Time Series Analysis: An Application of SARIMA Model in General Trade to Forecast Sales

[1]Ankita Vaswani
PGDM Scholar

[2]Dr. Poorna Chandra Prasad
Associate Professor- Marketing

[3]Dr. Prafulla Kumar Padhi
Assistant Professor- Marketing

[1]Pune Institute of Business Management (PIBM), Pune, India
Gut No. 605/1, Lavasa Road, Pirangut, Mulsi, Paud Road, Pune-412115, India

[2]Pune Institute of Business Management (PIBM), Pune, India

[3]Pune Institute of Business Management (PIBM), Pune, India

[1]P21d04ankita@pibm.net.in, [2]poornachandra.p@pibm.in, [3]prafulla.padhi@pibm.in

Abstracts:

General Trade is an essential piece of the FMCG business, as it assists with guaranteeing that items are accessible to customers in an ideal and productive way. Time Series (SARIMA) model is a significant model to help for forecasting the sales of Wholesale trade to Retail trade. To study the SARIMA model Time Series Approach, which helps to forecast the sales of Wholesale trade to Retail trade. A large data from a particular city to be prejudicated in the analysis of ARIMA, SARIMA models to make the analysis accurate abide in similar research. 48 data and 132 data are fitted to different sales forecasting models, whereas the SARIMA model is found best suitable model to forecast the sales forecasting of retail store sales by the wholesalers as well as the retailers also. The future scope of study aligns with the quality-seasonal effect on demand of retail trade.

Keywords: Time Series, SARIMA, seasonality, Wholesale trade, Retail Trade, SPSS, R, Python

1. Introduction

General Trade is an essential piece of the FMCG business, as it assists with guaranteeing that items are accessible to customers in an ideal and productive way. General exchange is around 90% of the general retail market, and incorporates regular Kirana shops, corner shops, general stores, mother and pop stores, container beedi shops and other little retail outlets arranged close to neighborhoods. With the section of current exchange things have extraordinarily changed, its portion of the market is reliably expanding. Huge foundation advancement, infectious insides, innovative turn of events, normalization, advertising and advancement has generally helped support its smooth infiltration. Be that as it may, the general exchange frameworks have continued as before. There has been no huge mechanical advancement in this field over beyond fifty years. Accordingly, there has been stagnation in the general exchange organization of India. Today, the producers and wholesalers are battling to stay up with the quick changing shopper requests. They are taking on different conflicts: On one side, they are battling Present day exchange while on the opposite side, they have Web based business organizations, that poor person recently made shopping helpful yet have likewise placed the best grouping of items before the clients. To restore and ad lib Broad Exchange, it expects to rebuild the framework with innovation support. Stock arranging is a prominent piece of dissemination divert in FMCG industry. Appropriate stock administration assists with adjusting market interest, depends vigorously on a precise gauge of future interest. Prescience of future deals is called as deals determining it assists in business with understanding future patterns and conduct it assists business with designating assets in a right amount brilliantly. What's more, to further develop stock arranging it additionally expect to comprehend the irregularity factor, Irregularity determining supports surveying the patterns and examples in the client's deals lifecycle, which assists with seeing precisely when to make explicit moves like purchasing stock, pushing market drives and running advancements.

From the above conversations, it is perfectly clear that deals estimating is really significant by and by and significantly to comprehend irregularity factor. The central goal of this study was to develop a model by which the we can ready to estimate deals involving verifiable information as a source of perspective through the utilization of Time series Approach and comprehend the way of behaving, patterns and example exposed to irregularity utilizing SARIMA model. There were different models additionally analyzed and assessed named Direct relapse, Strategic relapse, weighted normal model (up to 4 attentions) and SARIMA model.

2. Literature Review

2.1 Use of SARIMA and ARIMA model in Different Area

(Kumar & Anand, 2015) had used the time series ARIMA forecasting model for predicting sugarcane production in India. The effort is made to forecast sugarcane production for next 5 years, secondary data collected from DAC from 1950 to 2013 and developed ARIMA (2,1,0) model. The ARIMA (2,1,0) model predicted an increase in the production for year 2013, then a fall in 2014 and in subsequent years up to 2017, an overall increase in production (Table 5). The prediction for 2013 is resulted approximately 350 million tons (±6% at confidence interval 80%,

±9% at confidence interval 95% and ±13% at confidence interval 99.5%) and for 2014, the prediction is approximately 322 million tons (±11% at confidence interval 80%, ±17% at confidence interval 95% and ±24% at confidence interval 99.5%). Although, like any other predictive models in forecasting, ARIMA also has limitations on the accuracy of predictions yet it is used more widely for forecasting the future successive values in the time series.

(Wang et al., 2018) had used the time-series analysis of tuberculosis from 2005 to 2017 in China. SARIMA model and a hybrid model of SARIMA- GRNN (Generalized regression neural network model was used. Decreasing trend and seasonal variation were identified from 2005-2017 data in China, with a predominant peak in spring. A SARIMA model of ARIMA (0,1,1)(0,1,1)12 was identified. The mean error rate of the single SARIMA model and the SARIMA–GRNN combination model was 6.07% between 2.56%, and the determination coefficient was 0.73 between 0.94, respectively.

(Kam et al., 2010) had predicted the daily patient numbers for a regional emergency medical center by using time series analysis. The predicted number of the daily number of patients visiting the Emergency Department (ED) of a Korean hospital using 3 model average; 2) univariate seasonal auto-regressive integrated moving average (SARIMA); and 3) multivariate SARIMA, from above all three models the multivariate SARIMA found more appropriate and accurate. In this study, the number of patients visiting the ED daily was employed as a dependent variable, and the calendrical and meteorological information is utilized as independent variables for the construction of the forecasting model. In this study, the number of patients visiting the ED rises on Sundays and public holidays such as Seollal and Chuseok. This is because outpatient treatments are all closed on holidays, which means that the ED does double duty.

(Adanacioglu & Yercan, 2012) had analyzed the price of tomato at the wholesale level in Turkey: an application of SARIMA model. The objective of the study was to forecast the monthly tomato prices at wholesale level in Antalya, a city located in the Mediterranean Region, Turkey, on the basis of reported prices from 2000 to 2010. The highest tomato prices adjusted seasonally appear in October. SARIMA (1, 0, 0) (1, 1, 1)12 model was selected as the most suitable model to forecast of tomato prices.

(Milenković et al., 2018) had used the SARIMA modelling approach for railway passenger flow forecasting. The monthly data collected from Jan 2004- Jan 2014 to find out the seasonality of monthly flow of passengers. The results of models indicate that the SARIMA (0,1,0) × (0,1,1)12 is the most appropriate for modeling the rail passenger demand on Serbian railways as it shows positive results. (Arunraj et al., 2014) had used SARIMA model and observed the seasonality differences. The Seasonal Autoregressive Integrated Moving Average (SARIMA) and autoregressive integrated moving average (ARIMA) model is used in the field of patient observation, railway reservation, production process and wholesale trade in the various countries.

2.2. Use of SARIMA and ARIMA model in Business Application

(Pacco, 2022) has used the time series method for the simulation of the temperature control and irrigation time regulation in the production of the Tulips in Peru. Peiris, H. (2016) had used the ARIMA model in the tourism sector of Sri Lanka. They have predicted the consumer behavior

towards quality, safety, service quality, customer relationship, destination visit and the satisfaction level. The same method is used to check the food quality in the food supply chain (Melesse et al., 2022). Time series in data science is used to determine the quantity of the Oil & Gas production in the Mexican fields (Sánchez Morales & Soler Anguiano, 2022). The time series (SARIMA) model is used in the water management (Azad et al., 2022). The supply of drinking water level (RWL) expectation has turned into a provoking errand because of spatiotemporal changes in climatic circumstances and convoluted actual cycle. The Red Slopes Repository (RHR) is a significant wellspring of drinking and water system water supply in Thiruvallur area, Tamil Nadu, India. Additionally, it is expected to be changed over into the other useful administrations later on. Be that as it may, environmental change in the district is supposed to have resulted over the RHR's future possibilities. In traffic system, the same model is used to control the accident (Deretić et al., 2022). The forecasting of unemployment is also done through such models to set the fiscal policy of Greece (Dritsaki, 2016). Even though the inflation rate is forecasted by the said model (Otu et al., 2014), country like Nigeria.

2.3. Use of SARIMA and ARIMA model in Wholesale Trade

Taha Falatouri, Farzaneh Darbanian, Patrick Brandtner, Chibuzor Udokwu, (2022) have used the SARIMA model in retail trade connecting to the supply chain management of the company. All the distribution channel are inter-connected to the model starting from the carrying and forward agency (C&F), wholesalers and retailers. They have taken 37 months data and forecasted the seasonal behavior and demand of products in retail stores in terms of quantity and quality. The Canadian study of (Divisekara et al., 2021) is to display and estimate the red lentil costs utilizing the Occasional Autoregressive Coordinated Moving Typical model (SARIMA). Eight years of week-by-week information beginning from 2010 to 2019 which contain 521 perceptions, got from Saskatchewan.ca were utilized in this review. The typical red lentil cost in Saskatchewan was dollar 24.75 per 100 lb, and week by week costs were exceptionally fluctuating over the long run. The irregularity and instability of red lentils are demonstrated and determined by computing the occasional file and applying SARIMA models to the time series. The demand for the Asian tourist arrival is forecasted in Malaysia by using SARIMA (Nanthakumar et al., 2012).

3. Theoretical Background

The “Time series analysis: an application of SARIMA & ARIMA forecasting model in General Trade Store” is proposed in this study. The data analytics have become common in E-commerce and Modern Trade Retail Store but General Trade stores are lacking behind. This study aims to focus on the effective use of data analytics in General Trade store such as wholesaler and distributors and retailers. The prophecy of future sales is called as sales forecasting helps in business to understand future trends and behavior. it helps business to allocate resources in a right quantity at the right time. Hence, helps in overall inventory management. To perform sales forecasting, needs to require to study different tool. The ARIMA model tool has become the focus of data scientist to employ for sales forecasting. The time series analysis approach is banked on past figures for future augury. To enumerate a benefit of upsurge in Data Analytics with accessibility of historical sales data with respect to consumer buying behavior with respect to seasonality and future trends it now feasible for General Trade stores to forecast sales through SARIMA and ARIMA model (Tudor, 2022). The Google trend and machine learning are used to

forecast the sales (Retail et al., 2021).In a recent trend, the e-sales is depending on the sales forecasting (Tudor, 2022). As per the study of (Shah, 2020),the data scientist and managers are using different sales time-series forecasting methods such as “multiplicative Holt-Winters (HW), ad- dative HW, Seasonal Auto Regressive Integrated Moving Average (SARIMA) (A variant of Auto Regressive Integrated Moving Average (ARIMA)), Long Short-Term Memory Recurrent Neural Networks (LSTM) and the Prophet method by Facebook on thirty-two univariate sales time-series, data used to forecast sales is taken from time-series Data Library (TSDL)”, leads to advanced sales forecast. The ARIMA model is derived by using R, for forecasting the future value of commodities (Sagar Imambi et al., 2018).

3.1. Sales Forecasting

Forecasting is the activity of making predictions based past and present data. Businesses utilize forecasting to determine how to allocate their budgets or plan for anticipated expenses for an upcoming period of time. Prediction of future sales with the help of historical data is called as sales forecasting. It helps us to understand future trends and consumer behavior, it also helps in inventory management. (African et al., 2020) had used the SARIMA model to predict the business trends, affected by the consumer price, purchasing price and seasonality fluctuation.

3.2. SARIMA & ARIMA Model

A seasonal autoregressive integrated moving average model is a class of statistical model for analyzing and forecasting time series data. It is similar to ARIMA model, but one step ahead. It shows the importance of seasonality. Defining the ARIMA model: An autoregressive integrated moving average is a statistical analysis model that leverages time series data to forecast future trends. The wholesale price of tomatoes is studied by ARIMA and SARIMA model, based on the Nairobi wholesale tomato sales (Mathenge Mutwiri, 2019).

3.3. Time series analysis approach

Time series is the way of studying the data with respect to time in order to extract the meaningful information. The trend and development of the market performance is purely depending on the effects of secular trend, seasonal trend, cyclical fluctuation and irregular variations, on the business persuasion. The moving average, exponential smoothing and ARIMA are the suitable models of time series method.

3.4. Research Gap

From the above study, the adoption of SARIMA model for sales forecasting of retail trade is not found and aware to the trade. Herein, the “Time Series Analysis: An Application of SARIMA Model in General Trade to Forecast Sales” is taken as the main study of the research. Wherein, the seasonality effect and sales trend is proposed to visualize with the real-time data.

4. Methods

Data analyzing is the activity of analyzing raw data to make an insightful and informative decision and conclusions. It helps individual and organizations make sense of data. There are various types of data analysis, including descriptive, diagnostic, prescriptive and predictive analytics. Each type is used for specific purposes depending on the question a data analyst is

trying to answer. They use various tools and techniques to help organizations make decisions and succeed. (Soares et al., 2022) have tested different forecasting model to forecast the fashion trade retail business, based on the exponential smoothing, affected by seasonality in fashion trade. (Jeyakumar et al., 2017) have used the SARIMA model for product forecasting of appliances and consumables.

4.1 Inventory Management

Inventory management helps company to identify and maintain the requirement of stock in particular time duration. It tracks inventory from purchase to sale of goods, through this company ensures there's always enough stock to fulfill customers' requirement. It helps to manage surplus and shortage of stock. (Wicaksana, 2016) had predicted the monthly demand of vehicle for maintaining the inventory.

4.2. Linear Regression

Simple linear regression is a regression model that estimates the relationship between one independent variable and one dependent variable using a straight line. Both variables should be quantitative. Linear regression is commonly used for predictive analysis and modeling.

4.3. Logistic Regression

This type of statistical model (also known as *logit model*) is often used for classification and predictive analytics. Logistic regression estimates the probability of an event occurring, such as voted or didn't vote, based on a given dataset of independent variables

4.4 SARIMA Model

A seasonal autoregressive integrated moving average model is a class of statistical model for analyzing and forecasting time series data. It is similar to ARIMA model, but one step ahead. It shows the importance of seasonality. (Alano & Quinto, 2015) had forecasted the monthly domestic crude oil prices using time series SARIMA model in Philippines. In the CZECH Republic, SARIMA model is used in forecasting of demand in retail grocery stores (Patak, 2015).

4.5. Data Collection

In this particular session, we have aligned different models on the training data and went to check the accuracy (or error) on the training and test the data. The model which performs the best on the test data is an optimum model for us. The data of 48 retailers is taken in first stage and 132 data in second stage for the research. The random sample random is taken as the method of data collection, used in this study.

4.6. Objective of the Study

To study the SARIMA model Time Series Approach, which helps to forecast the sales of Wholesale trade to Retail trade.

4.7. Hypothesis of the Study

H- Time Series (SARIMA) model is not a significant model to help for forecasting the sales of Wholesale trade to Retail trade.

H_0- Time Series (SARIMA) model is not a significant model to help for forecasting the sales of Wholesale trade to Retail trade.
H_1- Time Series (SARIMA) model is a significant model to help for forecasting the sales of Wholesale trade to Retail trade.

5. Data Analysis

Two stage data analysis is made to make the result more prominent. In first stage 48 data are taken for analysis and in second stage 132 data are considered for analysis purposes. SPSS/Python/R are used to analyze and fit the models. The sales quantity of four months is taken for projective analysis, used in the Time Series Method. Starting from the year 1st January 2015 to 31st December 2019. The trend of the sales shown in Figure 1, and forecasted sales quantity (Kg) in year 30th April 2020 is 8620. The hypothesis is tested in Figure 2 and Figure 3 with 48 data and 132 data respectively. With 48 data, the model is not fit, shows Sigma 2 (0.000), but Prob(H) = 0.89. Whereas the result of 132 data shows, Sigma2 (0.000), Prob(H) = 0.23. The second test of 132 data is more appropriate to consider and fix the model in testing in comparison to test 1 of 48 data.
The standard equation of timeseries is considered in each step of the analysis.

$$\textbf{Equation 1}: \boldsymbol{Y(t)\beta + \varepsilon(t)}, \text{Where Y(t)} = \{\text{Yt.t} = 0, \pm 1, \pm 2, \ldots\}$$

Equation 2 SARIMA $\underbrace{(P, Q, R)}_{\text{Non-Seasonal}}\ \underbrace{(P, D, Q)}_{\text{Seasonal}}$

5.1. Hypotheses Testing

Figure 1: Read the data from the '.csv' file as a monthly Time Series

```
In [4]:  df = pd.read_csv('Sales_quantity.csv')
         df.head()
```

Out[4]:

	Sales_quantity
0	12729
1	11636
2	15922
3	15227
4	8620

```
In [5]:  Time_Stamp = pd.date_range(start='2015-01-01',periods=len(df),freq='M')
         Time_Stamp

Out[5]: DatetimeIndex(['2015-01-31', '2015-02-28', '2015-03-31', '2015-04-30',
                       '2015-05-31', '2015-06-30', '2015-07-31', '2015-08-31',
                       '2015-09-30', '2015-10-31', '2015-11-30', '2015-12-31',
                       '2016-01-31', '2016-02-29', '2016-03-31', '2016-04-30',
                       '2016-05-31', '2016-06-30', '2016-07-31', '2016-08-31',
                       '2016-09-30', '2016-10-31', '2016-11-30', '2016-12-31',
                       '2017-01-31', '2017-02-28', '2017-03-31', '2017-04-30',
                       '2017-05-31', '2017-06-30', '2017-07-31', '2017-08-31',
                       '2017-09-30', '2017-10-31', '2017-11-30', '2017-12-31',
                       '2018-01-31', '2018-02-28', '2018-03-31', '2018-04-30',
                       '2018-05-31', '2018-06-30', '2018-07-31', '2018-08-31',
                       '2018-09-30', '2018-10-31', '2018-11-30', '2018-12-31',
                       '2019-01-31', '2019-02-28', '2019-03-31', '2019-04-30',
                       '2019-05-31', '2019-06-30', '2019-07-31', '2019-08-31',
                       '2019-09-30', '2019-10-31', '2019-11-30', '2019-12-31',
                       '2020-01-31', '2020-02-29', '2020-03-31', '2020-04-30'],
                      dtype='datetime64[ns]', freq='M')
```

Source: Dr. Poorna Chandra Prasad

Figure 2 Time Series (SARIMA) model - Wholesale trade to Retail trade (48 data)

```
/Users/ssps/opt/anaconda3/lib/python3.8/site-packages/statsmodels/tsa/statespace/sarimax.py:
sonal ARMA. All parameters except for variances will be set to zeros.
  warn('Too few observations to estimate starting parameters%s.'
                                     SARIMAX Results
==========================================================================================
Dep. Variable:                     Sales_quantity   No. Observations:                   48
Model:             SARIMAX(3, 1, 3)x(0, 0, 3, 6)   Log Likelihood                -229.624
Date:                            Tue, 23 Feb 2021   AIC                            479.249
Time:                                    07:53:29   BIC                            491.437
Sample:                                01-31-2015   HQIC                           482.629
                                     - 12-31-2018
Covariance Type:                              opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.4033      2.528     -0.160      0.873      -5.358       4.552
ar.L2         -0.3966      2.539     -0.156      0.876      -5.372       4.579
ar.L3          0.6295      2.515      0.250      0.802      -4.300       5.559
ma.L1         -0.0044      2.937     -0.001      0.999      -5.761       5.753
ma.L2          0.2768      1.423      0.194      0.846      -2.513       3.066
ma.L3         -0.6402      2.086     -0.307      0.759      -4.728       3.448
ma.S.L6       -0.4915      0.894     -0.550      0.583      -2.244       1.261
ma.S.L12       0.6914      0.887      0.780      0.436      -1.047       2.429
ma.S.L18      -0.0899      0.915     -0.098      0.922      -1.883       1.703
sigma2      9.436e+06    4.4e-07   2.14e+13      0.000    9.44e+06    9.44e+06
===================================================================================
Ljung-Box (L1) (Q):                   1.10   Jarque-Bera (JB):                 0.53
Prob(Q):                              0.29   Prob(JB):                         0.77
Heteroskedasticity (H):               0.91   Skew:                             0.12
Prob(H) (two-sided):                  0.89   Kurtosis:                         2.33
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
[2] Covariance matrix is singular or near-singular, with condition number 2.08e+31. Standard
```

Figure 3 Time Series (SARIMA) model - Wholesale trade to Retail trade (132 data)

```
                                     SARIMAX Results
==========================================================================================
Dep. Variable:                          Sparkling   No. Observations:                  132
Model:             SARIMAX(2, 1, 3)x(2, 0, 3, 6)   Log Likelihood                -803.575
Date:                            Sun, 22 May 2022   AIC                           1629.151
Time:                                    21:54:15   BIC                           1658.755
Sample:                                01-01-1980   HQIC                          1641.156
                                     - 12-01-1990
Covariance Type:                              opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -1.7447      0.062    -27.921      0.000      -1.867      -1.622
ar.L2         -0.7869      0.067    -11.672      0.000      -0.919      -0.655
ma.L1          1.0844      0.161      6.715      0.000       0.768       1.401
ma.L2         -0.7525      0.122     -6.152      0.000      -0.992      -0.513
ma.L3         -0.8894      0.109     -8.158      0.000      -1.103      -0.676
ar.S.L6       -0.0108      0.029     -0.368      0.713      -0.068       0.047
ar.S.L12       1.0380      0.022     47.855      0.000       0.996       1.081
ma.S.L6        0.1216      0.179      0.679      0.497      -0.230       0.473
ma.S.L12      -0.5763      0.099     -5.841      0.000      -0.770      -0.383
ma.S.L18       0.0890      0.139      0.642      0.521      -0.183       0.361
sigma2      1.321e+05   1.81e-06   7.31e+10      0.000    1.32e+05    1.32e+05
===================================================================================
Ljung-Box (L1) (Q):                   0.01   Jarque-Bera (JB):                15.24
Prob(Q):                              0.93   Prob(JB):                         0.00
Heteroskedasticity (H):               1.50   Skew:                             0.38
Prob(H) (two-sided):                  0.23   Kurtosis:                         4.66
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
[2] Covariance matrix is singular or near-singular, with condition number 1.02e+26. Standard errors may be unstable.
```

Figure 4 Comparison of Different model - Wholesale trade to Retail trade (132 data)

	Test RMSE	RMSE	MAPE
RegressionOnTime	1798.200700	NaN	NaN
NaiveModel	3864.279352	NaN	NaN
SimpleAverageModel	1275.081804	NaN	NaN
2pointTrailingMovingAverage	811.178937	NaN	NaN
4pointTrailingMovingAverage	1184.213295	NaN	NaN
6pointTrailingMovingAverage	1337.200524	NaN	NaN
9pointTrailingMovingAverage	1422.653281	NaN	NaN
SARIMA(2,1,3)(2,0,3,6)	NaN	813.450419	0.357866

Source: Dr. Poorna Chandra Prasad

The SARIMA model along with other model are simultaneously used to test the hypothesis and found SARIMA matrix (2,1,3), (2,0,3,6) is positioning in second position (813.45,0.35) in comparison to other models. The normality and model fit are significant in the Figure 5.

Figure 5. Model Fit and Normality of Time Series (SARIMA) model - Wholesale trade to Retail trade (132 data)

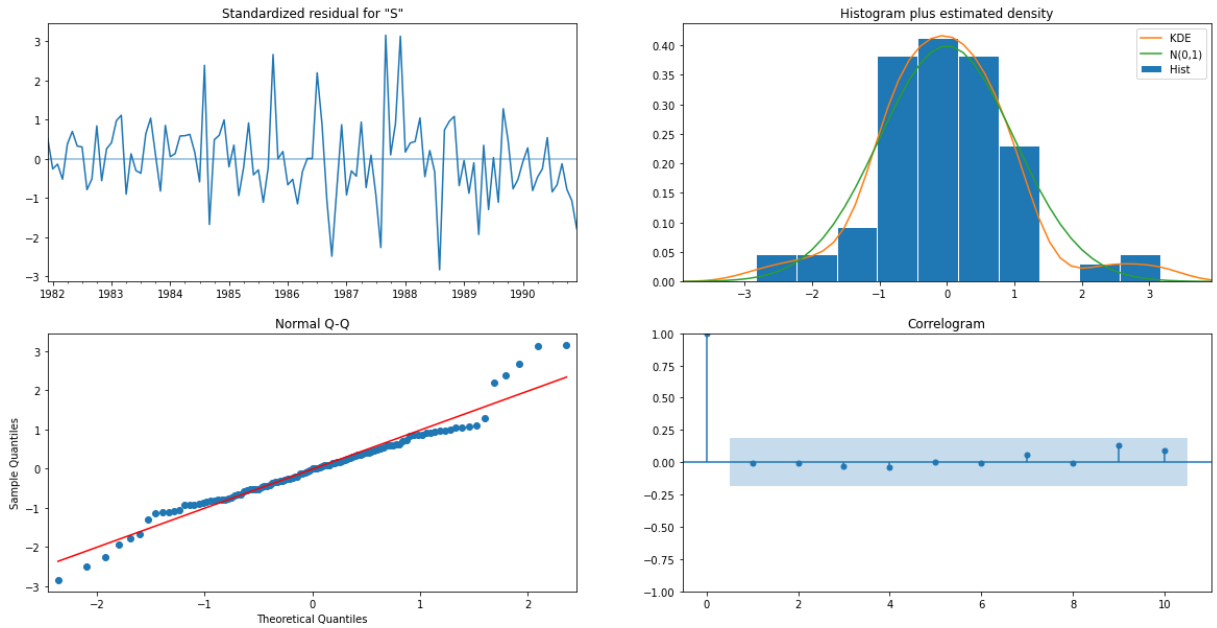


Figure 6. Time Series (SARIMA) model - Wholesale trade to Retail trade (132 data)

Out[236…

	param	seasonal	AIC
99	(1, 1, 2)	(0, 0, 3, 12)	14.000000
147	(2, 1, 1)	(0, 0, 3, 12)	14.000000
215	(3, 1, 1)	(1, 0, 3, 12)	208.047774
151	(2, 1, 1)	(1, 0, 3, 12)	347.672942
251	(3, 1, 3)	(2, 0, 3, 12)	349.867995

In [237…

```
import statsmodels.api as sm

auto_SARIMA = sm.tsa.statespace.SARIMAX(train['Sparkling'],
                                        order=(1, 1, 2),
                                        seasonal_order=(0, 0, 3, 12),
                                        enforce_stationarity=False,
                                        enforce_invertibility=False)
results_auto_SARIMA = auto_SARIMA.fit(maxiter=1000)
print(results_auto_SARIMA.summary())
```

The theoretical quantities and correlogram in Figure 5 estimates reliability of the data. In Figure 7, the sales data are taken as the input from the year 2015 to 2019. The output as the forecasted sales data of wholesale trade to the retail trade (132 data) are found for the year 2022. The AIC

(Akaike Information criterion) in between param 99 and 147 is 14.00, which is concentric of data set of 48 and 132. Nevertheless, the scatter of sales from wholesale to retail is much expedient.

Figure 7. Forecasting - Wholesale trade to Retail trade (132 data)

Proposed Data

Time_Stamp	Sales_quantity
2015-01-31	12729
2015-02-28	11636
2015-03-31	15922
2015-04-30	15227
2015-05-31	8620

Last few rows of Training Data

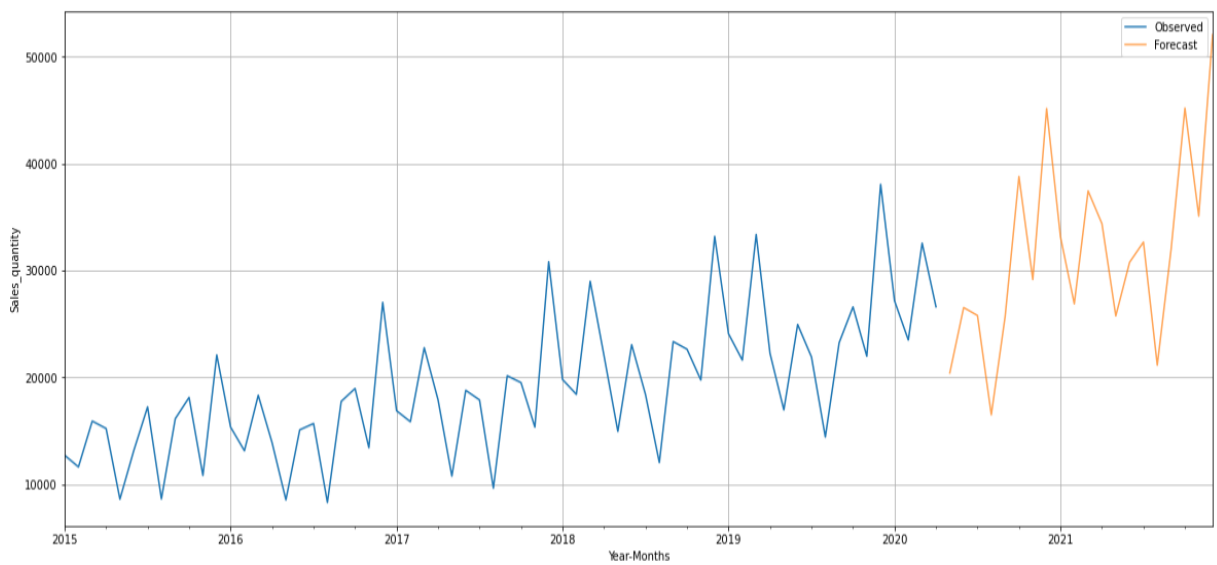
Time_Stamp	Sales_quantity
2018-08-31	12045
2018-09-30	23358
2018-10-31	22644
2018-11-30	19765
2018-12-31	33207

First few rows of Test Data

Forecasted Data

Time_Stamp	Sales_quantity
2019-12-31	38069
2020-01-31	27184
2020-02-29	23509
2020-03-31	32569
2020-04-30	26615

Figure 8. Forecasting - Wholesale trade to Retail trade (132 data)



Source: Dr. Poorna Chandra Prasad

Figure 9. Quantitative Result of Forecasting - Wholesale trade to Retail trade (132 data)

Result 1	Result 2

	Test RMSE	RMSE	MAPE
RegressionOnTime	1798.200700	NaN	NaN
NaiveModel	3864.279352	NaN	NaN
SimpleAverageModel	1275.081804	NaN	NaN
2pointTrailingMovingAverage	811.178937	NaN	NaN
4pointTrailingMovingAverage	1184.213295	NaN	NaN
6pointTrailingMovingAverage	1337.200524	NaN	NaN
9pointTrailingMovingAverage	1422.653281	NaN	NaN
SARIMA(2,1,3)(2,0,3,6)	NaN	813.450419	0.357866

```
resultsDf
```

	RMSE	MAPE
ARIMA(2,1,2)	4400.537137	14.606968
ARIMA(3,1,3)	4094.968771	14.383639
SARIMA(1,1,3)(3,0,3,6)	2388.755767	7.939640
SARIMA(3,1,3)(0,0,3,6)	2870.538437	9.367993

Source: Dr. Poorna Chandra Prasad

The quantitative result of Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) in Figure 9 has two results. In result 1, the SARIMA shows (RMSE, 813.45) and MAPE (0.35), which is very less error in comparison to results thereof. In Result 2, SARIMA (1,1,3) (0,0,3,6) is 7.93. The error in this model is lesser than other models, wherein tested.

6. Results and Discussion

Time Series (SARIMA) model is a significant model to help for forecasting the sales of Wholesale trade to Retail trade. From the above analysis, the result is clear which deals estimating is really significant by the lesser estimated error and significantly to realize irregularity factor. The mail objective of the study is to develop and utilize a proven model in wholesale trade, wherein the wholesalers will able to forecast the sales trend of coming years. The goal of this study was to develop a model by which the we can ready to estimate deals involving verifiable information as a source of perspective through the utilization of Time series Approach and comprehend the way of behaving, patterns and example exposed to irregularity utilizing SARIMA model. As per the result Time Series study, the seasonality variation is considered and well addressed to adhere the model. There were different models additionally analyzed and assessed named Direct relapse, Strategic relapse, weighted normal model (up to 4 attentions) and SARIMA model. The adoption of the SARIMA model in wholesale trade is more reliable in comparison to other models, best suited in time series methods.

7. Further Scope of Research

The forecasting of demand of quality, seasonal variance, adheres to wholesale-retail trade along with the demographic study is to be done in the future course of research. Even though, the area of study is purely based on the seasonality and quantity forecast by the wholesale trade of its retail traders, the qualitative prediction is not allowed in this model. So that, the future scope of the study peculates the forecasting of the quality, safety and security standard product. The secular trend, the seasonal trend, cyclical fluctuation and irregularities of variation to be considered in the future course of research. Apart from that, a large data from a particular city to be prejudicated on the analysis of ARIMA, SARIMA models to make the analysis accurate abide in similar research.

References

[1] Adanacioglu, H., & Yercan, M. (2012). An analysis of tomato prices at wholesale level in Turkey: An application of SARIMA model. *Custos e Agronegocio*, *8*(4), 52–75.

[2] African, S., Jean, B., Erero, L., & Makananisa, M. P. (2020). *Turkish Economic Review*. *7*(4).

[3] Alano, E. M. A. E. D., & Quinto, J. Q. (2015). *Modeling and Forecasting Foreign Trade of the Philippines using Time Series SARIMA Model Modeling and Forecasting Foreign Trade of the Philippines Using Time Series SARIMA Model*. *May*.

[4] Arunraj, N. S., Ahrens, D., Fernandes, M., & Müller, M. (2014). Time Series Sales Forecasting In Food Retail Industry. *Proceedings of the 34th International Symposium on Forecasting (ISF 2014), June 2014*. https://doi.org/10.13140/RG.2.1.4829.1607

[5] Azad, A. S., Sokkalingam, R., Daud, H., Adhikary, S. K., Khurshid, H., Mazlan, S. N. A., & Rabbani, M. B. A. (2022). Water Level Prediction through Hybrid SARIMA and ANN Models Based on Time Series Analysis: Red Hills Reservoir Case Study. *Sustainability (Switzerland)*, *14*(3). https://doi.org/10.3390/su14031843

[6] Deretić, N., Stanimirović, D., Al Awadh, M., Vujanović, N., & Djukić, A. (2022). SARIMA Modelling Approach for Forecasting of Traffic Accidents. *Sustainability (Switzerland)*, *14*(8). https://doi.org/10.3390/su14084403

[7] Divisekara, R. W., Jayasinghe, G. J. M. S. R., & Kumari, K. W. S. N. (2021). Forecasting the red lentils commodity market price using SARIMA models. *SN Business & Economics*, *1*(1), 1–13. https://doi.org/10.1007/s43546-020-00020-x

[8] Dritsaki, C. (2016). Forecast of Sarima Models: An Application to Unemployment Rates of Greece. *American Journal of Applied Mathematics and Statistics*, *4*(5), 136–148. https://doi.org/10.12691/ajams-4-5-1

[9] Jeyakumar, N., Jeyakumar, N., Program, M., & Seering, W. (2017). Signature redacted _ Signature redacted ___ Signat-ure redacted. 博士论文, *2000*, 1–195.

[10] Kam, H. J., Sung, J. O., & Park, R. W. (2010). Prediction of daily patient numbers for a regional emergency medical center using time series analysis. *Healthcare Informatics Research*, *16*(3), 158–165. https://doi.org/10.4258/hir.2010.16.3.158

[11] Kumar, M., & Anand, M. (2015). An Application of Time Series Arima Forecasting Model for Predicting Sugarcane Production in India. *Studies in Business and Economics*, 81–94.

[12] Mathenge Mutwiri, R. (2019). Forecasting of Tomatoes Wholesale Prices of Nairobi in Kenya: Time Series Analysis Using Sarima Model. *International Journal of Statistical Distributions and Applications*, *5*(3), 46. https://doi.org/10.11648/j.ijsd.20190503.11

[13] Melesse, T. Y., Bollo, M., Pasquale, V. Di, Centro, F., & Riemma, S. (2022). Machine Learning-Based Digital Twin for Monitoring Fruit Quality Evolution. *Procedia Computer Science*, *200*, 13–20. https://doi.org/10.1016/j.procs.2022.01.200

[14] Milenković, M., Švadlenka, L., Melichar, V., Bojović, N., & Avramović, Z. (2018). SARIMA modelling approach for railway passenger flow forecasting. *Transport*, *33*(5), 1113–1120. https://doi.org/10.3846/16484142.2016.1139623

[15] Nanthakumar, L., Subramaniam, T., & Kogid, M. (2012). Is “Malaysia Truly Asia”? Forecasting tourism demand from ASEAN using SARIMA approach. *Tourismos*, *7*(1), 367–381.

[16] Otu, A., George A., O., Jude, O., Hope Ifeyinwa, M., & Andrew I., I. (2014). Application of Sarima Models in Modelling and Forecasting Nigeria’s Inflation Rates.

American Journal of Applied Mathematics and Statistics, *2*(1), 16–28. https://doi.org/10.12691/ajams-2-1-4

[17] Pacco, H. C. (2022). Simulation of temperature control and irrigation time in the production of tulips using Fuzzy logic. *Procedia Computer Science*, *200*, 1–12. https://doi.org/10.1016/j.procs.2022.01.199

[18] Patak, M. (2015). Demand Forecasting in Retail Grocery Stores in the Czech Republic. *2nd International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM2015*, *2*(May 2017). https://doi.org/10.5593/sgemsocial2015/b22/s7.089

[19] Retail, F., Using, S., & Trends, G. (2021). *Forecasting Retail Sales Using Google Trends and Machine Learning*.

[20] Sagar Imambi, S., Vidyullatha, P., Santhi, M. V. B. T., & Haran Babu, P. (2018). Explore Big Data and forecasting future values using univariate ARIMA model in R. *International Journal of Engineering and Technology(UAE)*, *7*(March), 1107–1110. https://doi.org/10.14419/ijet.v7i2.7.12300

[21] Sánchez Morales, M. D. L. Á., & Soler Anguiano, F. I. (2022). Data science - Time series analysis of oil & gas production in mexican fields. *Procedia Computer Science*, *200*, 21–30. https://doi.org/10.1016/j.procs.2022.01.201

[22] Shah, V. (2020). A Comparative Study of Univariate Time-series Methods for Sales Forecasting. *University of Waterloo. Master Thesis*, 1–36.

[23] Soares, F. P., Master, A., Corte-real, C., Consult, S. R., Castro, F., Master, C., Engineering, M., & July, P. M. (2022). *Statistical forecasting models for fashion retailing industry*. *July*.

[24] Tudor, C. (2022). Integrated Framework to Assess the Extent of the Pandemic Impact on the Size and Structure of the E-Commerce Retail Sales Sector and Forecast Retail Trade E-Commerce. *Electronics (Switzerland)*, *11*(19). https://doi.org/10.3390/electronics11193194

[25] Wang, H., Tian, C. W., Wang, W. M., & Luo, X. M. (2018). Time-series analysis of tuberculosis from 2005 to 2017 in China. *Epidemiology and Infection*, *146*(8), 935–939. https://doi.org/10.1017/S0950268818001115

[26] Wicaksana, A. (2016). Comparative Study of Holt-Winters Triple Smoothing and Seasonal ARIMA: Forecasting Short Term Seasonal car sales in South Africa. *Risk Governance & Control: Financial Markets & Institutions /Https://Medium.Com/*, *6*(1), 1–18. https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf