Demand Forecasting of a Perishable Dairy Drink: An ARIMA Approach

T. Musora, Z. Chazuka, A. Jaison, J. Mapurisa, and J. Kamusha

School of Natural Sciences and Mathematics,Department of Mathematics,
Chinhoyi University of Technology ,
Pivate Bag-7724, Chinhoyi, Zimbabwe

Abstract. Any organization engaged in trading aims to maximize earnings while maintaining costs at their bare minimum. One of the inexpensive ways to accomplish this objective is through sales forecasting. Evidence from empirical literature has shown that sales forecasting frequently results in better customer service, fewer returns of goods, less dead stock, and effective production scheduling. Successful sales forecasting systems are essential for the food sector because of the limited shelf life of food goods and the significance of product quality. In this paper, we generated sales of forecasts for a perishable dairy drink using the famous ARIMA approach. We identified the ARIMA $(0,1,1)(0,1,1)_{12}$ as the proper model for modeling the daily sales forecast of the perishable drink. After performing model diagnostics, the model satisfied all the model assumptions, and a strong positive linear relationship ($R^2 > 0.9$) was observed when the actual daily sales were regressed against the forecasted values.

Keywords: Sales forecasting, ARIMA, Model Diagnostics, R^2- value.

1 Introduction

One of the critical ingredients in decision-making and business planning is forecasting. Demand forecasting is the process of predicting the most likely demand for a product in the future. [1] defines demand forecasting as a process of predicting expected demand, supply, and pricing for products based on historical data. Literature has shown that the main objective of forecasting is to minimize risk in decision-making. Some scholars argue that forecasting is the starting point in planning. According to [2], the success of a business depends on getting the correct forecasts. The author goes on to state that forecasting techniques can be split into two broad categories: qualitative and quantitative. Qualitative forecasting is mainly based on personal experience and opinion, while quantitative forecasting heavily relies on historical data.

Recent advancements in technology have seen authors like [3] using artificial intelligence to forecast the demand for horticultural products and concluding that machine learning outperforms classical forecasting on horticultural sales. However, classical forecasting methods such as the Auto-regressive Integrated Moving Average (ARIMA) and Exponential Smoothing (ES) are still heavily relied on in industry and academia as they form a baseline for these new methodologies. Despite their simple design, classical techniques have frequently demonstrated competitive performance.([4],[5], [6]).

[7] states that reliable forecasts are essential for a company to survive and grow. In a manufacturing environment, management must forecast the future demands for its products and provide the materials, labor, and capacity to fulfill these needs. These resources are

DOI:10.5121/ijdkp.2023.13203

secured before the customers place demands for the products on the firm. Forecasting is the heart and blood of any inventory control system. A firm with hundreds or thousands of items must anticipate the demand for its products in advance to match what is required and what it can supply to avoid the costs associated with overstocking or understocking. Management must plan several months for its inventory because of procurement lead times from suppliers. Hence, forecasts play a pivotal role in planning for inventory months ahead. Consequently, forecasts are used to assess whether or not to place an order with the supplier and the size of the order.

The arguments above show that any distribution company aims to determine an optimal supply of its products that minimizes costs and maximizes returns in the face of uncertainty. This paper is primarily concerned with quantitative forecasting using time series analysis. In this context, the time series represents the demands recorded over past time intervals. The forecasts are estimates of what is required over future time intervals and are generated using the flow of demands from the past. This paper proceeds as follows. Section 2 gives the literature review, some theoretical structures for exponential smoothing models, and autoregressive integrated moving average (ARIMA) models. Section 3 includes comprehensive empirical results and an analysis of orange drink circulation and results. Section 4 is the discussion and conclusion

2 Time Series Analysis and Modelling Strategy

Predicting future values of a time series is of great significance in multiple disciplines. Trend, cyclical, seasonal, and random components naturally appear in economic and commercial time series. Methods have been developed to capture these components by specifying and estimating statistical models. These methods comprise; log transformation, square root transformation exponential smoothing, and ARIMA, which are described by [9] and [10]. The authors demonstrate that, while requiring substantially more work, ARIMA provides, on average, more accurate out-of-sample forecasts than other smoothing approaches.

According to [11], Robert G. Brown developed exponential smoothing while working as an OR analyst for the US Navy in World War Two. [12] identifies that the more sophisticated exponential smoothing methods seek to isolate trends or seasonality from irregular variation. Where such patterns are found, the more advanced methods identify and model these patterns. The models can then incorporate those patterns into the forecast. Exponential smoothing uses weighted averages of past observations for forecasting. The effect of past observations is expected to decline exponentially over time[13] states that the exponential smoothing methods are relatively simple but robust approaches to forecasting. They are frequently used in business for forecasting demand for inventories. Three basic variations of exponential smoothing are simple exponential smoothing, trend-corrected exponential smoothing, and the Holt-Winters method. [14] states that the ARIMA method developed by [15] is one of the most noted models for time series data prediction and is often used in econometric research.

The ARIMA method stems from the autoregressive (AR) model and the moving average (MA) model, and the combination of the two gives the ARMA model. Evaluated against the early AR, MA, and ARMA models, the ARIMA model is more flexible in application and more accurate in the quality of the simulative or predictive results. [**?**] points out that in the ARIMA analysis, an identified underlying process is generated based on observations to a time series for generating a good model which shows the process-generating mechanism precisely. [17] and [18] states that the only problem with ARIMA modeling is that it is sophisticated in theory and requires a good understanding of mathematics. In other words, building an ARIMA model is not a trivial task as it needs training in statistical analysis, a good grasp of the field of application, and the availability of an easy-to-use but versatile specialized computer program. The Box-Jenkins method for modeling and forecasting time series data is amongst the large family of quantitative forecasting approaches established in the fields of operations research, statistics, and management science. Box-Jenkins models are also known as ”ARIMA” models, the acronym for Autoregressive Integrated Moving Average. This terminology is made clear in the following sections.

Exponential smoothing, linear regression, Bayesian forecasting, and generalized adaptive filtering are some of the other techniques which are termed ”extrapolative” forecasting [6] All the methods mentioned above have common elements, the first being they use historical data to try and explain what might occur in the future. Secondly, they use a single variable to predict the future values of the same variable, and they are referred to as univariate models.

2.1 ARIMA Model

The ARIMA model is an extension of the ARMA modelling the sense that by including auto-regression and moving average it has an extra function for differencing the time series. If a dataset exhibits long-term variations such as trends, seasonality and cyclic components, differencing a dataset in ARIMA allows the model to deal with them. Two common processes of ARIMA for identifying patterns in time-series data and forecasting are auto-regression and moving average.

2.2 Autoregressive Process

Most time series consist of elements that are serially dependent in the sense that one can estimate a coefficient or a set of coefficients that describe consecutive elements of the series from specific, time-lagged (previous) elements. Each observation of the time series is made up of random error components (random shock; a_t) and a linear combination of prior observations.

2.3 Moving Average Process

Independent from the autoregressive process, each element in the series can also be affected by the past errors (or random shock) that cannot be accounted for by the autoregressive component. Each observation of the time series is made up of a random error component (random shock, ϵ) and a linear combination of prior random shocks.

2.4 Autoregressive Integrated Moving Average Process, ARIMA (p, d, q)

A series X_t is called an autoregressive integrated moving average process of orders p, d, q, ARIMA(p, d, q), if $W_t = \nabla^d X_t$, where W_t is the differenced time series.

We may define the difference operator ∇ as $\nabla X_t = X_t - X_{t-1}$. Differencing a time series $\{X_t\}$ of length n produces a new time series $\{W_t\} = \left\{\nabla^d X_t\right\}$ of length $n-d$. If $\{Z_t\}$ is a purely random process with mean zero and variance σ_z^2, the general autoregressive integrated moving average process is of the form

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \ldots + \phi_p W_{t-p} + Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$$

In terms of the backward shift operator, the ARIMA(p, d, q) process is

$$\mathit{\Phi}_p(B) W_t = \mathit{\Theta}_q(B) Z_t$$

Remark: The autoregressive integrated moving average process is specifically for non-stationary time series. The differencing transformation is useful in reducing a non-stationary time series to a stationary one.

2.5 Seasonal Auto-regressive Integrated Moving Average Process.

Let s, be the number of observations per season. Then the time series, X_t, is called a seasonal autoregressive integrated moving average process of orders p, d, q, seasonal orders P, D, Q and seasonal period s, if it satisfies;

$$\phi_p(B) \mathit{\Phi}_P\left(B^s\right) \nabla^d \nabla_s^D X_t = \theta_q(B) \mathit{\Theta}_Q\left(B^s\right) Z_t$$

Where $\nabla_s^D X_t = \sum_{j=0}^{D} \binom{D}{j} X_{t-js}$, and $\phi_p(B)$ and $\theta_q(B)$ are polynomials in B of order p and q, that is ;

$$\begin{aligned} \phi_q(B) &= \left(1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q\right) \\ \theta_p(B) &= 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p \end{aligned}$$

We identified the stationary component of a data set by performing the Ljung and Box test. We tested this hypothesis by choosing a level of significance for the model adequacy and compared the computed Chi-square $\left(\chi^2\right)$ values with the $\left(\chi^2\right)$ values obtained from the table. If the calculated value is less than the actual $\left(\chi^2\right)$ value, then the model is adequate, otherwise not. The $Q(r)$ statistic is calculated by thefollowing formula:

$$Q_(r)) = n(n+2) \sum \frac{r^2(j)}{n-j}$$

where n is the number of observations in the series and $r(j)$ is the estimated correlation at lag j. Furthermore, we tested the data to specify the order of the regular and seasonal autoregressive and moving average polynomials necessary to adequately represent the time series model. For this purpose, model parameters were estimated using a maximum likelihood algorithm that minimized the sums of squared residuals and maximized the likelihood (probability) of the observed series. The maximum likelihood estimation is generally the preferred least square technique. The major tools used in the identification phase are plots of the series, correlograms (plots of autocorrelation and partial autocorrelation verses lag) of the autocorrelation function (ACF) and the partial autocorrelation function (PACF).The ACF and the PACF are the most important elements of time series analysis and forecasting. The ACF measures the amount of linear dependence between observations in a time series that are separated by a lag k. The PACF plot helps to determine how many autoregressive terms are necessary to reveal one or more of the following characteristics:time lags where high correlations appear, seasonality of the series, and trend either in the mean level or in the variance of the series. In diagnostic checking, the residuals from the fitted model were examined against their adequacy. This is usually done by correlation analysis through the residual ACF plots and by goodness-of-fit test using means of Chi-square statistics. At the forecasting stage, the estimated parameters were used to calculate new values of the time series with their confidence intervals for the predicted values.

2.6 Performance valuation

To choose the best model among the class of plausible model, the estimated parameters were tested for their validity using, ACF , PACF, Probability Plot and Histogram of residuals, a time series plot of observed and fitted values and other error statistics such as coefficient of determination(R^2) were analysed.

2.7 Data Source

The data used in this research is historical data of monthly sales of cases of the perishable drink from a small drink manufacturing company in Harare, Zimbabwe which among other products manufactures the perishable dairy drink. Each case contains 24 bottles of the drink. The company intends to minimise losses due to returns of the drinks as result of reduced shelf life.

3 Results and Analysis

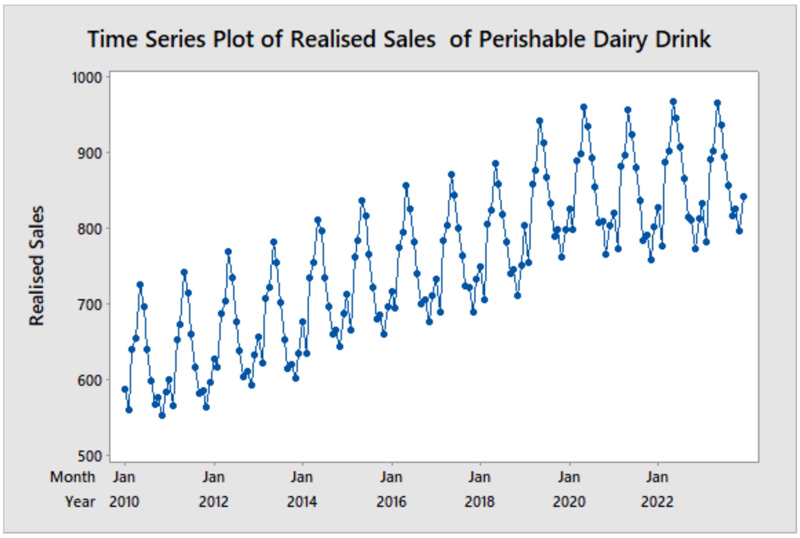

Fig. 1: Time Series Plot Of Demand of Perishable Dairy Drink- Original Data

Visual inspection of the plot shows that the series is dynamic. So need is there to transform the data so as to make it stationary.

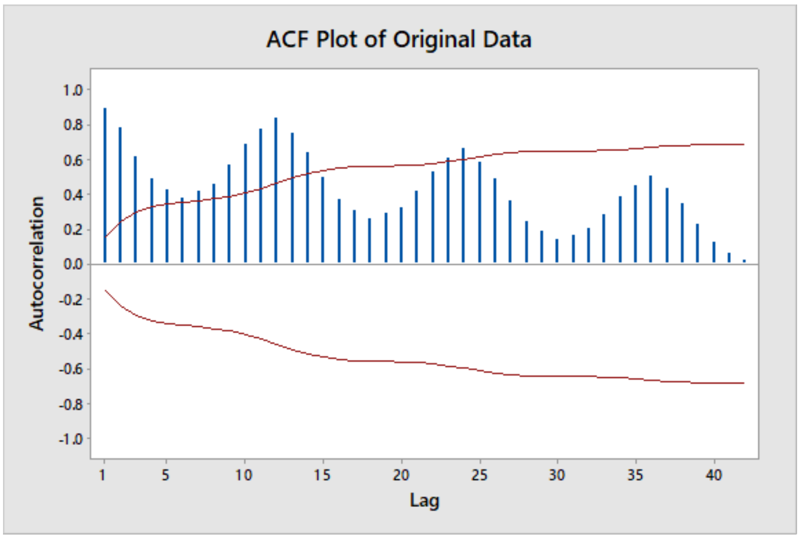

Fig. 2: ACF Plot of Original Data

ACF of most lags are very high, there is evidence of positive and negative autocorrelation. This is a typical ACF plot of a non stationery time series. Thus a model cannot be fitted at this stage. This further affirms need to transform the data.

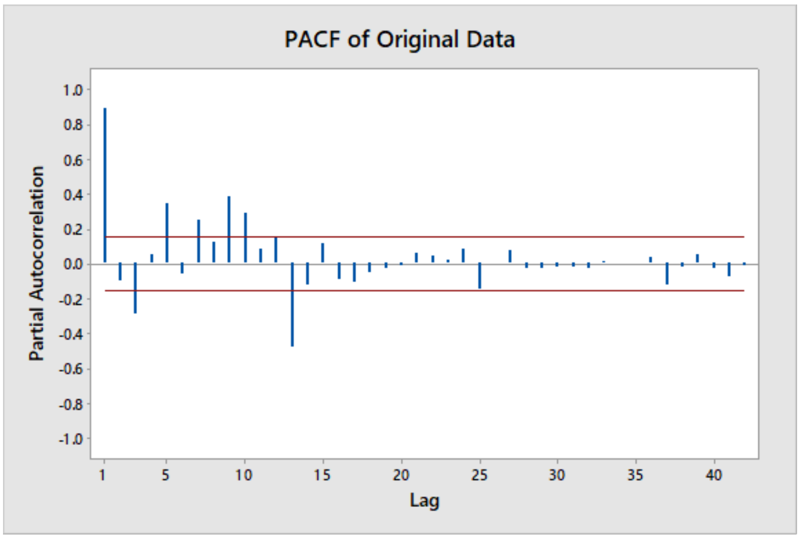

Fig. 3: PACF plot of Original Data

The PACF plot shows a number of significant spikes, which is typical of a non stationary series. Thus we have to transform the data to make it stationary.

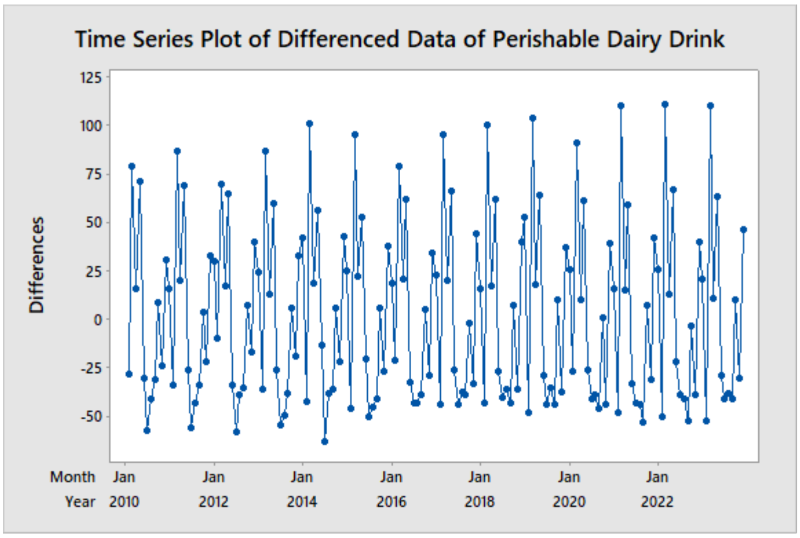

Fig. 4: Time series plot of Differenced Data of Perishable Dairy Drink

Visual inspection of the plot reveals that the differenced series fluctuates around zero, thus the data is now stationary

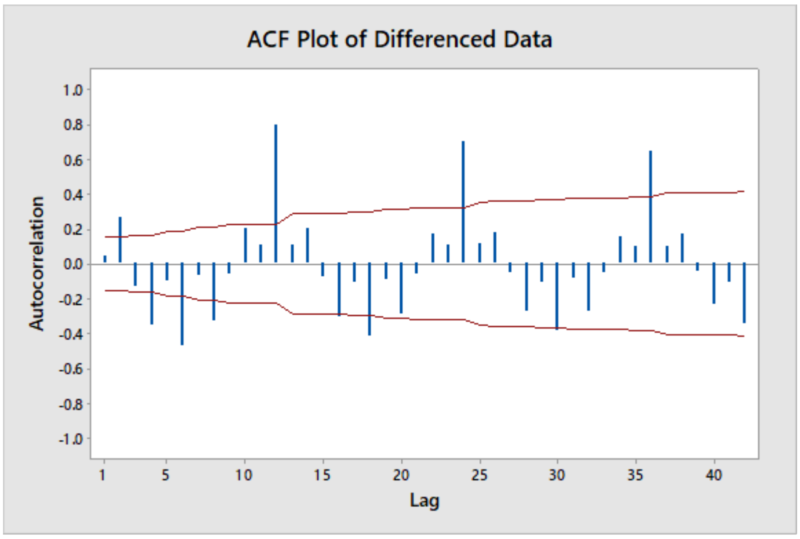

Fig. 5: ACF plot of Differenced Data

The ACF shows a significant spike at lag 2 and there is evidence of negative dumped oscillations with the rest of the ACF’s essentially zero, hence a seasonal ARIMA model is suggested

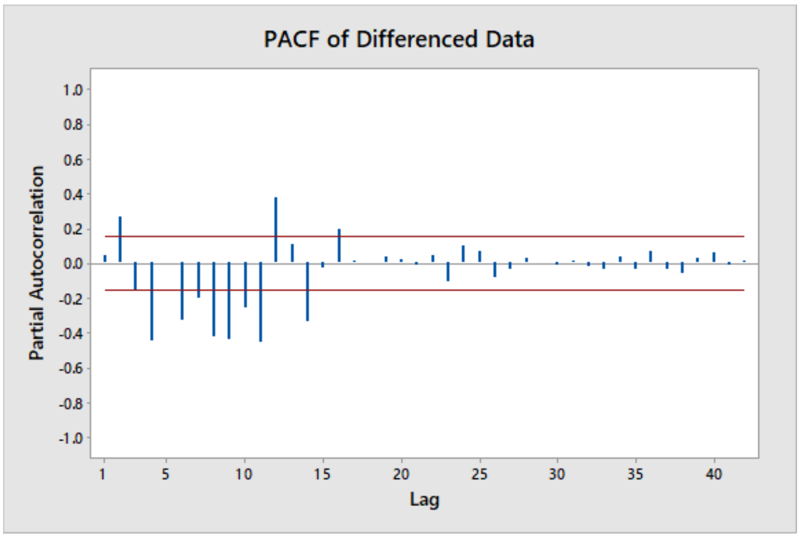

Fig. 6: PACF plot of Differenced Data

PACF plot shows a significant spike at lag 2 which is seasonal and there is evidence of negative dumped oscillations with the rest of the PACFs essentially zero, hence a seasonal ARIMA model is also suggested.

3.1 Parameter Estimation

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
MA 1	0.9707	0.0321	30.21	0.000
SMA 12	0.6533	0.0660	9.90	0.000
Constant	-0.00181	0.01501	-0.12	0.904

Lag	12	24	36	48
Chi-Square	99.2	232.2	326.0	405.3
DF	9	21	33	45
P-Value	0.000	0.000	0.000	0.000

Thus the fitted model is SARIMA$(0, 1, 1)(0, 1, 1)_{12}$

3.2 Model Diagnostics

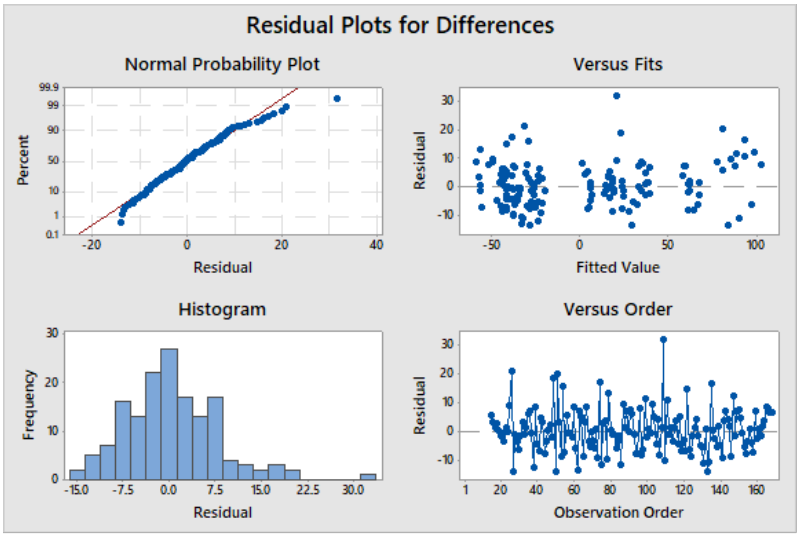

Fig. 7: Residual Plot for Differences

The normal probability plot is almost a straight line, an indication that the normality assumption has not been violated. A plot of residuals against fitted values shows no pattern and the histogram of residuals also indicates that the normality assumption has not been violated. Hence the fitted model is good and thus can be used for forecasting.

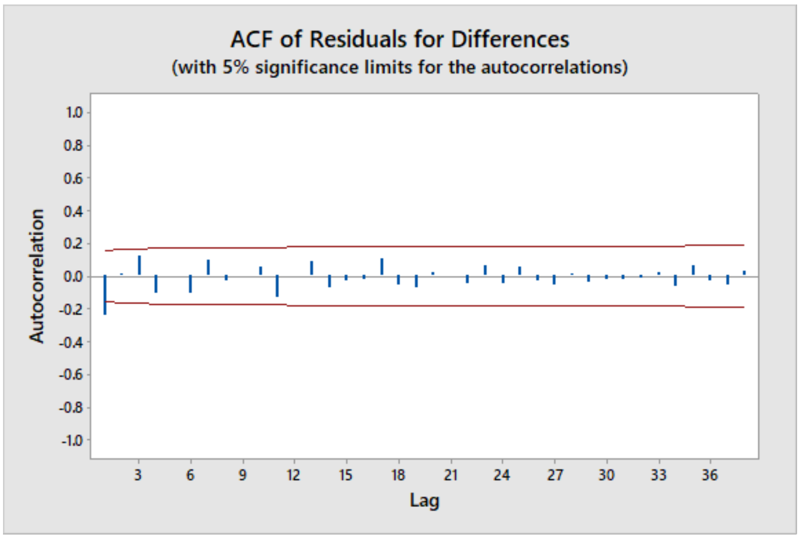

Fig. 8: ACF of residuals for Differences

Figure 8 ACF plot has no significant spikes suggesting that there might be no possible additional parameters which may have been omitted in this model.

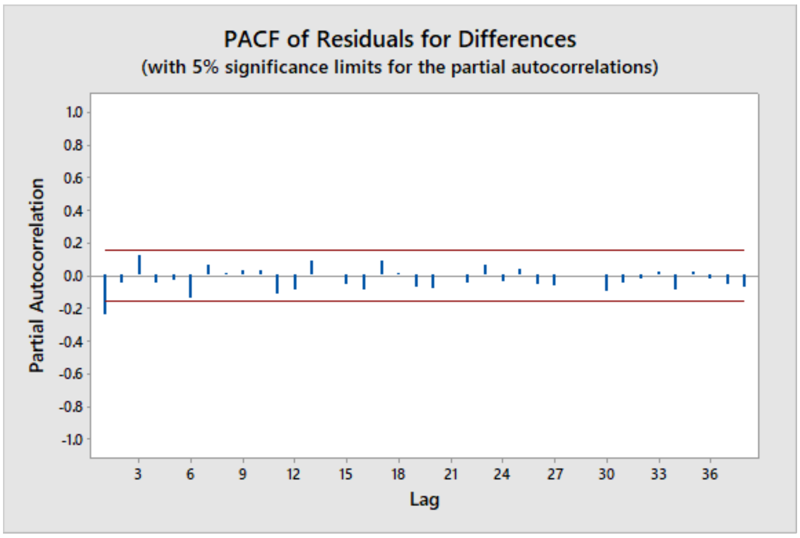

Fig. 9: PACF of residuals for Differences

Figure 9, The PACF plot of residuals refuses any significant spikes suggesting that there might be no possible additional parameters that may have been omitted in this model. Since the fitted model appears good enough, it can be used for forecasting future demand of the perishable dairy drink.

3.3 Inference Based on the Model

3.3.1 Forecasts from period 159

Period	Forecast	Lower	Upper	Actual
159	847.49	783.94	911.04	892.00
160	836.74	752.93	920.55	903.00
161	896.10	795.37	996.84	966.00
162	897.13	782.01	1012.25	937.00
163	881.36	753.45	1009.26	896.00
164	869.28	729.77	1008.80	858.00
165	831.16	683.39	983.87	817.00
166	808.89	670.91	991.40	827.00
167	826.39	639.32	978.55	797.00
168	828.86	639.23	1013.47	
169	830.30	609.70	1024.07	
170	831.37	591.24	1050.89	
171	833.43	573.95	1072.51	
172	834.99	557.99	1092.21	
173	836.55	543.07	1111.99	
174	838.11	529.03	1130.03	
175	839.67	515.73	1147.19	
176	841.23	503.09	1163.60	
177	842.79	491.03	1179.36	
178	842.79	479.47	1194.55	
179	844.35	487.93	1209.23	
180	845.91	468.36	1223.45	
181	847.47	457.67	1236.26	

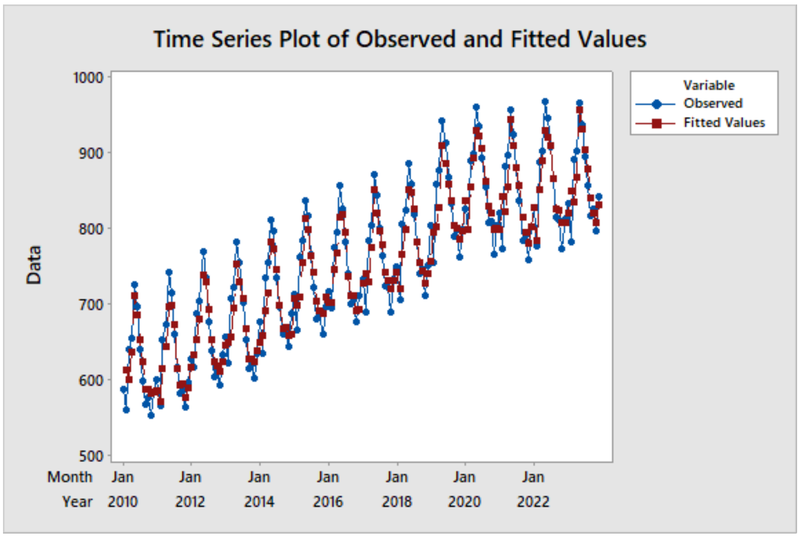

Fig. 10: Time Series plot of Observed and Fitted Values

The fitted values compare well with the observed values, thus the fitted model is reliable.

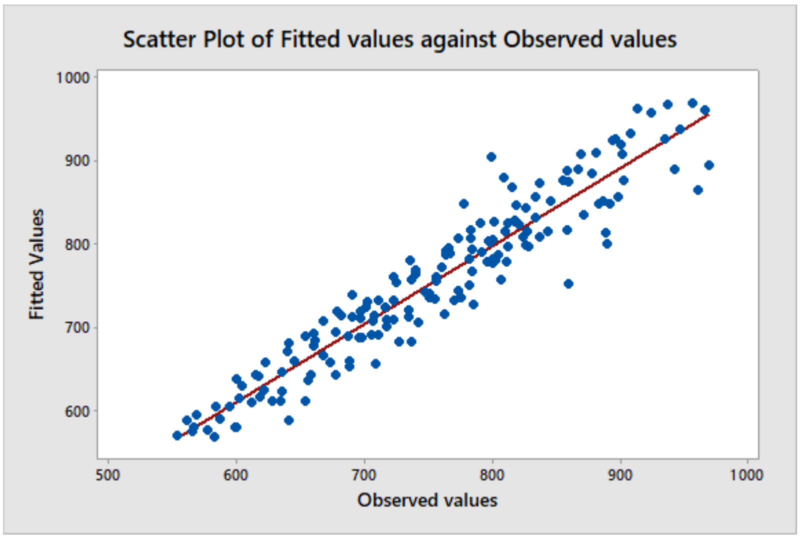

Fig. 11: Scatter plot of Fitted values against Observed Values

3.4 Regression Analysis: Fitted Values versus Sales

The scatter plot of fitted values against observed values suggests a positive linear relationship.

Method

Rows unused 1

Analysis of variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	1	150339	153391	1500.63	0.000
Error	165	165303	1002		
Total	166	1668694			

Model Summary

Analysis of variance Coefficients

R - Sq	R-sq(adj)	R-sq(pred)
90.09%	90.03%	89.84%

Regression Equation

$$\text{Fitted Values} = 48.4 + 0.9360 \times \text{Sales}$$

Term	**Sales**	**Se Coef**	**T-Value**	P-Value
Constant	48.4	18.4	2.62	0.009
Sales	0.9360	0.0242	38.74	0.000

The coefficient of determination value is 90.09% indicates that the fitted model accounts for about 91% of the variation in the fitted values. Thus the fitted seasonal ARIMA model which generated the fitted values must me appropriate and hence can be used to forecast sales values.

4 Discussions and Conclusions

This study demonstrates how ARIMA and Regression models are used to study and forecast sales for a particular company. This paper illustrates how the ARIMA methodology can be used to construct a model for sales forecasting. The $ARIMA(0,1,1)(0,1,1)_{12}$ predicted the data considerably well and gave reliable forecasts. With reference to the data presented, this model was the most appropriate in forecasting the demand, but could not tell why the sales will contain outliers. The Time Series forecasting system helped construct a model, the ARIMA time series, and the Regression, which is effective for forecasting and can be applied to other businesses to plan their sales. However, it would be interesting to do further research on the factors that influence the sales, such as the growth of the population of consumers, the industrial growth in the region, immigration, and so on; this would consolidate better this company's planning. forecasting.

References

1. Islek, I.; Oguducu, S, 2017. A Decision Support System for Demand Forecasting based on Classifier Ensemble. In: Communication Papers of the 2017 Federated Conference on Computer Science and Information Systems, M. Ganzha, L. Maciaszek, M. Paprzycki (eds). ACSIS, 13: 2017, p. 35–41.
2. Shim, J.K. (2009). Strategic Business Forecasting: Including Business Forecasting Tools and Applications, Global Professional Publishing.
3. Haselbeck, F. Killinger, J, Menard, K, Hannus T and Grim , G.D.(2020) Machine Learning Outperforms Classical Forecasting on Horticultural Sales Predictions
4. Kolassa, S. (2021). Commentary on the M5 forecasting competition. International Journal of Forecasting, http://dx.doi.org/10.1016/j.ijforecast.2021.08.006, Advance online publication
5. Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 competition: 100, 000 time series and 61 forecasting methods. International Journal of Forecasting, 36(1), 54–74. http://dx.doi.org/10.1016/j.ijforecast.2019.04.014.
6. Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2021). The M5 competition: Background, organization, and implementation. International Journal of Forecasting, http://dx.doi.org/10.1016/j.ijforecast.2021.07.007, Advance online publication
7. Thomopoulos, N.T. (1980). Applied Forecasting Methods, Prentice Hall.
8. Wei, W.W.S. (2006). Time Series Analysis: Univariate and Multivariate Methods, Second edition, Addison Wesley.
9. Granger, C.W.J. and Newbold, P. (1974). Spurious regressions in econometrics, Journal of Econometrics, 2, pp. 111–120.
10. Reid, D.J. (1975). A review of short-term projection techniques, Practical Aspects of Forecasting, H. A. Gordon, ed., London: Operational Research Society.
11. Gass, S.I. and Harris, C.M. (2000). Encyclopedia of operations research and management science, Centennial edition, Dordrecht, The Netherlands: Kluwer.
12. Yafee, R. and McGee, M. (2000). Introduction to Time series Analysis and Forecasting with Application of SAS and SPSS, Academic Press.

13. Gardner, E.S. (1985).Exponential smoothing: The state of the art, Journal of Forecasting, 4, pp. 1–28.

14. Ediger, V.S., Akar, S. and Ugurlu, B. (2006). Forecasting production of fossil fuel sources in Turkey using a comparative regression and ARIMA model, Energy Policy, 34, pp. 3836–3846.

15. Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control, Revised ed. Holden-Day, San Francisco, USA.

16. Gardner, E.S. (2006). Exponential smoothing: The state of the art – Part II, International Journal of Forecasting, 22, pp. 637–666.

17. Ho, S.L. and Xie, M. (1998). The use of ARIMA models for reliability forecasting and analysis, Computers Industrial Engineering, 35, pp. 213–216.

18. Melard, G. and Pasteels, J.M. (2000). Automatic ARIMA modeling including interventions, using time series expert software, International Journal of Forecasting, 16, pp. 497–508.

19. Billah, B., King, M.L., Snyder, R.D. and Koehler, A.D. (2006). Exponential smoothing model selection for forecasting, International Journal of Forecasting, 22, pp. 239–247.

20. Cho, S.H. and Song, I. (1996). A Formula for Computing the Autocorrelations of the AR Process, The Journal of the Acoustical Society of Korea, 15, 4–7.

21. Diebold, F.X. (2001). Elements of Forecasting, Second edition, Thomson Learning.

22. Enders, W. (1995). Applied Econometric Time Series, John Wiley and Sons, New York.

23. Eshel, G. (2003). The Yule Walker Equations for the AR Coefficients, Citeulikearticle- id: 763363.

24. Fuller, W.A. (1996). Introduction to Statistical Time Series, Second edition, John Wiley and Sons, New York.

Authors

T Musora Is a PhD student at Chinhoyi University of Technology, He is currently being supervised by Dr. F Matarise from the university of Zimbabwe.He holds a MSc in Operations Research from NUST, Zimbabwe, B.Sc., (Special Hons)in Operations Research & Statistics from NUST, Zimbabwe, BSc., Mathematics & Statistics with Zimbabwe Open University, Zimbabwe and has a Dip. Ed.from Gweru Teachers College, Zimbabwe. Mr Musora is also a lecturer in the Department of Mathematics at Chinhoyi University of Technology.

J Kamusha received a Msc In Mathematical Sciences from Stellenbosch University. He holds a Bachelor's degree in Mathematics majoring in Actuarial Science . Currently, he is Lecturer at Chinhoyi University of Technology. His research interests is in Graph Theory and Convolutional Neural Networks .

J Mapurisa received a MSc In Mathematics from University of Zimbabwe. He holds a B.Sc in Mathematics from the University of Zimbabwe . Currently, he is Lecturer at Chinhoyi University of Technology. His research interests is in Fluid Dynamics (flow in channels) .

A Jaison is a Lecturer in the Department of Mathematics at Chinhoyi University of Technology. He holds a MSc in Operations Research from NUST, Zimbabwe, B.Sc. Honours in Applied Mathematics from NUST, Zimbabwe. His research interest is in Multivariate Analysis, Financial and Statistical Modeling

Dr. Z Chazuka Holds a PhD in Mathematical Biology from University of South Africa. She also holds an MSc In operations Research from the National University of Science of Technology Zimbabwe, and a B.Sc Applied Mathematics from the National University of Science of Technology . Currently, she is lecturer at Chinhoyi University of Technology. Her research interests is in Mathematical Biology