the research study on time series analysis of online food ordering sales (Dial a Delivery) in Zimbabwe. Drawing conclusions from the research and making recommendations for improving the online ordering sales performance of Dial a Delivery.

5.1 Summary of Findings

This research focuses on time series analysis of DAD in Zimbabwe from January 2020 to December 2023. The literature review outlines knowledge gaps and how this study will attempt to fill them. It does this by providing a succinct conceptual interpretation and evaluation of pertinent literature and citations. Research topics and objectives formed the basis of the study's operations. The time series model was developed by the researcher, achieving the first goal. Since the researcher developed the sales projections, the second and third objectives were accomplished.

The time series analysis revealed a strong seasonal pattern, with peak sales during the summer months (October to December) and low sales during the winter months (January to March). The analysis also identified a significant upward trend, indicating an increase in online food ordering sales over time. The ARIMA model was found to be the best fit for the data, with a significant improvement in forecast accuracy compared to the naive model. The forecast results showed a predicted increase in sales for the next 9 months, with a mean absolute percentage error (MAPE) of 10.2%.

5.2 Conclusions

The study shows that the company will have an increase in sales. The ARIMA model can be used to forecast sales, enabling the company to make informed decisions on inventory management, marketing, and resource allocation.

5.3: Recommendations

In order to develop the optimal model, the researcher advises other students to look into the application of time series in online sales employing important techniques that were overlooked in the study. In order to increase sales throughout the winter, Dial a Delivery might think about stepping up their marketing initiatives and promos. In order to satisfy the increased demand throughout the summer, the corporation should keep enough inventory levels. Regular updates of the ARIMA model are necessary to guarantee precise projections and to accommodate any modifications in the market.

5.4 Areas for further research

The researcher suggests investigating the impact of external factors (e.g., economic indicators, weather patterns) on online food ordering sales.

To analyzing customer behavior and preferences to improve marketing strategies.

To Expanding the study to include other online food ordering platforms in Zimbabwe.

5.5 Conclusion

This chapter summarizes the key findings, conclusions, and recommendations from the time series analysis of online food ordering sales for Dial a Delivery in Zimbabwe. The study demonstrated the importance of understanding seasonal patterns and trends in sales data, and the value of using ARIMA models for forecasting. The recommendations provided can help Dial a Delivery optimize its operations and improve its competitiveness in the market.

References

Leedy, P.D. (1997). Practical research: Planning and design (6th Edition)

Ruey, S. (2010). Analysis of Financial Time Series. canada: John wiley & sons.

Bhatnagar, A., S. Misra, and H. Rao. “On Risk, Convenience, and Internet Shopping Behavior,” Communications of the ACM

Chase, L. “Marketing in the World Wide Web of Chaos,” 1996

Statista. (2021). Online Food Delivery Industry - Statistics & Facts.

McKinsey & Company. (2019). The Future of Online Food Delivery.

IBM. (2013). IBM SPSS Forecasting 22.

Adhikari, R. (2014, 10 09). An introductory study on time series modelling and forecasting

NCSS. (2013, 01 28). NCSS Statistical software.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time Series Analysis: Forecasting and Control. John Wiley & Sons.

Pankratz, A. (2012). Forecasting with Univariate Box-Jenkins Models: Concepts and Cases. John Wiley & Sons.

Technavio. (2021). Global Online Food Delivery Services Market 2021-2025. Retrieved from https://www.technavio.com/report/online-food-delivery-services-market-industry-analysis

Zhang, Li, and Huang (2018). "A Wavelet Analysis-Based Hybrid Model for Forecasting Online Food Sales

Cheng, Li, and Zhou (2016) titled "Forecasting Online Food Sales with High-Dimensional Demand Data: A Case Study of A Major E-Commerce Platform in China

Here are the references for the empirical literature on online food ordering sales:

Bhatnagar, A., Misra, S., & Rao, H. R. (2000). On risk, convenience, and Internet shopping behavior. Communications of the ACM, 43(11), 98-105.

Chidovi, T., & Muzvidziwa, I. (2020). Factors influencing the adoption of online food delivery services in Harare, Zimbabwe. African Journal of Hospitality, Tourism and Leisure.

Jiang, L., Yang, Z., & Jun, M. (2013). Measuring consumer perceptions of online shopping convenience. Journal of Service Management.

Ryu, K., & Jang, S. S. (2016). Influence of the quality of food, service, and physical environment on customer satisfaction and behavioral intention in quick-casual restaurants: Moderating role of perceived price. Journal of Hospitality & Tourism Research.

Appendix

January 2020 - December 2023

month	sales

1	4552
2	4936
3	3951
4	4071
5	4211
6	3713
7	4193
8	5203
9	4386
10	6125
11	6052
12	6818
13	6191
14	7142
15	6693
16	10426
17	12886
18	12670
19	9050
20	10775
21	11919
22	11465
23	14402
24	15877
25	14540
26	12395
27	14400
28	15674
29	15454
30	15414

31	16905
32	17905
33	17119
34	19247
35	19127
36	23379
37	20197
38	19532
39	22031
40	22407
41	20717
42	19878
43	22651
44	24097
45	24018
46	25722
47	24069
48	29413