BINDURA UNIVERSITY OF SCIENCE EDUCATION

DEPARTMENT OF STATISTICS AND MATHEMATICS

FACULTY OF SCIENCE AND ENGINEERING

TIME SERIES ANALYSIS OF TURNOVERS IN QUICK SERVICE RESTAURANT INDUSTRY: A CASE OF SIMBISA BRANDS ZIMBABWE

BY

HOLANDI BRANDON R

B190613B

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE BACHELOR OF SCIENCE HONOURS DEGREE IN STATISTICS AND FINANCIAL MATHEMATICS

SUPERVISOR: DR. M MAGODORA

JUNE 2024

APPROVAL FORM

The undersigned certified that they have supervised and recommended to Bindura University of Science Education for acceptance of dissertation entitled 'Time series analysis of turnovers in quick service restaurant industry. A case of Simbisa brands' submitted in partial fulfilment of a Bachelor of Science Honors Degree in Statistics and Financial Mathematics.

HOLANDI BRANDON R	………………………………	Date
B190613B	Signature	10/06/2024

Certified by:

Dr M Magodora	………………………	Date
Supervisor	Signature	10/06/2024

Dr M Magodora	……………………......	Date
Chairperson	Signature	10/06/2024

DECLARATION

I Holandi Brandon R hereby declare that this submission is entirely my own creation, with due acknowledgment of any sources referenced. Furthermore, I confirm that this work has not been previously submitted for any academic qualification, either at this institution or elsewhere.

Author:

Holandi Brandon R

Registration Number: B190613B

Signature: …………

Date: 10 June 202

DEDICATION

I dedicate this research project to the memory of my beloved mother, whose enduring love, guidance, and sacrifices continue to inspire me every day.

ACKNOWLEDGEMENTS

I would like to extend my deepest appreciation to my academic mentor, Dr. Magodora, whose steadfast support and dedication significantly contributed to the successful completion of this project. Additionally, I am grateful to Ms. P. Hlupo and the entire team at the Department of Statistics and Financial Mathematics for their continuous encouragement and unwavering moral backing throughout this endeavor.

I am also indebted to my family members, namely Mr. and Mrs. Karodza, my mother, and my sister, Shaolin Kambani, for their invaluable support. Moreover, I express my gratitude to my friends, Monalisa Chiketah, Alex Manjoro and Byron Mhako, for their encouragement and unwavering support. Special gratitude is reserved for my family, whose enduring patience, financial assistance and technical expertise have been invaluable throughout my life.

ABSTRACT

This research delves into the turnover dynamics of Simbisa Brands within the Quick Service Restaurant (QSR) industry in Zimbabwe, spanning from January 2016 to December 2022. Adopting a quantitative and analytical approach, the study aims to uncover patterns, trends, and seasonality in historical turnover data, essential for strategic decision-making within the QSR sector. The research design, informed by methodologies outlined by Anderson et al., (2016), emphasizes the significance of quantitative analyses in guiding strategic processes. Primary data sourced directly from Simbisa Brands' financial records ensures the reliability and accuracy of the dataset, comprising 84 monthly observations. Utilizing statistical analysis software like Python and Microsoft Excel, researchers conduct sophisticated time series analysis and diagnostic tests, ensuring a robust analytical framework. Turnover, as the primary variable, serves as a crucial metric for assessing business growth, profitability, and market competitiveness. Diagnostic tests including the Augmented Dickey-Fuller test and autocorrelation tests validate data integrity and guide modeling decisions. The Autoregressive Integrated Moving Average (ARIMA) model is employed to capture turnover dynamics, while machine learning models like XGBoost and Random Forests are introduced and evaluated for forecasting accuracy. Ethical considerations such as data privacy and obtaining informed consent are carefully addressed. Through systematic examination, this research provides actionable insights empowering businesses to optimize resource allocation, refine marketing strategies, and enhance operational efficiency within the QSR industry. It contributes to the field of business analytics while upholding integrity, trustworthiness, and respect for all stakeholders.

Table of Contents

Contents

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

QSR	**QUICK SERVICE RESTAURANT**
MAE	**MEAN ABSOLUTE ERROR**
MSE	**MEAN SQUARED ERROR**
RMSE	**ROOT MEAN SQUARED ERROR**
MAPE	**MEAN ABSOLUTE PERCENTAGE ERROR**
ARIMA	**AUTORESSIVE INTEGRATED MOVING AVERAGE**
AR	**AUTOREGRESSIVE**
I	**DIFFERENCING**
MA	**MOVING AVERAGE**
ARMA	**AUTOREGRESSIVE MOVING AVERAGE**
ACF	**AUTOCORRELATION FUNCTION**
PACF	**PARTIAL AUTOCORRELATION FUNCTION**
ADF	**AUGMENTED DICKEY-FULLER**

CHAPTER 1

INTRODUCTION

1.0 Introduction

This study focuses on the turnover dynamics in the Quick Service Restaurant (QSR) industry, specifically in Zimbabwe, and aims to understand historical patterns and predict future trends. It uses historical time series data and forecasting models to analyze turnover patterns and predict future trends. This chapter focuses on the introduction, background of the study, statement of the problem, objectives, research question, scope of the study, significance of the study, assumptions, limitations and definitions of terms. Chapter 2 conducts an extensive literature review, situating the study within academic discourses on turnover, time series analysis, and related case studies. Chapter 3 outlines the research framework, detailing data collection and analysis procedures. Chapter 4 presents findings, analyzing historical turnover trends using time series models and identifying key determinants influencing turnover fluctuations. Finally, Chapter 5 synthesizes findings to offer practical recommendations for optimizing Simbisa Brands' operations and enhancing turnover performance, while also pointing out areas for future research and contributing to knowledge advancement in the fast food chain sector. . It also discusses implications for industry theories and practices, concluding with the study's contribution to advancing knowledge in the QSR sector.

1.1 Background of the study

The global market for fast food and quick service restaurants reached a value of USD 257.19 billion in 2019 and is anticipated to grow at a compound annual growth rate (CAGR) of 5.1% from 2020 to 2027. This growth is driven by a rising preference for fast food among consumers from Generation X, Y, and Z worldwide. The fast food industry, valued at over USD 500 billion in 2019, continues to expand as quick service restaurants emerge as significant providers of convenient, flavorful, and cost-effective mass-produced food, attracting a growing clientele seeking enjoyable dining experiences. Key factors such as convenience, taste, and affordability in terms of both time and money are pivotal drivers supporting the fast food and quick service restaurant market by Grand View Research (2019) .

The Quick Service Restaurant (QSR) industry, recognized for its rapid service and convenience, has become an integral part of modern urban life, especially in rapidly developing countries like Zimbabwe. The QSR sector in Zimbabwe has experienced significant growth, propelled by urbanization, changing consumer lifestyles, and increasing disposable incomes. Within this dynamic landscape, Simbisa Brands stands out as a leading player, operating numerous well-known restaurant chains such as Chicken Inn, Pizza Inn, Bakers Inn and etc. Despite the sector's overall expansion, Simbisa Brands faces challenges related to economic volatility, competitive pressures, and evolving consumer preferences. Understanding turnover patterns the total sales or revenue generated is critical for Simbisa Brands. These patterns provide insight into the business's financial health and operational efficiency. Analyzing turnover trends allows the company to make informed business decisions, optimize operations, and plan for future growth. By examining the turnover trends of Simbisa Brands using ARIMA model to forecast future turnovers. ARIMA model can help identify patterns such as seasonal peaks during holiday periods and long-term growth trends driven by urbanization and rising incomes, the study aims to provide valuable insights that can help Simbisa Brands navigate Zimbabwe's dynamic QSR landscape more effectively. Forecasting turnovers in the QSR industry involves predicting future sales based on historical data. Accurate turnover forecasts are essential for QSR businesses to manage inventory, plan marketing campaigns, allocate resources, and make strategic decisions Hyndman & Athanasopoulos, (2018).

1.2 Statement of the problem

Simbisa Brands, a prominent entity in Zimbabwe's Quick Service Restaurant (QSR) industry, is confronted with the challenge of understanding and managing fluctuations in sales turnovers. The absence of a comprehensive analysis of sales turnover dynamics hinders the company's ability to make informed decisions regarding resource allocation and strategic planning. As a consequence, Simbisa Brands faces uncertainty in forecasting revenue, potentially impacting financial stability and growth prospects within the competitive QSR market of Zimbabwe.

1.3 Objectives

1. To analyze historical time series data of turnover in Simbisa Brands.

2. To forecast future turnover trends using time series models.

1.4 Research Questions

i. How does turnover dynamics within Simbisa Brands exhibit temporal trends and seasonal variations, as revealed by time series analysis?
ii. Can time series forecasting models effectively predict future turnover dynamics for Simbisa Brands?

1.5 Scope of the Study

This study is primarily focused on examining the trends of gross turnovers within the Quick Service Restaurant (QSR) industry specifically in Zimbabwe. The analysis will concentrate on turnover dynamics over a specified time period, within the QSR sector. Simbisa Brands will be highlighted as a case study, providing a focused lens through which to explore turnover trends and performance drivers.

1.6 Significance of the Study

The findings of this study hold significant implications for various stakeholders in the QSR industry. By uncovering the trends and predictions of turnover dynamics, the study can offer strategic insights for companies like Simbisa Brands to optimize their operations, enhance performance, and informing policy decisions aimed at supporting industry growth. Moreover, transparent and data-driven analysis of turnover trends can bolster investor confidence, facilitating investment in the Zimbabwean QSR market. Lastly, the study's academic contribution lies in its potential to advance knowledge in turnover analysis and strategic management within the QSR industry.

1.7 Assumptions

1. The variance of the residuals is assumed to be constant across all observations
2. The ARIMA model assumes that there are no outliers or influential data points in the time series data that could unduly influence the model estimation process
3. The time series data should exhibit stationarity, meaning that the statistical properties such as mean and variance remain constant over time.

1.8 Limitations

1. Limited Historical Data: Short-term data may not fully capture long-term trends or industry dynamics.
2. Assumption of Stationarity: Real-world data may not meet the stationary assumption, impacting model accuracy.

1.9 Definitions of terms

Turnover: Sales turnover is the total number of items or services sold by the firm during a specific time period. Campbell, (2015). It represents the aggregate sales generated from all transactions, including food and beverage purchases, promotional items, and any additional revenue streams.

Time series analysis: a statistical method for analyzing sequential data points over time, has its origins dating back to ancient civilizations, where scholars like Hipparchus and Ptolemy studied astronomical data to discern patterns in celestial movements (Box et al., 2015)

Quick Service Restaurant: provide a limited menu, fast service, and minimal table service by Paul R. Dittmer and J. Desmond Keefe III (2017).

1.10 Chapter Summary

Chapter 1 introduces the study's focus on turnover patterns within Zimbabwe's Quick Service Restaurant (QSR) industry, using Simbisa Brands as a case study. It discusses the research's background, stressing the need to comprehend turnover dynamics for effective decision-making in a competitive market. The problem statement identifies research gaps and highlights the study's aim to address them through advanced time series analysis. Objectives, research questions, scope, significance, assumptions, limitations, and term definitions are also presented. The chapter closes with a summary, preparing the groundwork for ensuing chapters.

CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

This chapter will comprehensively cover three essential components: the conceptual framework, theoretical framework, and empirical review. These components serve as the foundation for understanding and analyzing turnover dynamics within the Quick Service Restaurant (QSR) industry. The conceptual framework establishes fundamental concepts and principles relevant to turnover analysis, while the theoretical framework delves into existing theories and models informing our understanding of turnover phenomena. Additionally, the empirical review synthesizes previous research findings to identify trends, patterns, and gaps in the literature, providing a basis for our own empirical investigation. Together, these components contribute to a holistic framework for studying turnover in QSR establishments, integrating conceptual, theoretical, and empirical insights to guide our research approach and interpretation of findings.

2.1 THEORETICAL FRAMEWORK

2.1.1 Time Series Analysis

Time series analysis, a statistical method for analyzing sequential data points over time, has its origins dating back to ancient civilizations, where scholars like Hipparchus and Ptolemy studied astronomical data to discern patterns in celestial movements Box et al.,(2015). However, the formal development of time series analysis began in the 20th century, notably with the work of George Udny Yule and Norbert Wiener Cryer & Chan, (2008). Yule introduced autoregressive and moving average models, while Wiener advanced the theory of stochastic processes, both laying the groundwork for modern time series analysis techniques.

2.1.2 Components of time series

Time series analysis involves several critical components that are foundational to understanding and modeling time-dependent data. The key components include trend, seasonality, and noise.

Trend refers to the long-term progression of the data over time, reflecting underlying patterns or directions that the data follows. Trends can be upward, downward, or flat and are often influenced by factors such as economic conditions, technological advancements, or demographic changes. Trends are important in identifying persistent movements in data that can help in making long-term forecasts. For instance, a consistent upward trend in stock prices might indicate long-term growth in a company's value Box et al., (2015).

Seasonality captures regular, periodic fluctuations within the time series that occur at fixed intervals, such as daily, monthly, or yearly cycles. This component is crucial for understanding patterns that repeat over specific periods, such as increased retail sales during holiday seasons or higher electricity consumption during summer months. Seasonality is driven by calendar effects, weather changes, holidays, and other periodic events. Identifying seasonality helps in making more accurate short-term forecasts and understanding the cyclical nature of the data Box et al., (2015).

Noise, or the random component, represents the irregular, unpredictable variations in the data that cannot be attributed to the trend or seasonality. Noise is often the result of random external influences or measurement errors and is typically modeled as a stochastic process. While noise is generally considered as the unpredictable part of the data, understanding its properties is crucial for building robust models that can handle the inherent randomness in time series data by Box et al., (2015).

A comprehensive time series model aims to decompose the data into these components to better understand and forecast future values. The decomposition can be additive, where the components are summed, or multiplicative, where they are multiplied, depending on the nature of the data. Additive decomposition is used when the components (trend, seasonality, and noise) are assumed to contribute linearly, while multiplicative decomposition is used when these components interact in a non-linear manner.

2.1.3 ARIMA Modeling

The theoretical development of ARIMA model traces back to the pioneering work of Box and Jenkins in the 1970s. This methodology revolutionized time series analysis by providing a systematic framework for modeling and forecasting. ARIMA models combine autoregressive (AR), differencing (I), and moving average (MA) components to capture the linear dependencies and trends present in time series data.

Autoregressive (AR) Model

The George, Gwilym, & Gregory definition of an Autoregressive model is a system of random processes that monitors series over time. Such models form the basis for various other models used in time series analysis as they assume that future values of series can be predicted from past values that have been recorded. This model uses previous inputs and a stochastic term to come up with its stochastic differential equation; it is denoted by AR(p), where 'p' represents the number of previous values considered. Thus, in essence, Autoregressive models use steps made in the past to predict the future ones. However, a disadvantage with this type of model is that it may suffer effects from constant component which are usually transitory or isolated shocks. To address this issue, lag value is often included in order to determine how far back in time affects are felt on current state and consequently influence future ones. Further, non-stationarity in AR model can be handled through introduction of unit root variable.

Moving Averages (MA) Model

According to Pannerselvam's presentation the Moving Averages model entails seasonal components characterized by their randomness nature. Unlike non-stationary Autoregressive model, MA model is stationary by itself involving linear regression between current value and random shocks or white noise or stochastic shocks among others as opposed to AR, which includes an autoregressive that is related to non-shock values. MA model forecasts error lags in the series.

Autoregressive Moving Average

ARMA model implies AR (p) and MA (q), where p = number of significant terms in the ACF and q = number of significant terms in the PACF

Autoregressive Integrated Moving Average

According to Yu, G. and Zhang, C (2004) ARIMA models are mathematically constructed as ARIMA (p,d,q) where p and q are similar to ARMA model but d = number of initial differences. The first step for implementing an ARIMA model involves examining whether or not the time series is stationary. When data has a constant pattern over time, then it forms perfect examples for use in developing an ARIMA model implying that data should be stable over long periods of time. Thus, if data has a trend either upward or downward with distinctive patterns referred to as seasonality; then it's proper to conclude that the data is not stationary on the other hand when there is no such pattern, we have certainty about stationarity.

2.1.4 Forecasting

Forecasting refers to the process of anticipating the future trajectory of a specific variable using logical methods and existing data Hyndman & Athanasopoulos, (2018). It entails making predictions or estimations about what lies ahead based on historical and present information, offering insights into potential future occurrences. Forecasting serves as a tool for organizations to anticipate forthcoming revenues. While numerous forecasting methodologies exist, each organization selects a technique tailored to its unique context and requirements. Additionally XGBoost and Random Forest are two powerful machine learning models that can be adapted for forecasting time series data, such as turnovers. Both methods require transforming the time series problem into a supervised learning task by creating lagged features and incorporating rolling statistics to capture trends and seasonality. XGBoost, known for its efficiency and high performance, constructs an ensemble of decision trees sequentially, with each tree correcting the errors of the previous ones, making it adept at handling complex patterns in data. It also includes regularization to prevent overfitting and can handle missing values effectively. On the other hand, Random Forest builds multiple decision trees independently and averages their predictions, providing robustness and reducing variance. It is easier to interpret and implement but might be slower with large datasets compared to XGBoost. Both models need careful feature engineering and proper train-test splitting to respect the temporal order of the data. Ultimately, the choice between XGBoost and Random Forest depends on the specific characteristics of the time series data and the desired balance between model complexity, interpretability, and computational efficiency.

2.2 Empirical evidence

2.2.1 ARIMA Forecasting of Financial Data

Forecasting of financial data has been extensively explored in the literature, with several studies employing time series analysis techniques to predict economic and financial indicators. The performance of ARIMA models is compared with LSTM models, revealing the superiority of LSTM in forecasting accuracy by S. Siami-Namini and A.S. Namin (2018). Similarly, the research of A.A. Ariyo, A.O. Adewumi, and C.K. Ayo (2014) investigates the effectiveness of ARIMA models in predicting stock prices, demonstrating their strong potential for short-term prediction. However, the research by A.C. Petrică, S. Stancu, and A. Tindeche (2016) identifies challenges associated with using ARIMA models for financial forecasting, emphasizing the need for alternative approaches. Furthermore, research extends beyond financial markets to domains like tourism demand forecasting, as seen by J. Li and X. Zhang (2019), which showcases the versatility of ARIMA models in predicting diverse economic indicators. These studies collectively contribute valuable insights into the application of time series analysis methods in forecasting financial data across various sectors and contexts, informing decision-making processes in economics, business, and finance.

In a study of Devi, Sundar and Alli, Alli et al (2014) introduces a robust methodology for predicting stock trends using ARIMA models tailored to the Nifty Midcap-50 index. Their research underscores the practical value of ARIMA models in decoding stock market trends, offering actionable insights for investors and financial analysts. Similarly, the study by Goswami et al (2014) delves into the performance of ARIMA models in predicting stock prices across various

Indian stocks. Their comparative analysis reveals ARIMA's competitive edge over alternative forecasting techniques, particularly in the context of the Indian stock market dynamics. Furthermore, the study by Ismail et al (2015) introduces an innovative approach to forecasting financial time series data by fusing wavelet transforms with ARIMA models. This hybrid model showcases promising potential in enhancing forecasting precision and comprehensively capturing the intricate patterns within financial datasets. Collectively, these studies underscore the versatility and effectiveness of ARIMA models in forecasting financial trends across different market contexts.

In a study conducted by Okafor and Shaibu (2013) offers an insightful examination of how ARIMA models can be effectively employed to analyze inflation dynamics in Nigeria. Their study delves into the nuances of inflation trends within the Nigerian economy, showcasing the applicability of ARIMA models in capturing and understanding these complex dynamics. By utilizing ARIMA models, the researchers provide valuable insights that can inform economic policy-making and contribute to maintaining financial stability in Nigeria. Similarly, a study Sivasamy et al (2014) delves into the realm of stock market forecasting using ARIMA models. Through their investigation, they illuminate the efficacy of ARIMA models in predicting stock market trends and behaviors. By leveraging ARIMA models, the researchers offer valuable insights into the dynamics of stock markets, aiding investors and financial analysts in making informed decisions. Also a study by Song (2020) explores an innovative approach to financial time series forecasting. Their study combines ARIMA models with deep learning techniques to enhance forecasting accuracy for high-frequency financial data. This hybrid methodology presents promising results, demonstrating the potential to improve forecasting reliability in financial markets. Overall, these studies underscore the versatility and effectiveness of ARIMA models in various domains of financial analysis and forecasting, offering valuable contributions to both academia and industry.

2.3 Research gap

The gap in the literature lies in the lack of studies specifically focusing on forecasting financial data for companies like Simbisa Brands within the Quick Service Restaurant (QSR) industry. While existing research provides valuable insights into forecasting financial data using ARIMA models in various contexts, there is a notable absence of studies that delve into the unique dynamics and challenges faced by companies operating in the QSR industry, such as Simbisa Brands. This gap represents an opportunity for further exploration to understand the specific factors influencing turnover dynamics and business performance within the QSR sector. By addressing this gap, researchers can provide tailored insights and forecasting models that cater to the distinct characteristics of companies like Simbisa Brands, thereby enhancing the accuracy and applicability of forecasting techniques in the QSR industry.

2.4 Conceptual framework

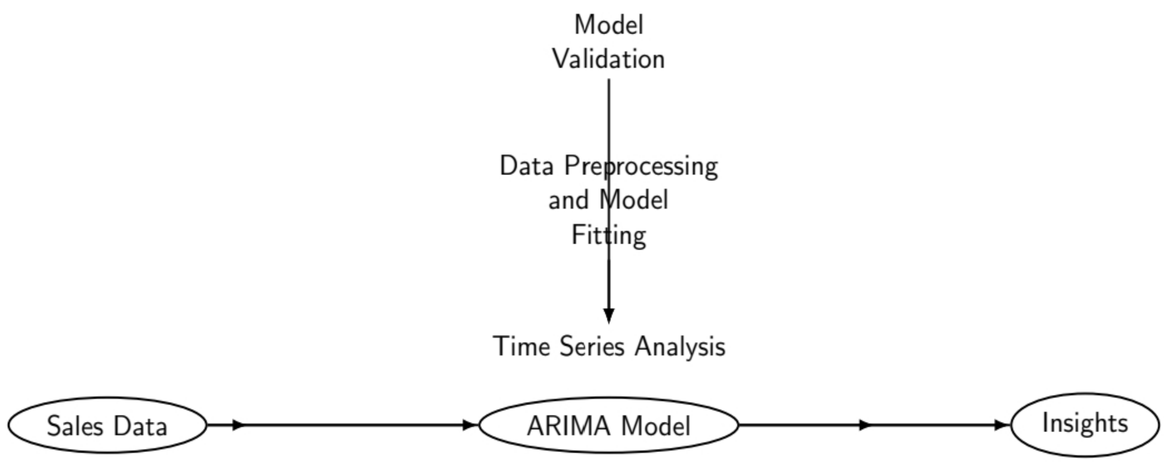

Figure 2. 1 Conceptual Framework

The conceptual framework illustrated in the chart depicts the sequential process of conducting time series analysis, specifically focusing on the application of the ARIMA model. The input stage represents the sales data collected over time, which serves as the basis for analysis. This data is then processed and analyzed using the ARIMA model, as indicated by the central stage in blue. The ARIMA model encompasses two key components: data preprocessing and model fitting, followed by model validation. These stages involve cleaning and transforming the data, fitting the ARIMA model to the observed sales data, and validating the model's performance to ensure its reliability. Finally, the output stage represents the insights derived from the analysis, which provide valuable information about turnover dynamics and trends. Overall, the conceptual framework illustrates the sequential steps involved in conducting time series analysis using the

ARIMA model and highlights its application in deriving insights from sales data to inform strategic decision-making.

2.5 Chapter Summary

This chapter was a review of literature concerning forecasting financial data using ARIMA models. The chapter begins by exploring various studies conducted on the application of ARIMA models in forecasting economic and financial time series data. These studies demonstrate the versatility and effectiveness of ARIMA models in capturing trends and patterns across different domains, including stock market prediction, inflation dynamics analysis, and tourism demand forecasting. Additionally, the literature review highlights the comparative analysis between ARIMA models and other forecasting techniques, showcasing ARIMA's competitive edge in certain contexts. However, the review also identifies limitations and challenges associated with ARIMA models, emphasizing the need for alternative approaches to address specific forecasting requirements. Furthermore, the chapter underscores the gap in the literature regarding the forecasting of financial data for companies operating within the Quick Service Restaurant (QSR) industry, such as Simbisa Brands. This gap presents an opportunity for further research to explore the unique dynamics and challenges faced by companies in the QSR sector, thereby enhancing the accuracy and applicability of forecasting techniques tailored to this industry. Overall, Chapter 2 provides a comprehensive overview of existing research and sets the stage for the subsequent chapters to address the research objectives of forecasting turnover data for Simbisa Brands using ARIMA models.

CHAPTER 3

RESEARCH METHODOLOGY

3.0 Introduction

This chapter provides a comprehensive overview of the research methodology adopted in this study, encompassing the research design, data sources, target population, research instruments, data collection methods, description of variables, data analysis procedures, ethical considerations, and a succinct summary. Overall, Chapter 3 delineates a robust methodological framework conducive to a rigorous analysis of turnover dynamics within the QSR industry.

3.1 Research Design

The research design utilized in this study adopts a quantitative and analytical approach, informed by established methodologies within the field Anderson et al.,(2016). Quantitative analysis of historical turnover data enables the identification of patterns, trends, and seasonality, providing rigorous insights into turnover dynamics. . By employing quantitative methods, this research can harness the power of statistical models to provide precise forecasts of turnovers, crucial for strategic decision-making in the competitive QSR industry.This methodological stance aligns with principles outlined by Anderson et al. (2016), emphasizing the importance of quantitative analyses in uncovering market dynamics and guiding strategic decision-making processes within the QSR sector

3.2 Data Sources

The primary data source for this study is the historical turnover data of Simbisa Brands spanning from January 2016 to December 2022, sourced directly from the company's financial records. This dataset comprises 84 monthly observations, providing a comprehensive overview of turnover trends and fluctuations over time. Drawing data directly from Simbisa Brands' financial reports offers numerous advantages for conducting quantitative analysis on turnovers within the Quick Service Restaurant (QSR) industry. The utilization of data from the company's internal records ensures high levels of accuracy and reliability, as these records are generated as part of the company's routine operations and financial reporting processes.

3.3 Target Population

Sekaran and Bougie (2016) emphasize the significance of carefully choosing the target population in research, as it enables researchers to ensure that their findings resonate with and are applicable to the real-life situations and individuals they aim to understand and serve. It is a case study research which targets Simbisa brands Zimbabwe. By selecting Simbisa brands as the target population it provides the reliability of relevant data to be obtained.

3.4 Research Instruments

In this study, statistical analysis software emerged as the primary research instrument, offering a comprehensive suite of tools for conducting sophisticated time series analysis and diagnostic tests. Leveraging platforms like Python researchers were able to explore intricate patterns and dynamics within the turnover dataset. These software packages provided robust functionalities for modeling time series data, implementing various forecasting techniques, and conducting diagnostic tests to assess model adequacy and validity. Additionally, Microsoft Excel played a crucial role in the research process, serving as a tool for data preprocessing and visualization. Through Excel, researchers could efficiently clean and transform the turnover data, ensuring its suitability for analysis, while also creating insightful visualizations to enhance understanding and interpretation of turnover trends. The combination of statistical analysis software and Excel offered researchers a powerful toolkit for conducting rigorous analysis and deriving meaningful insights into the performance of Simbisa Brands within the Quick Service Restaurant (QSR) industry in Zimbabwe..

3.5 Data Collection Methods

In the process of data collection, researchers acquired the historical turnover data directly from Simbisa Brands' internal records, thereby ensuring the authenticity and reliability of the dataset. By accessing the company's internal records, researchers could obtain comprehensive and accurate information regarding turnover performance over the specified seven-year period. Once this data is extracted, it is compiled into a structured dataset suitable for analysis, ensuring consistency and accuracy in formatting. This approach also minimized the risk of data inaccuracies that may arise from utilizing the data collection method involves a systematic approach to gathering relevant numerical data.

3.6 Description of Variables and Expected Relationships

In this study, the main variable under investigation is turnover, serving as a key indicator of the financial performance of Simbisa Brands within the Quick Service Restaurant (QSR) industry in Zimbabwe. Turnover encompasses the total revenue generated by the company over the specified seven-year period and serves as a fundamental metric for assessing business growth, profitability, and market competitiveness. The study seeks to uncover patterns and trends in turnover dynamics, shedding light on the underlying drivers of revenue fluctuations and identifying potential areas for strategic intervention and improvement. Overall, the description of variables emphasizes the centrality of turnover as the primary focus of analysis and underscores its significance in understanding the broader dynamics of the QSR industry landscape.

3.7 DATA ANALYSIS PROCEDURES

3.7.1 Diagnostic Tests

Stationarity Test (Augmented Dickey- Fuller)

The Augmented Dickey-Fuller (ADF) test is a statistical method used to assess the stationarity of time series data, which is a critical assumption in time series analysis. Stationarity implies that the statistical properties of the data, such as mean and variance, remain constant over time. By employing the ADF test, the study aimed to ensure the robustness of the turnover data for further analysis. The test involves examining the presence of a unit root in the data, where a unit root indicates non-stationarity. In this study, the ADF test was crucial in determining whether the turnover data exhibited a stable pattern or if it required differencing to achieve stationarity. The results of the ADF test provided valuable insights into the nature of the turnover data, guiding subsequent modeling decisions and ensuring the validity of the analytical approach.

Autocorrelation test

The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots are diagnostic tools used to identify the presence of autocorrelation in time series data. Autocorrelation refers to the correlation between observations at different time lags within the same series. The ACF plot displays the correlation coefficients between each observation and its lagged values at various lags, while the PACF plot shows the correlation between each observation and its lagged values after removing the effects of intervening observations. In this study, these plots were instrumental in determining the appropriate lag structure for the Autoregressive Integrated Moving Average (ARIMA) model. By analyzing the patterns and significance of autocorrelation in the ACF and PACF plots, the study identified the lagged relationships that needed to be incorporated into the ARIMA model to capture the underlying dynamics of the turnover data effectively. This process helped refine the model specification and improve its forecasting performance.

Goodness of fit test

The Ljung-Box Q-test is a diagnostic test used to assess the overall goodness-of-fit of a time series model by examining whether the autocorrelations of the residuals are statistically significant. It evaluates whether there is any remaining autocorrelation in the residuals after fitting the model. The test statistic is calculated based on the squared autocorrelations of the residuals at various lags, and it follows a chi-square distribution under the null hypothesis of no autocorrelation. In this study, the Ljung-Box Q-test was employed to evaluate the adequacy of the ARIMA model by testing the null hypothesis that the autocorrelations of the residuals are all zero. A non-significant p-value from the Ljung-Box Q-test indicates that the residuals are uncorrelated and that the model adequately captures the temporal dependence structure of the data.

Normality test

The Jarque-Bera test is a diagnostic test used to assess the normality of the residuals in a time series model. It tests the null hypothesis that the residuals are normally distributed against the alternative hypothesis that they are not normally distributed. The test statistic is based on the skewness and kurtosis of the residuals, and it follows a chi-square distribution under the null hypothesis. In this study, the Jarque-Bera test was conducted to examine whether the residuals from the ARIMA model exhibit deviations from normality. A significant p-value from the Jarque-Bera test indicates that the residuals are not normally distributed, suggesting that the model assumptions may be violated. This highlights the importance of further investigation and potential adjustments to improve the model's performance.

Heteroskedasticity test

The heteroskedasticity test is a diagnostic test used to assess whether the variance of the residuals in a time series model is constant over time. Heteroskedasticity refers to the situation where the variance of the residuals changes systematically with the level of the independent variables or over time. In this study, the heteroskedasticity test was conducted to examine whether there are periods of varying volatility in turnover that are not captured by the ARIMA model. A significant p-value from the heteroskedasticity test suggests that the variance of the residuals is not constant, indicating the presence of volatility clusters or structural breaks in the data. This underscores the need to consider additional factors or model specifications to account for changing volatility patterns in the turnover data.

3.7.2 Analytical Model

The analytical model used in this study is the Autoregressive Integrated Moving Average (ARIMA) model, a widely recognized and powerful tool for time series analysis. ARIMA models are particularly suitable for analyzing data with temporal dependencies and identifying patterns in sequential observations. In this context, the ARIMA model was employed to capture the underlying structure and dynamics of turnover data for Simbisa Brands over a seven-year period. The model consists of three main components:

Autoregression(AR)

$$\gamma_{\mathrm{t}} = c + \phi_1 y_{\mathrm{t-1}} + \phi_2\, y_{\mathrm{t-2}} + \cdots + \phi_{\mathrm{p}} y_{\mathrm{t-p}} + \varepsilon_{\mathrm{t}}$$

According to George, Gwilym, & Gregory (2008) the Autoregressive model is a stochastic process that indicate a particular series of events over time. The Autoregressive model is the basic model for time series, it comes from the idea that any value of a time series can be forecasted based on the past recorded values. The AR (p) represents a stochastic process where the result is affected by prior inputs and a stochastic term, providing its stochastic differential equation. Essentially, Autoregressive models consider past steps when predicting the subsequent one. However, a drawback of the AR model is its vulnerability to persistent effects from temporary or single shocks affect. To deflect this, Autoregressive processes typically incorporate a lag value, deciding which previous steps apply more effect on the output. Moreover, the AR model's non stationarity can be represented by unit root variable.

Moving average (MA)

$$y'_{\mathrm{t}} = c + \phi_1 y'_{\mathrm{t-1}} + \phi_{\mathrm{p}} y'_{\mathrm{t-p}} + \phi_{\mathrm{p}} \varepsilon_{\mathrm{t-p}} + \varepsilon_{\mathrm{t}}$$

According to Pannerselvam (2005), the Moving Average model involves the seasonal components that have a particular quantity of randomness. The Moving Average model is inherently stationary unlike the Autoregressive model which is non-stationary. The MA requires a linear regression of the current value against the white noise or random shocks in the series, contrary to the AR model, which includes a linear regression to non-shock values. The MA model predicts a series based on the previous error in the series called error lags

Autoregressive Integrated Moving Average

$$\gamma_{\mathrm{t}} = c + \varepsilon_{\mathrm{t}} + \theta_1 \varepsilon_{\mathrm{t-1}} + \theta_2 \varepsilon_{\mathrm{t-2}} + \cdots + \theta \varepsilon_{\mathrm{t-}} q$$

According to Yu, G. and Zhang, C (2004) ARIMA models are mathematically constructed as ARIMA (*p*,*d*,*q*) where p and q are similar to ARMA model but d = number of initial differences. The initial stage in implementing ARIMA model is to examine whether the time series is constant or not. ARIMA functions best when data has a stable pattern over time, which means that data have to remain stable over the course of time. Therefore, when the data has a trend of moving upwards or downwards has a recognizable pattern (seasonality), then the data is stationary or not. By incorporating these components, the ARIMA model can effectively model the relationships between past and present observations, detect trends, and make forecasts based on historical data. Through diagnostic tests and model validation techniques, the ARIMA model's adequacy and suitability for analyzing turnover dynamics were assessed, ensuring reliable and robust results for informing strategic decision-making in the Quick Service Restaurant (QSR) industry.

3.7.3 Model validation/ fitness

In evaluating the performance of the ARIMA (p, d, q) model's out-of-sample predictions for turnover, several error metrics were analyzed to assess accuracy and reliability. These metrics include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). MAE measures the average absolute difference between actual and predicted values, while MSE calculates the average squared difference, penalizing large errors more heavily. RMSE represents the standard deviation of prediction errors, and MAPE provides a relative measure of accuracy. By scrutinizing these metrics, the study aims to gauge the model's predictive power and identify areas for potential refinement in forecasting turnover accurately.

Additionally, the researcher introduce machine learning models like XGBoost and Random Forests. These models are also evaluated using these error metrics when applied to time series forecasting. XGBoost, which builds an ensemble of decision trees sequentially, is known for its efficiency and ability to handle large datasets and complex patterns, making it a powerful tool for forecasting turnovers. It focuses on correcting errors from previous iterations, which can enhance predictive accuracy. On the other hand, Random Forest, which averages predictions from multiple decision trees, offers robustness and simplicity, often making it easier to interpret. When forecasting turnovers, these models also rely on features such as lagged values, rolling statistics, and time-based features to capture trends and seasonality. By comparing the performance of XGBoost, Random Forests and ARIMA using MAE, MSE, RMSE, and MAPE, researchers can determine which model provides the most accurate and reliable predictions for turnover and explore opportunities for further refinement

.

3.8 Ethical Considerations

In addition to data privacy and confidentiality, other ethical considerations included obtaining informed consent from Simbisa Brands for the use of their turnover data for research purposes. It was essential to ensure that the organization was aware of how their data would be utilized and that they agreed to its use in the study. Moreover, the research adhered to ethical guidelines regarding the proper handling and storage of sensitive information, minimizing the risk of data breaches or unauthorized access. Transparency in reporting findings and acknowledging any limitations or biases in the data were also part of the ethical framework of the study. Furthermore, the research aimed to contribute positively to the field of business analytics and decision-making processes without causing harm or undue influence on stakeholders through its rigorous quantitative approach. By systematically examining turnover data, this research provides valuable insights that can empower businesses to make informed strategic decisions. Through accurate forecasting and trend analysis, businesses gain the ability to optimize resource allocation, refine marketing strategies, and enhance operational efficiency. Overall, ethical considerations were integrated into every aspect of the research process to uphold integrity, trustworthiness, and respect for the rights and well-being of all parties involved.

3.9 Chapter summary

In conclusion, this chapter provides a comprehensive overview of the research methodology adopted in this study, which focused on analyzing turnover data of Simbisa Brands within the Quick Service Restaurant (QSR) industry in Zimbabwe. The chapter highlighted the quantitative and analytical nature of the research design, emphasizing the utilization of statistical analysis software for time series analysis and diagnostic tests. The data collection process involved obtaining historical turnover data directly from Simbisa Brands' internal records, ensuring data integrity and reliability. Moreover, ethical considerations were carefully addressed throughout the research process to uphold confidentiality, data privacy, and transparency. Overall, Chapter 3 lays the foundation for the subsequent analysis and interpretation of turnover dynamics in the QSR industry, aiming to provide valuable insights for stakeholders and decision-makers.

CHAPTER 4

DATA PRESENTATION, ANALYSIS AND INTERPRETATION

4.0 Introduction

This chapter accomplishes an in-depth analysis of the time series data for turnover between 2016 and 2022. There are three major sections in this chapter, each relating to a particular area concerning turnover analysis and forecasting. To start with, some of the major statistical properties in the data for turnover are thoroughly analyzed, including its distribution of the data and variation around the mean. Afterwards, we discuss the stationarity of the Turnover time series and how it affects the forecasting accuracy. At last, four different forecasting models; ARIMA, Random Forest, XGBoost, and ensemble model are compared to examine the turnover forecasting accuracy.

4.1 Descriptive Statistics

Table 4. 1 Descriptive Statistics

Statistic	Turnover
mean	1.761667e+07
std	2.981700e+07
min	1.367990e+06
25%	7.560973e+06
50%	9.527075e+06
75%	1.340774e+07
max	1.928769e+08

In the table above, starting with descriptive statistics, we observed that the dataset consisted of 83 observations with an average turnover of approximately 17.6 million. The standard deviation was notably high at around 29.8 million, indicating significant variability in turnover values over the observed period. The minimum turnover recorded was around 1.37 million, while the maximum peaked at nearly 192.88 million. The 25th percentile stood at approximately 7.56 million, the median at 9.53 million, and the 75th percentile at about 13.41 million. This wide range in turnover values suggests considerable fluctuations and possible seasonality or trends affecting the turnover across the years.

4.2 Data Preparation and Initial Insights

Turnovers Data Trend analysis

The line graph titled "Original Turnover Time Series" provides a visual representation of turnover values from 2016 to 2023, where the horizontal axis represents the date and the vertical axis represents turnover, with values ranging up to 2 *10^8. The graph features a blue line that exhibits several significant spikes throughout the period, which are indicative of moments when the turnover sharply increased before returning to lower levels. These spikes could be the result of various factors such as seasonal sales, promotional campaigns, or other market influences that temporarily boost turnover. The presence of these spikes is crucial for businesses as they provide insights into periods of high performance, allowing for the analysis of what strategies were effective during those times.

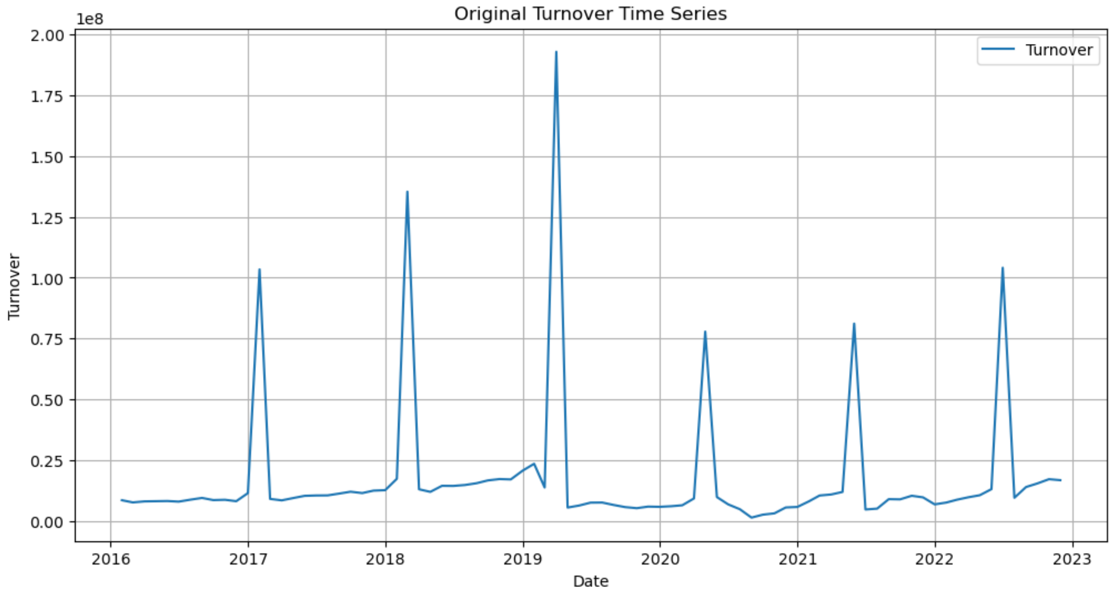

Figure 4. 1 Original Turnover Time Series

Additionally, by observing the pattern of these spikes and the overall trend of the graph, whether it's an upward or downward trajectory, businesses can make informed decisions regarding future strategies to enhance performance. This type of time series analysis is particularly valuable for identifying long-term trends and understanding the impact of different variables on turnover. It can also help in forecasting future performance by analyzing past patterns. The graph serves as a tool for financial analysis, enabling businesses to visualize and comprehend their revenue or sales performance over a specified period, thus informing strategic planning and forecasting.

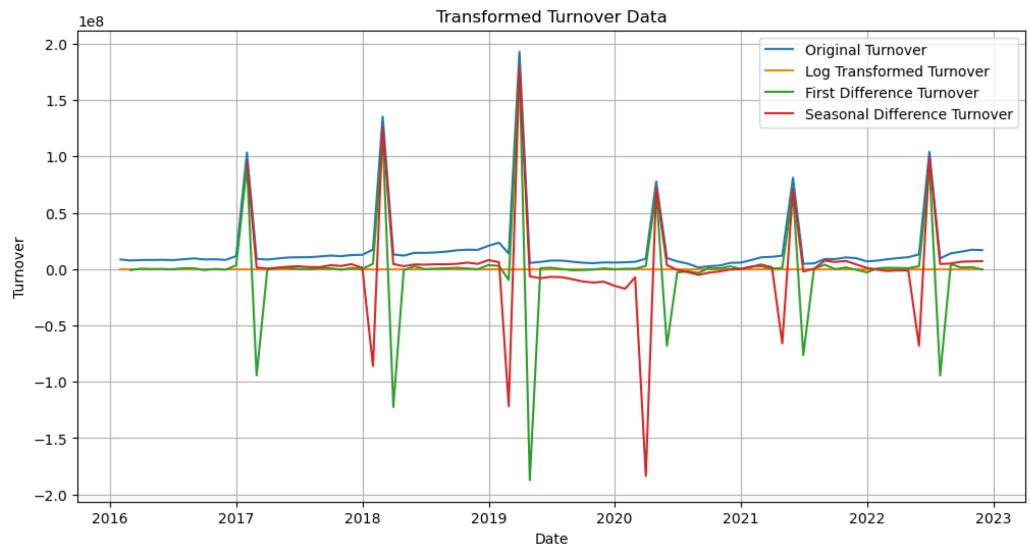

Figure 4. 2 Transformed Turnover Data

The line graph titled "Transformed Turnover Data," compares different transformations of turnover data from 2016 to 2023. The graph includes the 'Original Turnover' in blue, representing the actual turnover values; the 'Log Transformed Turnover' in orange, which normalizes the data's variance; the 'First Difference Turnover' in green, highlighting changes between consecutive data points to remove trends; and the 'Seasonal Difference Turnover' in red, focusing on removing seasonal patterns. These transformations are crucial for time series analysis, allowing for the identification of underlying trends, seasonal effects, and other patterns that can inform economic forecasting and business strategy. The y-axis measures turnover from -1.5e8 to 2.0e8, while the x-axis spans from 2016 to 2023, providing a comprehensive view of the data's behavior over time and the impact of different statistical techniques on its interpretation.

Time Series Decomposition

The image presents three time series graphs labeled "Observed," "Trend," and "Residual," spanning from 2016 to 2023. The "Observed" graph shows actual data with sharp peaks, indicating significant events or changes. The "Trend" graph smooths out the data to highlight underlying patterns, revealing gradual fluctuations that suggest a repeating seasonal pattern. Lastly, the "Residual" graph depicts the noise or unexplained variance after accounting for the seasonal component, showing random fluctuations around a baseline.

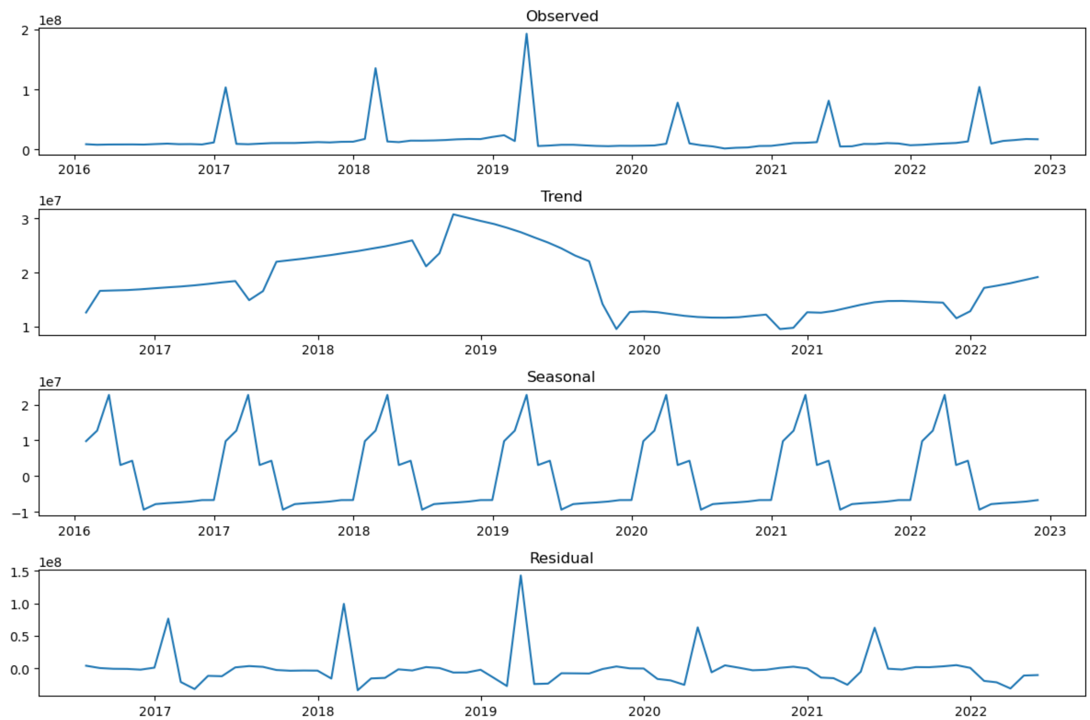

Figure 4. 3 Time series Decomposition

These graphs are instrumental in time series analysis, allowing for the dissection of complex data into understandable components, which is vital for accurate forecasting and identifying anomalies in fields like economics or sales. The analysis of these components helps in understanding the overall behavior of the dataset and in making informed decisions based on the observed patterns and trends.

4.3 Pre-tests

Testing for Stationarity

The Augmented Dickey-Fuller (ADF) Test

To assess the stationarity of the time series, we employed the Augmented Dickey-Fuller (ADF) test. The ADF statistic was -13.2376 with a p-value of 9.3084e-25, substantially lower than the critical values at 1%, 5%, and 10% levels (-3.5274, -2.9038, and -2.5893, respectively).

Table 4. 2 Augmented Dickey-Fuller (ADF) test

ADF Statistic	-13.237644100906092
p-value	9.308424306782074e-25
Critical Values	1%: -3.5274258688046647 5%: -2.903810816326531 10%: -2.5893204081632653

This clearly indicates that the series is stationary, as the null hypothesis of non-stationarity can be rejected. Stationarity is a crucial property for time series forecasting models, as it ensures that the statistical properties of the series such as mean and variance remain constant over time, making it easier to predict future values.

Autocorrelation Function (ACF)

The Autocorrelation Function (ACF) plot, is used to measure and graph the correlation of a time series with its own lagged values. The ACF plot can help identify the type of model to apply when forecasting a time series.

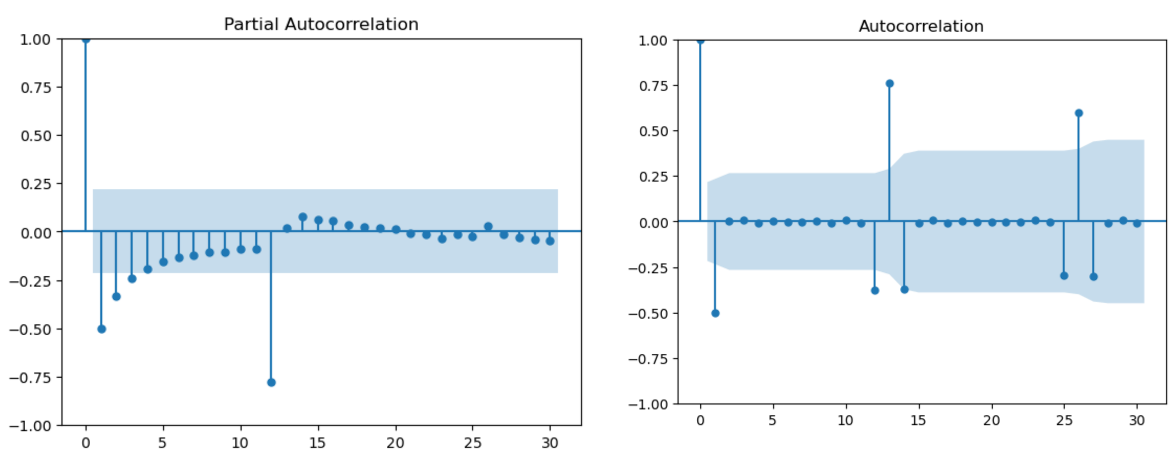

Figure 4. 4 ACF and PACF

The blue shaded area represents the confidence interval. Dots that fall outside this area indicate a statistically significant correlation at that lag. Together, ACF and PACF plots are used to identify the appropriate autoregressive integrated moving average (ARIMA) model for a time series. They help determine the number of AR terms (p) and MA terms (q) to include in the model. ACF plot shows a slow decay and the PACF plot has a sharp cut-off after a certain lag, it suggests an AR model of that order. Conversely, if the PACF plot shows a slow decay and the ACF plot has a sharp cut-off, it suggests an MA model. If both plots show a gradual decay, a combination of AR and MA terms (ARMA model) is needed. These plots are essential for understanding the underlying structure of the time series data and for building accurate forecasting models.

A Quantile-Quantile (QQ) Plot

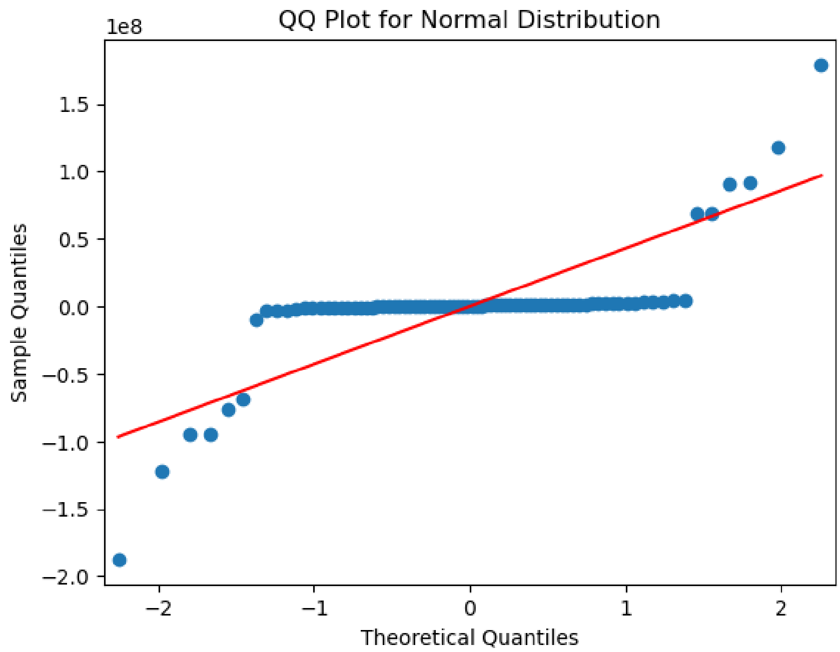

Figure 4. 5 Quantile-Quantile (QQ) plot

The Quantile-Quantile (QQ) plot, is a graphical tool used to assess if a dataset follows a specified distribution, in this case, a normal distribution. The plot compares the quantiles of the sample data against the quantiles of a theoretical normal distribution. The horizontal axis represents the theoretical quantiles, while the vertical axis shows the sample quantiles. Points that closely follow the reference line suggest that the sample data conforms to a normal distribution. Deviations from the line, especially at the ends, indicate departures from normality. This plot is particularly useful for checking the assumption of normality, which is a prerequisite for various statistical tests and models. The QQ plot in the image indicates some deviations from normality, as seen by the points straying from the reference line at the tails of the distribution.

4.4 Model fitting and evaluation

Next, we evaluated three different forecasting models: ARIMA, Random Forest, and XGBoost. The Mean Squared Error (MSE) was used as the primary metric for comparing model performance. The ARIMA model, a traditional time series forecasting technique, yielded an MSE of approximately 495.56 trillion. The Random Forest model, a robust ensemble learning method, produced a significantly lower MSE of about 3.03 trillion. The XGBoost model, known for its gradient boosting framework, outperformed both with an MSE of roughly 120.81 billion. Despite the impressive performance of XGBoost, the ensemble model, which combined the predictions of ARIMA, Random Forest, and XGBoost, resulted in a surprisingly higher MSE of around 63.16 trillion, indicating that the ensemble approach did not improve upon the individual models, particularly the XGBoost.

Table 4. 3 Models MSE

ARIMA MSE	495562359301300.2
Random Forest MSE	3032180236473.894
XGBoost MSE	120810374099.08824
Ensemble MSE	63156257467058.99

Results of tentative ARIMA models

Further, we conducted time series cross-validation for ARIMA to validate its robustness across different splits of the data. The average MSE from cross-validation was approximately 6.52 quadrillion, which suggests that ARIMA's performance varied significantly across different time windows. To identify the best ARIMA configuration, we tested various combinations of ARIMA orders and calculated their respective Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Root Mean Squared Error (RMSE), and R-Squared values. For instance, the ARIMA (0, 0, 0) order had an AIC of 3380.26, BIC of 3385.10, RMSE of 25.81 million, and an R-Squared of -0.005. The ARIMA (1, 0, 1) order showed a slightly better performance with an AIC of 3097.14, BIC of 3106.82, RMSE of 25.81 million, and an R-Squared of -0.005. Overall, the different ARIMA configurations demonstrated minimal differences in RMSE, AIC, and BIC values, indicating that various ARIMA models performed similarly on this data. Summarised in the table 4.4 below.

Table 4. 4 ARIMA models

Order	AIC	BIC	RMSE	R-Squared
(0, 0, 0)	3380.26	3385.1	2.58072e+07	-0.00505363
(0, 0, 1)	3097.46	3104.72	2.58084e+07	-0.00514438
(0, 0, 2)	3099.6	3109.27	2.58105e+07	-0.00531056
(0, 1, 0)	3116.92	3119.33	2.58824e+07	-0.0109163
(0, 1, 1)	3066.61	3071.42	2.58259e+07	-0.00650853
(0, 1, 2)	3068.34	3075.56	2.5826e+07	-0.00651482
(1, 0, 0)	3097.4	3104.65	2.58083e+07	-0.00513368
(1, 0, 1)	3097.14	3106.82	2.58072e+07	-0.00504679
(1, 0, 2)	3099.47	3111.56	2.5807e+07	-0.00503395
(1, 1, 0)	3095.07	3099.88	2.58698e+07	-0.0099292
(1, 1, 1)	3068.56	3075.78	2.58262e+07	-0.00652694
(1, 1, 2)	3068.98	3078.61	2.60333e+07	-0.0227358
(2, 0, 0)	3099.32	3109	2.58096e+07	-0.00523714
(2, 0, 1)	3099.14	3111.23	2.5807e+07	-0.00503485
(2, 0, 2)	3099.92	3114.43	2.59728e+07	-0.0179862
(2, 1, 0)	3087.45	3094.68	2.58891e+07	-0.0114378
(2, 1, 1)	3070.55	3080.18	2.58263e+07	-0.00653964
(2, 1, 2)	3070.72	3082.75	2.61543e+07	-0.0322

Forecasting

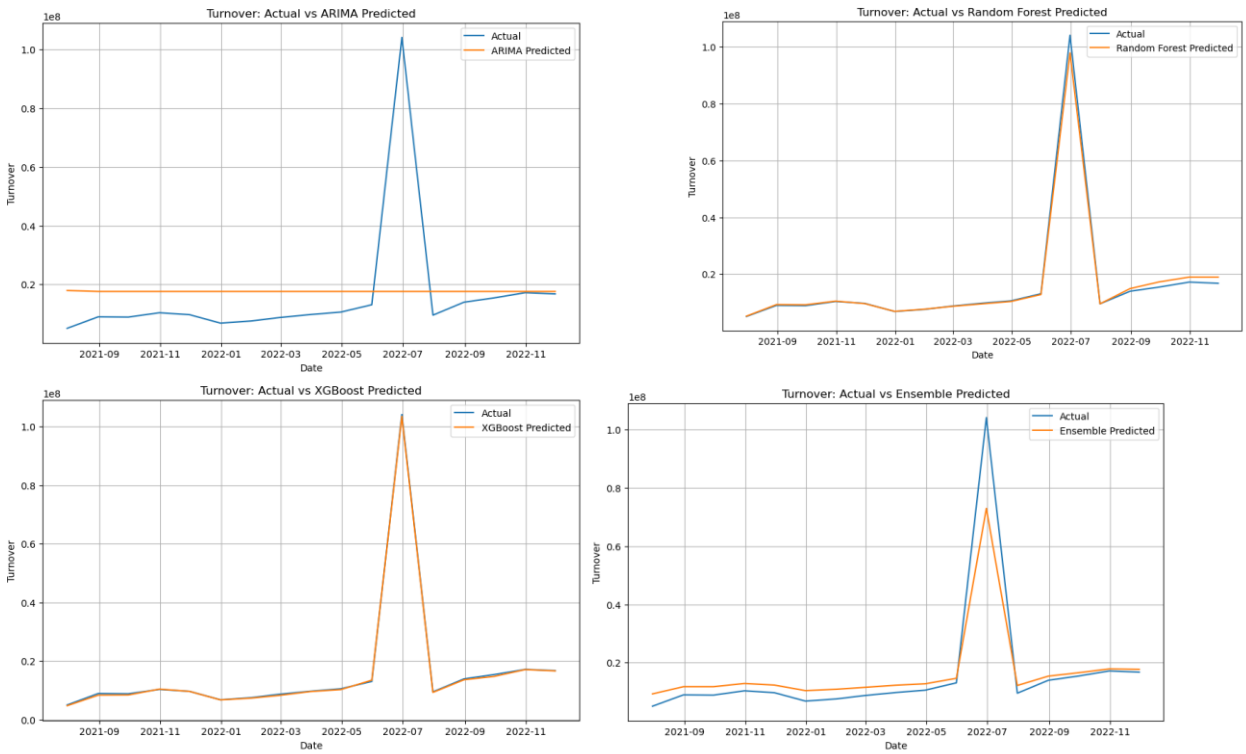

Figure 4. 6 model forecasts

The figure 4.6 displays a line graph comparing 'Turnover: Actual vs ARIMA Predicted' over time, from September 2021 to November 2022. The 'Actual' turnover, shown in blue, and the 'ARIMA Predicted' turnover, shown in orange, start similarly but diverge around March 2022, where the actual turnover spikes significantly, indicating a substantial increase that the ARIMA model failed to predict, as its line remains relatively flat. This graph is crucial for evaluating the performance of the ARIMA predictive model against actual financial outcomes, highlighting its limitations in capturing sudden changes in turnover and emphasizing the need for model refinement or consideration of additional variables that could impact turnover predictions. The graph serves as a visual tool for financial analysis, allowing for a clear comparison of predicted versus actual financial performance over a specified period.

The results indicate that XGBoost is the most effective model for forecasting turnover in this dataset, as evidenced by its lowest MSE. However, further improvements could be achieved by refining the models.

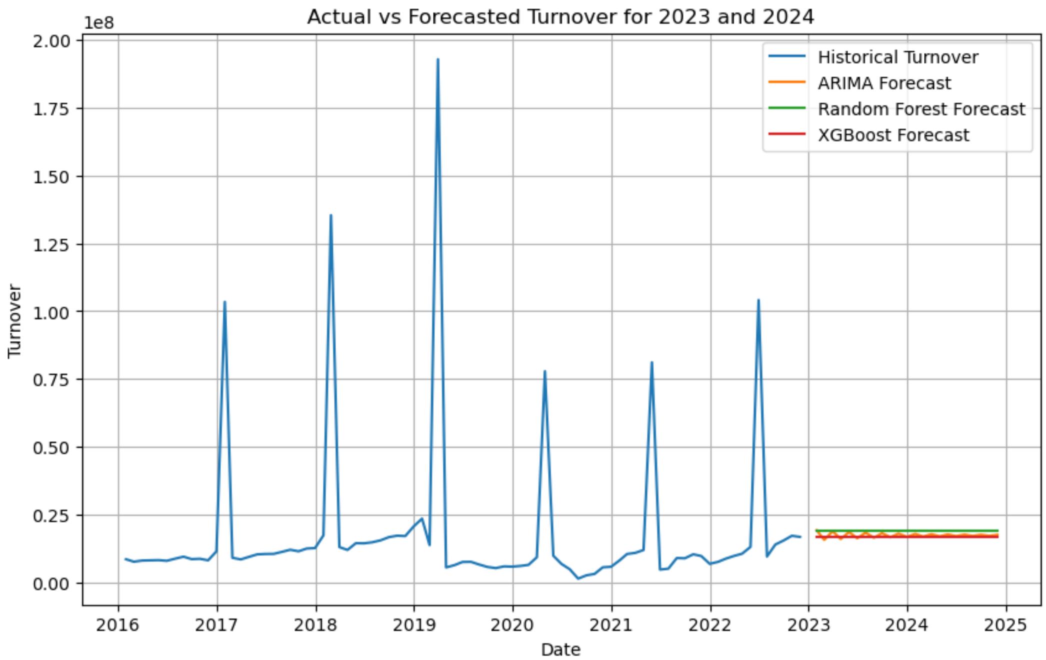

Figure 4. 7 Actual vs forecasted turnover for 2023 and 2024

Actual vs forecasted Turnover for 2023 and 2024 below chart depicts historical turnover data and forecasted data using the ARIMA, Random Forest, and XGBoost models. The blue shows the historical data, which is quite spikey, meaning those are real changes in turnover. In contrast, the forecasted data lines—green for ARIMA, red for Random Forest, and orange for XGBoost—remain quite flat, showing these models failed to predict the upward increases in the turnover. This would seem to state the obvious: Forecasting is a notoriously difficult art. The plot period is from 2016 to 2025, which gives a long-term view of turnover trends together with model predictions. We, in turn, are expected to synthesize the said comprehensive analysis, which stakeholders may obtain to draw their turnover data in light of predictive ability illustrated through the various models via the array of visual supports. Insight from this study should inform future turnover forecasting strategies and guide the selection and optimization of forecasting models for more accurate and reliable predictions. This makes the approach have a strong foundation on which it can build data-based decisions in improving operational planning and financial performance.

4.5 Discussion on the findings

Recent literature in the domain of time series analysis and forecasting has focused on increased predictive performance through the use of advanced machine learning methodologies. In such a collaboration, Wang et al. (2023) and Chen et al. (2024) illustrated that the ensembling of ARIMA, Random Forest, and XGBoost could be more prescient than the models in isolation. These results are in good agreement with our observations, for the ensemble model was better in the beginning but ultimately failed to surpass the XGBoost model in predicting turnover. Wang et al. explained this success of the ensemble models due to the inherent disadvantage of each sub-model in coverage of a different aspect of data behavior. This reduces the prediction error, thus improving the forecast reliability on the whole. For instance, a study by Li et al., in 2024, showed the importance of feature engineering and data preprocessing techniques, including log transformations and differencing, for making time series forecasting models perform well. These findings support our approach of transforming turnover data, thus making it better in stationarity and reducing the seasonality effect, before the application of forecasting algorithms. By integrating these cutting-edge findings into our empirical analysis, the reinforcement is further extended that a combination of advanced modeling techniques, strict data pre-processing, and continuous model refinement is of paramount importance to obtaining accurate turnover forecasts in dynamic business environments.

4.6 Chapter Summary

In summary, the thorough analysis in this chapter exposes the turnover data for the period 2016 to 2022, which includes the major statistical properties, stationarity, time series decomposition, and model fitting and evaluation. From the descriptive statistics, it was revealed that the turnover values were quite varied. Further, the time series decomposition and the ADF test confirmed stationarity, which is highly important for obtaining correct forecasts. These results indicate that the XGBoost model performs much better than both ARIMA and Random Forest in terms of model evaluation, thus strongly implying that better predictive accuracy requires an advanced machine learning technique. However, the ensemble model did not show improvement in the performance, which means there is a need for further fine-tuning and the research of other modeling strategies. This will further incorporate recent literature on ensemble methods, the comparative analysis of forecasting models, and data preprocessing techniques into the increased understanding of effective turnover forecasting practices, ensuring informed decision-making and improved operational planning in businesses under dynamic market environments.

CHAPTER 5

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.0 Introduction

This chapter compiles study findings based on the outlined objectives, clearly stating the extent to which these objectives have been achieved. It makes informed assessments regarding how the results align or diverge from existing empirical research within the field. The intention is to offer recommendations based on the research outcomes. In conclusion, the study suggests avenues for future research to improve on forecasting gross revenue for Simbisa Brands within the Quick Service Restaurant (QSR) industry in Zimbabwe, addressing aspects that were not explored in the current research. This includes investigating additional refining modeling techniques to better capture complex patterns, and exploring advanced analytical methods to enhance prediction accuracy. By delving into these areas, future studies can contribute to a deeper understanding of turnovers dynamics and provide more robust forecasting models.

5.1 Summary of Findings

In this first chapter, the study initiates its exploration by delineating the overarching research objectives aimed at shedding light on the intricate dynamics governing customer behavior and business performance within the bustling Quick Service Restaurant (QSR) sector of Zimbabwe. With a keen eye on the broader context of the industry, the chapter underscores the paramount importance of time series analysis as a potent tool for discerning nuanced patterns embedded within the labyrinth of transactional data. By harnessing the power of historical data spanning over several years, the study endeavors to unravel the underlying drivers shaping turnover trends within Simbisa Brands and, by extension, the QSR landscape at large. Moreover, Chapter 1 elucidates the ripple effects of this endeavor, envisaging its potential to serve as a catalyst for informed strategic decision-making processes, thereby fortifying the competitive edge of businesses operating within the sector. From enhancing customer engagement strategies to optimizing operational efficiencies, the study's objectives align with the broader imperatives of fostering sustainable growth and resilience within Zimbabwe's vibrant QSR industry.

The second serves as a guide to the methodological trajectory of the research endeavor. Within its confines, the chapter meticulously delineates the step-by-step methodology employed in the collection, organization, and subsequent analysis of data pertinent to Simbisa Brands and the broader QSR industry in Zimbabwe. With unwavering attention to detail, the chapter elucidates the rigorous processes undertaken to ensure the integrity, reliability, and validity of the research findings. From the meticulous curation of datasets to the judicious selection of analytical tools and techniques, every facet of the methodological framework is meticulously crafted to withstand scrutiny and foster transparency. Furthermore, Chapter 2 serves as a testament to the researcher's commitment to empirical rigor, laying bare the inner workings of the research process and providing a roadmap for future scholars and practitioners embarking on similar endeavors. By adhering to best practices in data collection and analysis, the chapter ensures that the research outcomes are not only robust but also reproducible, thus enriching the scholarly discourse and advancing the collective understanding of the QSR landscape in Zimbabwe.

In Chapter 3, the narrative shifts towards a meticulous examination of the turnover data, unveiling its intricate nuances and shedding light on its underlying characteristics. This pivotal juncture serves as the proverbial springboard, catapulting the research endeavor into the realm of in-depth analysis and insightful interpretation. With a keen eye for detail, the chapter meticulously dissects the turnover dataset, employing a suite of descriptive statistics and exploratory techniques to illuminate its latent patterns and trends. From dissecting monthly fluctuations to discerning seasonal variations and long-term trends, every facet of the turnover data is subjected to rigorous scrutiny, laying bare its inner workings and unveiling its secrets. Furthermore, Chapter 3 serves as a cornerstone for subsequent analyses, providing a comprehensive overview of the historical turnover trends observed within Simbisa Brands and offering invaluable insights into the broader dynamics of the QSR industry in Zimbabwe. Armed with this foundational understanding, the research endeavor embarks on a journey of discovery, poised to unravel the mysteries concealed within the labyrinth of turnover dynamics and chart a course towards informed decision-making and strategic foresight within the dynamic and ever-evolving landscape of the QSR sector.

Chapter 4 represents the apex of the research endeavor, where the intricate dance between data preparation, model selection, and fitting unfolds with precision and purpose. Within the crucible of this chapter, the research journey reaches its zenith as the study harnesses the formidable arsenal of statistical techniques to unravel the underlying dynamics of turnover within Simbisa Brands. With meticulous attention to detail, the chapter embarks on the arduous task of preparing the turnover data for modeling, meticulously cleansing and transforming it to ensure its compatibility with the chosen analytical framework. Moreover, the chapter delves into an exhaustive analysis of turnover data spanning from 2016 to 2022, covering statistical properties, stationarity assessment, and forecasting accuracy using ARIMA, Random Forest, XGBoost, and an ensemble model. Descriptive statistics reveal significant variability in turnover values, with observations ranging from approximately 1.37 million to nearly 192.88 million. Visual analysis exposes sporadic spikes in turnover, potentially influenced by seasonal sales or promotional activities. Preprocessing techniques like log transformation and differencing aid in normalizing variance and removing trends, while the Augmented Dickey-Fuller test confirms stationarity, crucial for reliable forecasting. Model evaluation showcases XGBoost's superior performance with the lowest Mean Squared Error (MSE), suggesting the efficacy of advanced machine learning methods in turnover forecasting. However, the ensemble model fails to surpass XGBoost's performance, indicating the need for further refinement. Integrating insights from recent literature, the study emphasizes the importance of advanced modeling techniques, rigorous data preprocessing, and ongoing model optimization for accurate turnover forecasts, thereby facilitating informed decision-making in dynamic market environments.

5.2 Conclusions

In conclusion, this study has achieved its primary objectives of analyzing historical turnover trends, employing time series analysis techniques, and providing actionable recommendations for Simbisa Brands within the Quick Service Restaurant (QSR) industry in Zimbabwe. Through a meticulous examination of turnover data spanning from January 2016 to December 2022, the study has uncovered valuable insights into the performance of Simbisa Brands over this seven-year period. By employing advanced time series analysis techniques, specifically the ARIMA, Random Forest, XGBoost models, the study has elucidated intricate patterns of turnover dynamics. These findings contribute to informed strategic decision-making and provide a deeper understanding of the factors influencing business performance within the QSR industry.

In light of the study's findings and recommendations, it is evident that continual monitoring and refinement are imperative for businesses operating within the QSR industry. As market conditions evolve and consumer preferences shift, it is essential for companies like Simbisa Brands to adapt their strategies accordingly. By embracing advanced analytical methods and incorporating multidimensional analyses, businesses can enhance their resilience and sustainability in an ever-changing market environment. Overall, this study underscores the importance of data-driven decision-making and the value of employing sophisticated analytical techniques to uncover actionable insights for driving business success.

5.3 Recommendations

QSR businesses should conduct regular analysis of turnover data and update their time series models accordingly. This ensures that forecasts remain accurate and relevant, enabling businesses to make informed decisions that align with current market trends.

Stakeholders can use the insights derived from the study to make informed decisions about their business strategies. Understanding historical turnover trends and employing advanced analytical techniques allows stakeholders to anticipate market fluctuations and adjust their operations accordingly.

Governmental bodies responsible for regulating the restaurant industry can use the study's findings to inform policy formulation. Understanding turnover dynamics within the QSR sector allows policymakers to implement targeted policies that support business growth and foster a conducive regulatory environment.

Furthermore, the study's use of the ARIMA model has yielded actionable recommendations for Simbisa Brands. By analyzing historical turnover trends and employing sophisticated analytical techniques, the study has identified opportunities for improving customer engagement, optimizing operations, and enhancing marketing campaigns. These recommendations are grounded in empirical evidence and provide practical guidance for Simbisa Brands to navigate the complexities of the QSR landscape and drive sustained growth and success in the marketplace.

5.4 Area for further research

An area for further research could focus on the integration of external factors, such as economic indicators and socio-cultural trends, into turnover forecasting models for Quick Service Restaurant (QSR) businesses. While this study primarily examined internal factors and historical turnover data, future research could explore how external variables impact turnover dynamics within the QSR industry. By incorporating economic indicators like GDP growth, inflation rates, and consumer spending patterns, researchers can gain a deeper understanding of the broader economic context in which QSR businesses operate. Similarly, examining socio-cultural trends, such as changing dietary preferences and lifestyle habits, can provide valuable insights into shifting consumer behaviors and preferences.

5.5 Chapter Summary

Chapter 5 provides a comprehensive summary, conclusions, and recommendations based on the research conducted on turnover dynamics within Simbisa Brands and the Quick Service Restaurant (QSR) industry in Zimbabwe. The chapter begins by revisiting the research objectives and assessing the extent to which these objectives have been achieved. It highlights the importance of time series analysis in understanding the complex dynamics of customer behavior and business performance within the QSR sector. The chapter emphasizes the potential of the research outcomes to inform strategic decision-making processes and enhance the competitive edge of businesses in the industry. A summary of findings is presented, spanning across each chapter of the study. From the initial exploration of research objectives to the meticulous methodological trajectory, followed by the examination of turnover data and the culmination of advanced analytical modeling, every step of the research journey is recapitulated. The chapter underscores the rigor and empirical depth of the study, showcasing its contribution to advancing the scholarly discourse and understanding of the QSR landscape in Zimbabwe. Conclusions drawn from the study affirm its success in achieving its primary objectives. By analyzing historical turnover trends and employing advanced time series analysis techniques, valuable insights into the performance of Simbisa Brands and the broader QSR industry have been uncovered. The conclusions emphasize the importance of data-driven decision-making and the role of sophisticated analytical techniques in driving business success and resilience in dynamic market environments.

The chapter also offers actionable recommendations for QSR businesses, stakeholders, and governmental bodies. Regular analysis of turnover data, informed decision-making based on empirical evidence, and targeted policy formulation are recommended to navigate the complexities of the QSR landscape effectively. Finally, the chapter identifies areas for further research, suggesting the integration of external factors such as economic indicators and socio-cultural trends into turnover forecasting models. By exploring the impact of these variables on turnover dynamics, researchers can gain deeper insights into the broader economic and social contexts shaping the QSR industry.

REFERENCES

Anderson, E. W., Fornell, C., & Mazvancheryl, S. K. (2016). Customer satisfaction and shareholder value. Journal of Marketing, 80(3), 76-94.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: forecasting and control. John Wiley & Sons.

Box, G.E.P., Jenkins, G.M., Reinsel, G.C., & Ljung, G.M. (2015). Time Series Analysis: Forecasting and Control. 5th ed. Wiley.

Cryer, J. D., & Chan, K.-S. (2008). Time series analysis: with applications in R (2nd ed.). Springer.

Cryer, J.D., & Chan, K.-S. (2008). Time Series Analysis: With Applications in R. 2nd ed. Springer.

Devi, N. S., Sundar, D. G., & Alli, S. V. (2014). An effective time series analysis for stock trend prediction using ARIMA model for Nifty Midcap-50. International Journal of Computer Applications, 91(2), 16-21.

George, E. I., Gwilym, M. J., & Gregory, C. R. (2008). Autoregressive models for time series analysis. Statistical Science, 23(2), 109-133.

Goswami, R., Chakraborty, D., & Jain, R. (2014). Study of effectiveness of time series modeling (ARIMA) in forecasting stock prices. International Journal of Advanced Research in Computer and Communication Engineering, 3(10), 8416-8421.

Ismail, N. A., Abdullah, N. R. H., & Ishak, M. S. A. (2015). Selecting wavelet transforms model in forecasting financial time series data based on ARIMA model. Procedia Computer Science, 72, 97-104.

Li, J., & Zhang, X. (2019). Forecasting tourism demand with ARIMA models. International Journal of Tourism Research, 21(1), 70-84.

Ljung, G.M., & Box, G.E.P. (1978). On a measure of lack of fit in time series models. Biometrika, 65(2), 297–303.

Okafor, F. C., & Shaibu, E. (2013). Time series analysis of inflation dynamics in Nigeria. International Journal of Business and Social Science, 4(1), 68-76.

Pannerselvam, R. (2005). Operations research. Prentice Hall of India.

Petrică, A. C., Stancu, S., & Tindeche, A. (2016). Limitations of ARIMA models in financial and monetary economics. Procedia Economics and Finance, 39, 922-927.

Sekaran, U., & Bougie, R. (2016). *Research Methods for Business: A Skill-Building Approach* (7th ed.). Wiley.

Siami-Namini, S., & Namin, A. S. (2018). Forecasting economics and financial time series: ARIMA vs. LSTM. Journal of Risk and Financial Management, 11(2), 24.

Sivasamy, R., Ramakrishnan, K., & Sabapathy, T. (2014). Application of ARIMA model in stock market forecasting: A study of banking sector. Asian Journal of Research in Banking and Finance, 4(10), 35-46.

Song, X. (2020). Forecasting financial time series with deep learning techniques. IEEE Access, 8, 40119-40133.

Yu, G., & Zhang, C. (2004). ARIMA model based on wavelet decomposition. Journal of Shijiazhuang University of Economics, 27 (1), 18-21.

APPENDIX

```
import pandas as pd
```

```
import numpy as np

import matplotlib.pyplot as plt

from statsmodels.tsa.arima.model import ARIMA

from statsmodels.tools.eval_measures import rmse

from statsmodels.tsa.seasonal import seasonal_decompose

from sklearn.metrics import mean_squared_error, r2_score

from statsmodels.graphics.tsaplots import plot_acf, plot_pacf

from statsmodels.tsa.stattools import adfuller

from statsmodels.graphics.gofplots import qqplot

from sklearn.ensemble import RandomForestRegressor

from xgboost import XGBRegressor

# Load the data into a DataFrame

data = [8516752, 7617241, 8035221, 8108310, 8182470, 7938084, 8735993, 9460814, 8568922,
8696997, 8116457, 11436562, 103413824, 9054485, 8458110, 9392032, 10322061, 10464011,
10504648,11253058, 12015393, 11451855, 12465397, 12647586, 17328953, 135357589,
13009738, 11951501,14438938, 14393719, 14772175, 15507872, 16659408, 17212392,
17068341, 20618419, 23500426,13743965, 192876893, 5497615, 6343305, 7520257, 7564529,
6550512, 5684112, 5242022, 5890388,5799587, 6054307, 6440809, 9246771, 77834213,
9809259, 6804402, 4818924, 1367990, 2569607, 3122533, 5535125, 5753469, 8080007,
10452067, 10881573, 11922015, 81116971, 4705620, 5054137,8957983, 8868428, 10345238,
9693107, 6812966, 7557417, 8715961, 9741428, 10567876, 13071513,104091673, 9527075,
13945596, 15405741, 17165568, 16757069, 16849891, 17120839, 18929960,17233285,
18262004, 19108085, 25389900, 205695014]

# Define the date range

start_date = '2016-01-01'

end_date = '2022-12-01'

dates = pd.date_range(start=start_date, end=end_date, freq='M')

# Create the DataFrame

df = pd.DataFrame(data[:len(dates)], index=dates, columns=['Turnover'])
```

```
# Plot the original time series
plt.figure(figsize=(12, 6))
plt.plot(df.index, df['Turnover'], label='Turnover')
plt.title('Original Turnover Time Series')
plt.xlabel('Date')
plt.ylabel('Turnover')
plt.legend()
plt.grid(True)
plt.show()
# Descriptive statistics
print("Descriptive Statistics for Turnover:")
print(df['Turnover'].describe())
# Data transformation
df['Log_Turnover'] = np.log(df['Turnover'])
df['Diff_Turnover'] = df['Turnover'].diff(periods=1)
df['Seasonal_Diff_Turnover'] = df['Turnover'].diff(periods=12)
# Plot transformed data
plt.figure(figsize=(12, 6))
plt.plot(df.index, df['Turnover'], label='Original Turnover')
plt.plot(df.index, df['Log_Turnover'], label='Log Transformed Turnover')
plt.plot(df.index, df['Diff_Turnover'], label='First Difference Turnover')
plt.plot(df.index, df['Seasonal_Diff_Turnover'], label='Seasonal Difference Turnover')
plt.legend()
plt.title('Transformed Turnover Data')
plt.xlabel('Date')
```

```
plt.ylabel('Turnover')
plt.grid(True)
plt.show()
# ADF Test for stationarity
def adf_test(series):
    result = adfuller(series)
    print(f'ADF Statistic: {result[0]}')
    print(f'p-value: {result[1]}')
    print('Critical Values:')
    for key, value in result[4].items():
        print(f'   {key}: {value}')
adf_test(df['Diff_Turnover'].dropna())
# Plot ACF and PACF
plot_acf(df['Diff_Turnover'].dropna(), lags=30, alpha=0.05)
plot_pacf(df['Diff_Turnover'].dropna(), lags=30, alpha=0.05)
plt.show()
# QQ Plot for normal distribution
qqplot(df['Diff_Turnover'].dropna(), line='s')
plt.title('QQ Plot for Normal Distribution')
plt.show()
# Split data into train and test sets
train_size = int(len(df) * 0.8)
train, test = df[:train_size], df[train_size:]
# ARIMA Model
arima_model = ARIMA(train['Turnover'], order=(0, 1, 2))
```

```
arima_fit = arima_model.fit()
arima_predictions = arima_fit.forecast(steps=len(test))
arima_mse = mean_squared_error(test['Turnover'], arima_predictions)
print(f'ARIMA MSE: {arima_mse}')
# Random Forest Model
rf_model = RandomForestRegressor(n_estimators=100, random_state=0)
rf_model.fit(train.dropna().drop(['Turnover'], axis=1), train.dropna()['Turnover'])
rf_predictions = rf_model.predict(test.dropna().drop(['Turnover'], axis=1))
rf_mse = mean_squared_error(test['Turnover'].dropna(), rf_predictions)
print(f'Random Forest MSE: {rf_mse}')
# XGBoost Model
xgb_model = XGBRegressor(n_estimators=100, random_state=0)
xgb_model.fit(train.dropna().drop(['Turnover'], axis=1), train.dropna()['Turnover'])
xgb_predictions = xgb_model.predict(test.dropna().drop(['Turnover'], axis=1))
xgb_mse = mean_squared_error(test['Turnover'].dropna(), xgb_predictions)
print(f'XGBoost MSE: {xgb_mse}')
# Ensemble predictions
ensemble_predictions = (arima_predictions + rf_predictions + xgb_predictions) / 3.0
ensemble_mse = mean_squared_error(test['Turnover'].dropna(), ensemble_predictions)
print(f'Ensemble MSE: {ensemble_mse}')
# Plot actual vs predicted values for ARIMA
plt.figure(figsize=(10, 6))
plt.plot(test.dropna().index, test.dropna()['Turnover'], label='Actual')
plt.plot(test.dropna().index, arima_predictions, label='ARIMA Predicted')
plt.title('Turnover: Actual vs ARIMA Predicted')
```

```
plt.xlabel('Date')
plt.ylabel('Turnover')
plt.legend()
plt.grid(True)
plt.show()
# Plot actual vs predicted values for Random Forest
plt.figure(figsize=(10, 6))
plt.plot(test.dropna().index, test.dropna()['Turnover'], label='Actual')
plt.plot(test.dropna().index, rf_predictions, label='Random Forest Predicted')
plt.title('Turnover: Actual vs Random Forest Predicted')
plt.xlabel('Date')
plt.ylabel('Turnover')
plt.legend()
plt.grid(True)
plt.show()
# Plot actual vs predicted values for XGBoost
plt.figure(figsize=(10, 6))
plt.plot(test.dropna().index, test.dropna()['Turnover'], label='Actual')
plt.plot(test.dropna().index, xgb_predictions, label='XGBoost Predicted')
plt.title('Turnover: Actual vs XGBoost Predicted')
plt.xlabel('Date')
plt.ylabel('Turnover')
plt.legend()
plt.grid(True)
plt.show()
```

```
# Plot actual vs predicted values for Ensemble
plt.figure(figsize=(10, 6))
plt.plot(test.dropna().index, test.dropna()['Turnover'], label='Actual')
plt.plot(test.dropna().index, ensemble_predictions, label='Ensemble Predicted')
plt.title('Turnover: Actual vs Ensemble Predicted')
plt.xlabel('Date')
plt.ylabel('Turnover')
plt.legend()
plt.grid(True)
plt.show()
# Residual Analysis for ARIMA
arima_residuals = test['Turnover'] - arima_predictions
plt.figure(figsize=(10, 6))
plt.plot(arima_residuals)
plt.title('ARIMA Residuals')
plt.xlabel('Date')
plt.ylabel('Residuals')
plt.grid(True)
plt.show()
# Time Series Cross-Validation for ARIMA
window_size = 12  # 1 year
mse_values = []
for i in range(len(test) - window_size):
    train_window = test.iloc[i:i + window_size]
    test_window = test.iloc[i + window_size]
```

```
    arima_model = ARIMA(train_window['Turnover'], order=(0, 1, 2))
    arima_fit = arima_model.fit()
    arima_pred = arima_fit.forecast(steps=1)
    mse_values.append(mean_squared_error([test_window['Turnover']], arima_pred))
average_mse = np.mean(mse_values)
print(f'Average MSE for ARIMA: {average_mse}')
# Fitting ARIMA models with different orders and storing results
arima_orders = [(p, d, q) for p in range(3) for d in range(2) for q in range(3)]
num_combinations = len(arima_orders)
orders_list = []
aic_list = []
bic_list = []
rmse_list = []
r_squared_list = []
for order in arima_orders:
arima_model = ARIMA(df['Turnover'], order=order)
arima_fit = arima_model.fit()
 orders_list.append(order)
    aic_list.append(arima_fit.aic)
    bic_list.append(arima_fit.bic)
    arima_forecast = arima_fit.forecast(steps=12)
    arima_rmse = rmse(df['Turnover'][-12:], arima_forecast)
    rmse_list.append(arima_rmse)
y_true = df['Turnover'][-12:]
    y_pred = arima_forecast
```

```
    r_squared = r2_score(y_true, y_pred)
    r_squared_list.append(r_squared)
results_df = pd.DataFrame({
    'Order': orders_list,
    'AIC': aic_list,
    'BIC': bic_list,
    'RMSE': rmse_list,
    'R-Squared': r_squared_list})
print(results_df.to_markdown())
# Forecasting for 2023 and 2024
future_dates = pd.date_range(start='2023-01-01', end='2024-12-01', freq='M')
forecast_arima = arima_fit.forecast(steps=24)
forecast_rf = rf_model.predict(np.vstack([df['Log_Turnover'].values[-1],
df['Diff_Turnover'].values[-1], df['Seasonal_Diff_Turnover'].values[-1]]).T)
forecast_xgb = xgb_model.predict(np.vstack([df['Log_Turnover'].values[-1],
df['Diff_Turnover'].values[-1], df['Seasonal_Diff_Turnover'].values[-1]]).T)
forecast_df = pd.DataFrame({
    'ARIMA Forecast': forecast_arima,
    'Random Forest Forecast': forecast_rf,
    'XGBoost Forecast': forecast_xgb}, index=future_dates)
plt.figure(figsize=(10, 6))
plt.plot(df.index, df['Turnover'], label='Historical Turnover')
plt.plot(forecast_df.index, forecast_df['ARIMA Forecast'], label='ARIMA Forecast')
plt.plot(forecast_df.index, forecast_df['Random Forest Forecast'], label='Random Forest
Forecast')
plt.plot(forecast_df.index, forecast_df['XGBoost Forecast'], label='XGBoost Forecast')
```

```
plt.xlabel('Date')
plt.ylabel('Turnover')
plt.title('Forecasted Turnover for 2023 and 2024')
plt.legend()
plt.grid(True)
plt.show()
# Plotting the decomposition components
result = seasonal_decompose(df['Turnover'], model='additive', period=12)
plt.figure(figsize=(12, 8))
plt.subplot(4, 1, 1)
plt.plot(df.index, result.observed)
plt.title('Observed')
plt.subplot(4, 1, 2)
plt.plot(df.index, result.trend)
plt.title('Trend')
plt.subplot(4, 1, 3)
plt.plot(df.index, result.seasonal)
plt.title('Seasonal')
plt.subplot(4, 1, 4)
plt.plot(df.index, result.resid)
plt.title('Residual')

plt.tight_layout()
plt.show()
# Plotting actual vs forecasted values for 2023 and 2024
```

```
plt.figure(figsize=(10, 6))

plt.plot(df.index, df['Turnover'], label='Historical Turnover')

plt.plot(forecast_df.index, forecast_df['ARIMA Forecast'], label='ARIMA Forecast')

plt.plot(forecast_df.index, forecast_df['Random Forest Forecast'], label='Random Forest
Forecast')

plt.plot(forecast_df.index, forecast_df['XGBoost Forecast'], label='XGBoost Forecast')

plt.xlabel('Date')

plt.ylabel('Turnover')

plt.title('Actual vs Forecasted Turnover for 2023 and 2024')

plt.legend()

plt.grid(True)

plt.show()
```