IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

Seasonal And Trend Analysis of Drug Sales using Time Series Analysis

Vimal Kumar E S[1], Dr Chitra K[2], Mrs. Harilakshmi VM[3]
Student MCA, IVth Semester[1]
Associate Professor, Department of MCA[2]
Assistant Professor, Department of MCA[3]
Dayananda Sagar Academy of Technology and Management, Udaypura, Bangalore, Karnataka, India
vimalkumarchinnu786@gmail.com, chitra-mca@dsatm.edu.in
harilakshmi-mca@dsatm.edu.in

Abstract: *The SARIMA model is to analyse and forecast seasonal drug sales data, enhancing accuracy and understanding the complex of seasonal and trend patterns in fields like pharmaceutical sales, unlike the ARIMA model that only handles non-seasonal data. The model aims to improve predictions and provide deeper insights into the intricate seasonal and trend patterns in drug sales data. This study analyses drug sales behaviour across different seasons using drug store data from 1992 to 2010, grouping data by seasonal periods. Time series analysis is a method used to analyse data points over time to identify trends, patterns, and seasonal variations, providing valuable insights into drug sales. The study emphasizes the significance of drug sales data in real-time monitoring of best values of p,d,q activity*

Keywords: Drugs, Sales, SARIMA, Seasonal

I. INTRODUCTION

Machine learning, a subset of AI, enables computers to learn from data and make predictions without explicit programming. In the "Seasonal and Trend Analysis of Drug Sales: Using Time Series Analysis" project, machine learning algorithms are used to analyze and predict seasonal trends in drug sales data. The project uses machine learning to analyze and predict seasonal trends in drug sales data, using historical data to identify patterns and make accurate predictions about future sales. Machine learning algorithms are trained on extensive drug sales data from various seasons to identify patterns and make predictions. The project uses machine learning algorithms to analyze and predict seasonal trends in drug sales, providing accurate forecasting and deeper insights into the dynamics of the industry. The project uses the SARIMA algorithm to analyze daily drug sales, enhancing market understanding and predicting future trends. The SARIMA algorithm, an extension of the ARIMA model, is used to analyze historical drug sales data, revealing seasonal patterns and trends. This machine learning approach, a subset of AI, aids in accurate modeling and forecasting of complex seasonal data.

1.1 PROBLEM STATEMENT

Traditional forecasting models like ARIMA fail to account for seasonal demand fluctuations in retail and healthcare sectors, resulting in suboptimal predictions and potential financial losses. In the pharmaceutical industry, understanding the patterns and trends in drug sales is crucial for inventory management, supply chain optimization, and strategic planning. Drug sales data often exhibit complex patterns that include both seasonal effects and underlying trends. Effective analysis of this data can provide valuable insights into sales fluctuations, help predict future demands, and inform business decisions. However, analyzing time series data for drug sales can be challenging due to the presence of multiple interacting factors such as seasonal variations, long-term trends, and irregular fluctuations. Problem Context The primary objective of this project is to conduct a detailed seasonal and trend analysis of drug sales using the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The SARIMA model is particularly suited for time series data

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

1.2. LITERATURE SURVEY

Amy O'Donnell et al., The paper examines the impact of Minimum Unit Pricing (MUP) on alcohol consumption and public health in Scotland. It found that MUP led to increased purchase prices, reduced weekly alcohol purchases, and non-significant increases in alcohol expenditure. The policy targeted cheaper and stronger alcohol products, resulting in a 7.6% reduction in purchases [1].

Bunyanoot Kitsomcheep, Retail stores face challenges in using data collected from databases to gain a competitive advantage. They need to plan product arrangement, home page adjustments, and customer flow. Market Basket Analysis (MBA) is a data mining technique that helps companies determine which products customers purchase together, aiming to generate more sales through promotions, cross-sales, product recommendations, and product placement. The most popular association rule for MBA is used for sales transactions. Rapid Miner Studio can be used to perform MBA [2].

Deanna J. Buehrle et al., The study found a significant decrease in fluoroquinolone prescriptions in the US following FDA safety warnings in 2016 and 2018, attributed to stewardship efforts, with increased prescriptions among infectious diseases specialists and nurse practitioners [3].

Jason Brownlee on August 21,2019 in Time Series , Seasonal ARIMA is a seasonal component-supporting version of ARIMA for univariate time series data. It adds three new hyperparameters for autoregression, differencing, and moving average, and an additional parameter for the seasonality period [4].

Mathilde Pivette, The study uses drug sales data, particularly non-prescribed drug sales, to detect gastroenteritis epidemics earlier than traditional surveillance methods. A model with a 30% pruning value and 99% prediction interval successfully detected seasonal epidemics 1.6 weeks earlier than traditional methods. Geographic analyses showed high correlations between prescribed drug sales and reported cases, suggesting drug sales data can be a sensitive, specific, and timely indicator [5].

Pascal Beaudeau, Alain Le Tertre, This study looks into the relationship between the incidence of acute gastroenteritis (AGE) in the lower section of Le Havre and water turbidity. Analysis was done on medicine sales data from 1994 to 2000. The completed water turbidity and daily drug sales were shown to be connected, with settlement significantly changing risk. The timely adaptation of therapy to turbidity conditions is crucial in mitigating the infectious risk associated with karstic waters.[6].

Parekh, and William H. Shrank , The study examines the US's rising prescription drug costs from 2005 to 2016, focusing on new and existing drugs. Brand-name costs grew due to inflation, while generic and specialty drug costs increased due to new product entry. The study highlights policy needs [7].

Racquel Jandoc Andrea M. Burden et al., Interrupted time series analysis is a widely used quasi-experimental design used to evaluate the impact of interventions when a randomized controlled trial is not feasible. It uses aggregate data collected before and after an intervention, assuming trends before the intervention could be extrapolated to predict future trends [8].

II. METHODOLOGY

1. Data Collection: Time series data refers to regular observations or measurements over time, sourced from various sources like sensors, financial markets, weather stations, and economic indicators.
2. Data Preprocessing: Time series data often requires preprocessing to ensure stationarity through techniques like differencing or transformation,eliminating outliers, managing missing values, and handling missing values.
3. Exploratory Data Analysis (EDA): EDA is a statistical technique that uses line, scatter, autocorrelation, and decomposition plots to visually analyze time series data to find seasonality, trends, and patterns.
4. Modeling: Time series models, such as ARIMA, SARIMA and ETS are used to represent data structure and capture linear dependencies between observations.
5. Model Evaluation: The performance of a model is evaluated using metrics like MAE, MSE, or MAPE, after fitting it to the data.
6. Forecasting: Selecting and validating a model allows for future forecasting, providing valuable insights for decision-making in finance, economics, healthcare, and environmental science domains.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

Actions taken in SARIMA Modelling
Start:
1. Input: Time series data {y_t}, seasonal period m
2. Initialize: Set initial parameters and conditions
Step 1: Identify Seasonality
- Determine seasonal period m from the data
Step 2: Stationarity Check
- Perform differencing until stationary:
- Non-seasonal differencing d times
Step 3: Autocorrelation and Partial Autocorrelation Analysis
- Compute ACF and PACF:
- Identify AR order p from PACF
- Identify MA order q from ACF
Step 4: Seasonal Differencing
- Perform seasonal differencing:
- Seasonal differencing D times
Step 5: Seasonal Autocorrelation and Partial Autocorrelation Analysis
- Compute seasonal ACF and PACF:
- Identify seasonal AR order P from seasonal PACF
- Identify seasonal MA order Q from seasonal ACF
Step 6: Model Selection
- Choose SARIMA (p, d, q) (P, D, Q) model based on identified parameters
Step 7: Parameter Estimation
- Estimate model parameters using maximum likelihood estimation (MLE) or other methods
Step 8: Model Fitting
- Fit SARIMA model to the time series data
End
Input and Output Statements
Input:
- Time series data {y_t}, seasonal period m
This algorithm helps in translating head movements into cursor movements on the screen, ensuring smooth and stable control by accounting for small, unintentional movements through a neutral zone and using complementary filtering to reduce noise.

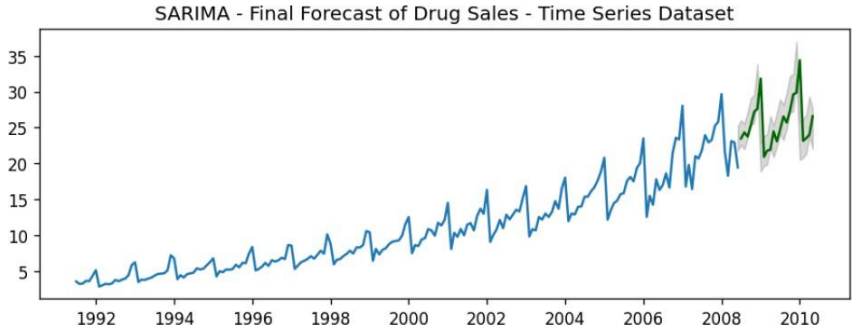

Figure 1 . SARIMA - Final Forecast of Drug Sales - Time Series Dataset

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

III. RESULTS AND DISCUSSIONS

The text provides a detailed analysis of a dataset for drug sales, detailing its time period, observations, and key statistics. It then presents the time series decomposition, trends analysis, seasonal patterns, model selection and validation, and forecasting. The analysis also discusses the selection of the appropriate time series model, its fitting results, and performance metrics used to validate the model. The model's accuracy is assessed by comparing forecasted values with actual data. The summary discusses the findings of a study, highlighting trends and seasonal patterns. It also discusses the practical implications for stakeholders like pharmaceutical companies and healthcare providers, suggesting informed decisions about inventory management and marketing strategies. It acknowledges limitations in the analysis and suggests areas for future research, such as exploring additional factors influencing drug sales.

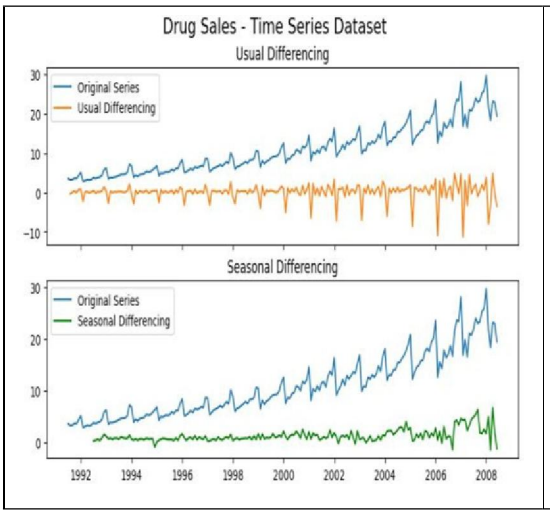

Figure 2. Seasonal Comparison

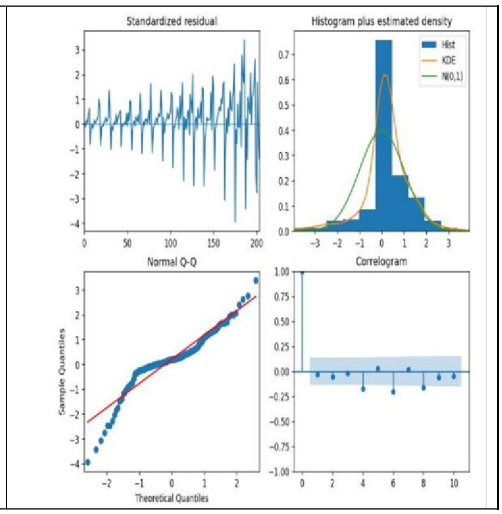

Figure 3. Residual plots in model

In figure 2 the blue line represents the original drug sales time series data, showing an upward trend and potential seasonal fluctuations. The orange line represents the usual differencing, which removes the trend component. The bottom plot shows the seasonal differencing, which removes seasonality. Differencing is important for time series forecasting models as it ensures data is stationary, ensuring its statistical properties don't change over time. SARIMA models, which rely on stationary data, often use differencing as a crucial step before fitting.

Figure 3 Depicts the plot shows the SARIMA model's standardized residuals, a random difference between actual values and predictions. It shows a histogram and probability density function superimposed on a normal distribution, aiming to determine if the residuals are normally distributed. The plot compares standardized residual quantiles to standard normal distribution quantiles, with points close to the diagonal line indicating normal distribution. The plot also displays autocorrelations, indicating independence or model inaccuracies.

IV. CONCLUSION

This paper introduces the study of time series analysis to identify seasonal trends in drug sales. Techniques like decomposition and SARIMA modeling reveal seasonal fluctuations. This helps in informed decision-making, resource allocation, healthcare delivery, and policy planning. The SARIMA model provided accurate forecasts, aiding inventory management and marketing strategies. Future research could incorporate external factors, comparative analyses, and regional sales data for more localized insights.

REFERENCES

[1]. Amy O'Donnell,1 Peter Anderson,1,2 Eva Jané-Llopis,2,3,4 Jakob Manthey,5,6 Eileen Kaner,1 Immediate impact of minimum unit pricing on alcohol purchases in Scotland: controlled interrupted time series analysis for 2015-18

[2]. Bunyanoot Kitsomcheep, Market Basket Analysis Of Drug Store In Seasonal Period

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 6, November 2024

[3]. Deanna J. Buehrle,a Marilyn M. Wagener,b Cornelius J. Clancya,b Outpatient Fluoroquinolone Prescription Fills in the United States, 2014 to 2020: Assessing the Impact of Food and Drug Administration Safety Warnings
[4]. Jason Brownlee on August 21,2019 in Time Series ,A Gentle Introduction to SARIMA for Time Series Forecasting in Python
[5]. Mathilde Pivettea,b,c,∗, Judith E. Muellera,d, Pascal Crépeya,e, Avner Bar-Henb Surveillance of gastrointestinal disease in France using drug sales data
[6]. Pascal Beaudeau, Alain Le Tertre, A time series study of drug sales and turbidity of tap water in Le Havre, France
[7]. Parekh, and William H. Shrank The Contribution Of New Product Entry Versus Existing Product Inflation In The Rising Costs Of Drugs
[8]. Racquel Jandoca, Andrea M. Burdena, Muhammad Mamdania,b,c, Linda E. Levesqued, Suzanne M. Cadarettea, Interrupted time series analysis in drug utilization research is increasing: systematic review and recommendations