Modeling the Impact of Holidays and Events on Retail Demand Forecasting in Online Marketing Campaigns using Intervention Analysis

Jeffri Prayitno Bangkit Saputra[1,*], Aayush Kumar[2]

[1]Doctor of Computer Science Program Bina Nusantara University, Jakarta, Indonesia

[2]Computer Science and Engineering Department, Indian Institute of Technology, Kanpur, India

ABSTRACT

This study explores the impact of holidays and events on retail demand forecasting using intervention analysis within a SARIMAX model framework. Retail demand forecasting is critical for inventory management and supply chain optimization. Traditional forecasting models often struggle to account for irregular events like holidays, leading to inaccuracies. This study aims to address these limitations by incorporating holidays and events as exogenous variables in the forecasting model. The dataset, consisting of retail sales records across multiple product categories, was preprocessed to handle missing values and standardize date formats. Binary indicators for state holidays and school holidays were created, along with temporal features like the day of the week and hour of the day. The stationarity of the time series was confirmed using the Augmented Dickey-Fuller (ADF) test, with a statistic of -48.67066391486136 and a p-value of 0.0. The SARIMAX model (1, 1, 1)x(1, 1, 1, 24) was developed and evaluated. The model achieved an Akaike Information Criterion (AIC) of 363321.861 and a Bayesian Information Criterion (BIC) of 363375.269. Key coefficients included the state holiday variable at 0 (p-value: 1.000000) and the school holiday variable at 165.2158 (p-value: 0.919689), though neither were statistically significant. Diagnostic checks revealed significant non-normality and heteroscedasticity in the residuals. Forecasting accuracy was assessed using Mean Absolute Error (MAE: 8057.069376036054) and Mean Squared Error (MSE: 809008799.3517022). The Mean Absolute Percentage Error (MAPE) was not computable due to division by zero. Visualizations comparing forecasted versus actual demand highlighted the model’s strengths in capturing general trends and seasonal patterns but indicated challenges in accurately predicting demand during holidays and events. The study underscores the importance of incorporating holidays and events into demand forecasting models and suggests further refinement and the inclusion of additional variables for improved accuracy. Future research should explore alternative modeling approaches and validate findings across multiple datasets to enhance the generalizability and robustness of the forecasting tools.

Keywords Retail Demand Forecasting, SARIMAX, Intervention Analysis, Holidays, Events, Time Series Analysis, Retail Sales

Submitted 10 January 2024
Accepted 20 April 2024
Published 1 June 2024

Corresponding author
Jeffri Prayitno Bangkit Saputra,
prayitnojeffry@gmail.com

Additional Information and Declarations can be found on page 162

DOI: 10.47738/jdmdc.v1i2.9

INTRODUCTION

Retail demand forecasting is a crucial process in the retail industry, essential for financial performance, supply chain efficiency, and overall business success [1]. Accurate forecasting directly impacts a retailer's profitability and competitive position, as poor forecasting can lead to understock or overstock situations [2]. In the apparel retail sector, demand forecasting is particularly crucial for

How to cite this article: J.P.B. Saputra and A. Kumar, “Modeling the Impact of Holidays and Events on Retail Demand Forecasting in Online Marketing Campaigns using Intervention Analysis,” *J. Digit. Mark. Digit. Curr.*, vol. 1, no. 2, pp. 144-164, 2024.

sustaining business operations, especially when dealing with a variety of products and intermittent demand patterns [3]. This process is essential for making informed business decisions and ensuring efficient supply chain management. By analyzing historical sales data, market trends, and various external factors, retailers can anticipate demand fluctuations and align their inventory and procurement strategies accordingly.

Accurate demand forecasting is crucial for retailers to effectively manage their inventory. By utilizing precise forecasts of future demand, retailers can make informed decisions regarding inventory levels, such as setting safety stocks and optimizing replenishment strategies [4]. Forecasting models, including time series techniques like SARIMA and ETS, play a significant role in predicting demand for various products, aiding in reducing wastage and enhancing supply chain management [5]. The importance of accurate sales forecasts for efficient inventory management has long been acknowledged in the retail industry [2]. This not only improves operational efficiency but also enhances customer satisfaction by ensuring product availability.

The significance of retail demand forecasting in supply chain management cannot be overstated. In a competitive retail environment, the ability to predict demand accurately is a key differentiator. It enables retailers to plan their logistics and distribution activities more effectively, ensuring that products are available at the right place and time. Traditional methods for demand forecasting include time series analysis, moving averages, exponential smoothing, and causal models. Time series analysis, such as ARIMA (AutoRegressive Integrated Moving Average) models, involves analyzing historical data to identify patterns and project future demand. Moving averages and exponential smoothing techniques smooth out short-term fluctuations to highlight longer-term trends. Causal models, on the other hand, incorporate external factors like promotions, pricing changes, and economic indicators to predict demand. Each of these methods has its strengths and limitations, often depending on the context and data availability.

However, predicting retail demand is fraught with challenges. One significant challenge is the inherent volatility and seasonality in retail demand. Retail sales are often influenced by seasonal trends, such as increased demand during holiday seasons or special promotional periods. These fluctuations can be substantial and unpredictable, making it difficult to maintain accurate forecasts. Additionally, external factors such as economic conditions, competitive actions, and changes in consumer preferences can introduce further uncertainty and complexity into the forecasting process. Economic downturns or booms, competitor promotions, and shifting consumer tastes can all impact demand in ways that are hard to predict with traditional forecasting models.

Moreover, irregular events such as holidays, special events, and public holidays pose a unique challenge to demand forecasting. These events can cause sudden and significant spikes or drops in demand that traditional models may not capture effectively. For example, holidays like Christmas or Black Friday typically see a surge in retail sales, while other events might lead to reduced shopping activity. Standard forecasting techniques, which rely on regular patterns and continuity in data, often struggle to accommodate these irregularities. This necessitates the use of more advanced forecasting methods that can incorporate the effects of such events, improving the accuracy and reliability of demand predictions.

Holidays and events have a significant impact on consumer behavior. Research indicates that personality traits can predict how individuals respond to the demands of the holiday season, highlighting the connection between personality and consumer behavior [6]. Studies have also shown that holiday season discounts influence consumer behavior, emphasizing the importance of understanding purchasing patterns during this time [7]. The rise in holiday travel demand, driven by activities like sightseeing, shopping, and family gatherings, has attracted attention from policymakers and researchers, underscoring the necessity of analyzing travel patterns during holidays [8].

During these periods, consumers tend to spend more on gifts, decorations, travel, and other holiday-specific items. For instance, Christmas, Black Friday, and national holidays like Independence Day are known for substantial spikes in retail sales as consumers engage in celebratory shopping, take advantage of discounts, and purchase goods in preparation for gatherings and festivities. This increased spending during holidays can result in significant deviations from normal demand patterns, making it critical for retailers to anticipate and prepare for these fluctuations.

Conversely, some events may cause a decline in retail demand. For example, during school holidays, families might go on vacations, leading to reduced local shopping activity. Similarly, certain public holidays that do not involve gift-giving or celebrations might see a drop in retail sales as consumers stay home and limit their expenditures. Understanding these shifts in consumer behavior is crucial for retailers aiming to optimize their inventory and sales strategies around such events.

Traditional forecasting models, while useful, often fall short in capturing the effects of holidays and events on retail demand. These models typically rely on historical sales data and assume a degree of regularity and continuity in the demand patterns. However, holidays and special events introduce irregularities that disrupt these patterns, leading to inaccuracies in predictions. Standard methods like ARIMA, moving averages, and exponential smoothing do not inherently account for these abrupt changes, resulting in forecasts that may either overestimate or underestimate demand during these critical periods.

Intervention analysis emerges as a powerful method to address these limitations. This advanced technique incorporates exogenous variables representing holidays and events directly into the forecasting model. By doing so, it allows for adjustments based on the presence and expected impact of these irregular events. For example, intervention analysis can include binary indicators for holidays, special events, or promotional periods, enabling the model to differentiate between normal and event-driven demand.

The benefits of using intervention analysis in retail demand forecasting are substantial. By improving the accuracy of demand forecasts, retailers can manage their inventory more effectively, ensuring that they have sufficient stock to meet increased demand during peak periods while avoiding overstock during low-demand periods. Enhanced forecasting accuracy also aids in strategic planning for promotional events, allowing retailers to capitalize on increased consumer spending during holidays and events. Overall, intervention analysis provides a robust framework for incorporating the complexities of consumer behavior during holidays and events, leading to more reliable and actionable demand forecasts.

The primary objective of this study is to develop a robust forecasting model that explicitly incorporates the effects of holidays and events on retail demand. Traditional forecasting models often fail to account for these irregularities, leading to inaccuracies. Therefore, a critical step involves identifying relevant holidays and events that significantly impact consumer behavior and retail demand. These events will be represented as binary indicators in the dataset, marking the presence or absence of a particular event on a given day. For instance, a binary indicator for Christmas would be '1' on December 25th and '0' on all other days.

The framework chosen for this study is the ARIMAX (Autoregressive Integrated Moving Average with Exogenous Variables) model. This model extends the traditional ARIMA model by incorporating exogenous variables, which, in this case, are the binary indicators for holidays and events. The ARIMAX model is well-suited for this application as it can handle both the regular components of the time series data (such as trends and seasonality) and the external shocks introduced by holidays and events. By integrating these indicators into the model, we aim to capture the additional variance in demand caused by these events, leading to more accurate forecasts.

To determine the effectiveness of the intervention analysis model, we will assess its performance using standard forecasting metrics. These metrics include Mean Absolute Error (MAE), which measures the average magnitude of the errors in a set of predictions, without considering their direction; Mean Squared Error (MSE), which squares the average of the errors to give more weight to larger errors; and Mean Absolute Percentage Error (MAPE), which provides a percentage error measurement that is useful for comparing forecast accuracy across different scales.

A comparative analysis will be conducted to evaluate the improvements offered by the intervention model over traditional forecasting models. This involves comparing the ARIMAX model's performance metrics with those of models that do not include exogenous variables for holidays and events, such as standard ARIMA models. By demonstrating superior accuracy and reliability, the intervention model can be validated as a more effective tool for retail demand forecasting.

In addition to statistical performance metrics, the practical implications of the enhanced forecasting model will be discussed. Improved demand forecasts have significant benefits for retailers, including more effective inventory management, reduced holding costs, and better preparedness for peak shopping periods. This can lead to increased customer satisfaction by ensuring product availability during high-demand periods and minimizing stockouts. Additionally, strategic planning capabilities are enhanced, allowing retailers to optimize promotional activities and resource allocation. The practical benefits underscore the value of incorporating holidays and events into demand forecasting models, providing a compelling case for the adoption of advanced forecasting techniques in the retail industry.

Literature Review

Retail Demand Forecasting

Retail demand forecasting is a well-researched area within supply chain management and retail operations, with numerous methodologies developed to predict future demand. Traditional methods primarily include time series analysis, which involves the examination of historical data to identify patterns and project future trends. Among these methods, ARIMA (AutoRegressive Integrated Moving Average) models are particularly prominent. ARIMA models are used to analyze and forecast time series data by accounting for various components, such as autocorrelation, differencing to achieve stationarity, and moving averages to smooth out fluctuations. These models are valued for their flexibility and ability to handle various types of data trends and seasonal patterns.

Another commonly used approach is the moving average method, which calculates the average demand over a specified number of past periods to predict future demand. This method is simple and easy to implement but can lag behind trends and react slowly to changes in the data. Exponential smoothing is a more advanced technique that applies decreasing weights to past observations, allowing the model to be more responsive to recent changes in the data. It includes methods like Simple Exponential Smoothing (SES), Holt's Linear Trend Model, and Holt-Winters Seasonal Model, each adding layers of complexity to handle trends and seasonality more effectively.

Causal models are also employed in retail demand forecasting. These models incorporate external factors, such as economic indicators, promotional activities, and pricing strategies, to predict demand. Causal models are particularly useful when the demand is influenced by identifiable external variables. Despite their sophistication, these models require extensive data on the external factors and their relationships with demand, which can be a limitation.

While these traditional methods are widely used and have proven effective in many scenarios, they have notable limitations in handling irregular events like holidays. Standard time series models assume that future patterns will resemble past patterns, with regular intervals and consistent seasonality. However, holidays and special events introduce anomalies that disrupt these patterns, causing sudden and significant deviations in demand that traditional models are not designed to handle. For example, an ARIMA model might predict demand based on historical sales data, but it would not automatically account for the impact of an upcoming holiday unless the holiday effects are explicitly included in the model.

The inability to capture these irregular events can lead to inaccurate forecasts and suboptimal inventory management. During holidays or special events, consumer behavior often changes dramatically, resulting in either spikes or drops in demand. Retailers relying on traditional forecasting methods may either overestimate or underestimate the actual demand, leading to issues such as stockouts, excess inventory, and missed sales opportunities. Therefore, there is a growing recognition of the need for advanced forecasting techniques that can incorporate the effects of holidays and events, improving the accuracy and reliability of retail demand forecasts.

Recent research has started to address these limitations by exploring models that include exogenous variables representing holidays and events. Intervention analysis, for instance, is one such advanced technique that adjusts the forecast based on the presence of irregular events. This approach has shown promise in various fields, including retail, where accounting for holidays and events can significantly enhance forecast accuracy. By incorporating these factors into the forecasting models, researchers aim to bridge the gap between traditional methods and the complex realities of retail demand, providing more reliable tools for inventory and supply chain management.

Impact of Holidays and Events

The impact of holidays and events on consumer purchasing behavior has been a significant area of study within retail and consumer behavior research. Numerous studies have highlighted how these special periods influence buying patterns, often resulting in marked deviations from regular demand trends.

One of the primary ways holidays and events affect consumer behavior is through changes in spending habits. Consumer spending tends to rise significantly during holidays such as Christmas, Thanksgiving, and Black Friday. Black Friday, in particular, is highlighted as one of the most popular shopping days globally, with a surge in consumer purchasing behavior, especially in the United States where it follows Thanksgiving [9]. Studies have delved into the rituals and perceived value associated with Black Friday, emphasizing its significance in driving consumer behavior and loyalty [10]. Research has also explored the sentiment and behavior response of consumers towards Black Friday deals, indicating the event's global marketing opportunities and strategies aimed at boosting sales [11]. The anticipation of these holidays and the associated promotional activities further amplify consumer spending.

Beyond these commonly studied holidays, there are also less frequent but highly impactful events. For instance, sporting events like the Super Bowl can lead to increased demand for specific products such as snacks, beverages, and entertainment equipment. Cultural events and festivals, depending on their scale and nature, can also influence retail demand in unique ways. For example, the Chinese New Year is associated with increased consumer spending on gifts, food, and decorations, significantly boosting retail sales in regions where this festival is celebrated.

Intervention Analysis

Intervention analysis in statistics involves employing various statistical methods to evaluate the impact of interventions on specific outcomes. Segmented regression analysis is a commonly used statistical method in interrupted time series studies to estimate the intervention effect by comparing post-intervention trends with pre-intervention trends [12], [13]. This approach allows for a rigorous assessment of the effectiveness of interventions in fields such as pre-hospital ambulance care and quality improvement project [14].

Furthermore, statistical analysis plays a crucial role in determining the efficacy of interventions in randomized controlled trials. Usage metrics of web-based interventions can inform statistical analyses to estimate the efficacy of interventions received, going beyond simple effectiveness assessments [15]. Similarly, statistical methods like principal stratification are utilized in environmental intervention studies to analyze the effects of interventions on health outcomes and environmental factors [16].

The fundamental concept of intervention analysis is to model the impact of an external event as an intervention in the time series. This is typically done by introducing binary or dummy variables that indicate the occurrence of the event.

For instance, if a significant holiday like Christmas affects retail sales, a binary variable representing the presence of Christmas can be added to the model. This variable takes the value of 1 during the holiday period and 0 otherwise. By incorporating such variables, the model can separate the regular time series components (like trends and seasonality) from the irregular effects caused by the interventions.

Intervention analysis is often applied using the ARIMAX (Autoregressive Integrated Moving Average with Exogenous Variables) framework. The ARIMAX model extends the traditional ARIMA model by including exogenous variables. This allows the model to account for both the internal structure of the time series data and the external impacts of the interventions. The exogenous variables can capture sudden changes in the level, slope, or seasonal pattern of the series due to the intervention.

Gap in Literature

Despite the extensive research on retail demand forecasting and the proven benefits of intervention analysis, several gaps remain in the existing literature that this study aims to address. These gaps highlight the need for further exploration and refinement of forecasting models, particularly in the context of accounting for holidays and events.

One significant gap in the literature is the limited integration of intervention analysis with high-frequency retail data. Most studies have focused on daily or weekly sales data, which, while useful, may not capture the more granular fluctuations that occur within shorter time intervals. High-frequency data, such as hourly sales data, can provide deeper insights into consumer behavior patterns during holidays and events. However, there is a lack of research on how intervention analysis can be effectively applied to such high-frequency data. This study aims to fill this gap by developing and testing models that incorporate hourly sales data, providing a more detailed and accurate representation of demand fluctuations.

Another gap is the underutilization of comprehensive feature engineering in conjunction with intervention analysis. While many studies incorporate basic binary indicators for holidays and events, there is potential for more sophisticated feature engineering techniques to enhance model performance. For instance, creating interaction terms between different types of holidays, incorporating lagged effects of events, and accounting for varying intensities of promotional activities can provide richer information to the model. This study will explore advanced feature engineering methods to improve the representation of holidays and events in the forecasting model.

Furthermore, there is a need for comparative studies that rigorously evaluate the performance of intervention analysis against other advanced forecasting methods. While intervention analysis has shown promise, its comparative effectiveness relative to machine learning approaches, such as random forests, gradient boosting, and neural networks, remains underexplored. By conducting a thorough comparative analysis, this study aims to determine the conditions under which intervention analysis outperforms or complements these machine learning techniques. This can provide valuable insights for practitioners on selecting the most appropriate forecasting method based on their specific context and data characteristics.

The literature also reveals a gap in the practical application and validation of these models in real-world retail environments. Many studies focus on theoretical or simulated data, with limited validation using actual retail sales data. This disconnects between research and practice can hinder the adoption of advanced forecasting models by retailers. This study addresses this gap by utilizing real-world retail sales data to develop and validate the intervention analysis model, ensuring its practical relevance and applicability.

Lastly, there is a need for more research on the long-term impact of incorporating holidays and events into demand forecasting models. While short-term improvements in forecast accuracy are well-documented, the long-term benefits, such as sustained inventory optimization, improved customer satisfaction, and enhanced strategic planning capabilities, are less explored. This study will investigate both the immediate and long-term impacts of using intervention analysis, providing a holistic view of its benefits for retailers.

Method

Research method for this study detailed and shown in figure 1 below:

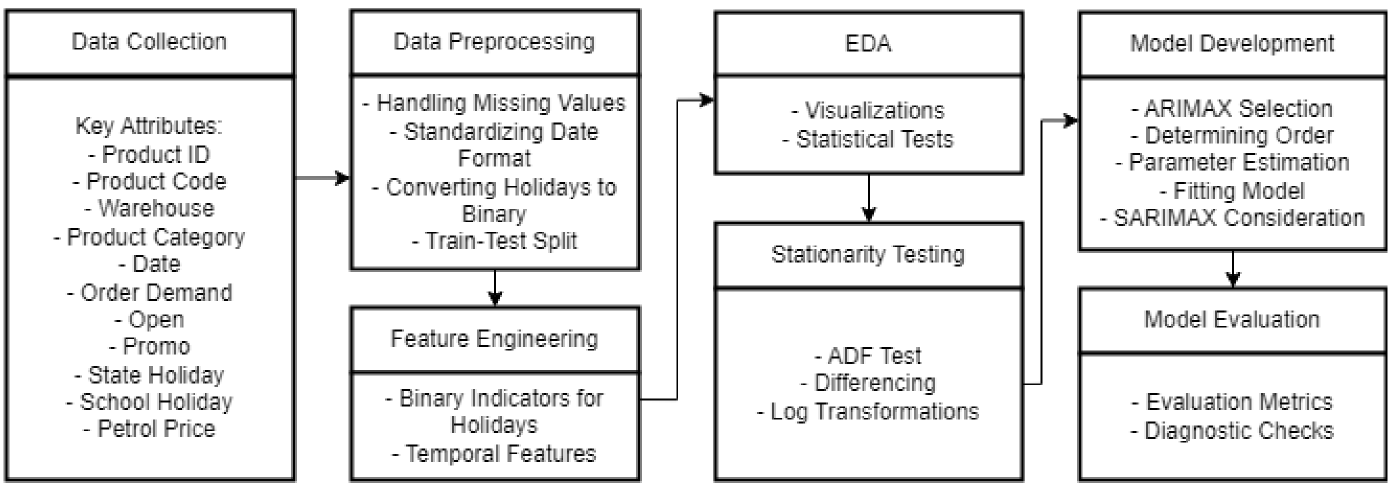

Figure 1 Research Method

Data Collection

The dataset used in this study consists of detailed retail sales records from a retail chain, covering a wide range of product categories such as electronics, apparel, home goods, and groceries. This comprehensive dataset spans multiple years, capturing both regular sales patterns and the impacts of various holidays and special events, allowing for a robust analysis of seasonal trends and specific event effects on retail demand. Each record includes key attributes such as Product ID, Product Code, Warehouse, Product Category, Date, Order Demand, Open (store status), Promo (promotion status), State Holiday, School Holiday, and Petrol Price, providing a rich set of variables for analysis and enabling the modeling of demand with respect to various influencing factors.

In preparing the data for analysis, several preprocessing steps were undertaken. Firstly, missing values in the dataset were addressed. For numeric columns, missing values were filled using the median value of the respective columns to maintain the central tendency without being influenced by outliers. For categorical columns, missing values were filled using the mode, reflecting the most common category. The 'Date' column, which originally contained dates in various formats, was standardized to a consistent datetime format using Python's pandas library, ensuring uniformity across the dataset, which is crucial for accurate time series analysis.

To incorporate holidays and events into the model, the 'State Holiday' and 'School Holiday' columns were converted to binary indicators. Non-numeric values in these columns were transformed into integers (1 for holiday, 0 for non-holiday), facilitating their use as exogenous variables in the modeling process. Additionally, feature engineering was performed to create new variables that capture the effects of temporal elements on demand. This included extracting the day of the week from the 'Date' column to account for weekly sales patterns and extracting the hour of the day to capture intraday variations in demand.

For initial testing and model development, a 10% random sample of the data was used to make the dataset more manageable. This approach ensured that the model could be efficiently developed and validated before scaling up to the full dataset. The stationarity of the time series was assessed using the Augmented Dickey-Fuller (ADF) test, and first-order differencing was applied to achieve stationarity where necessary, as this is a prerequisite for effective ARIMA modeling.

Finally, the dataset was split into training and testing sets based on time, with the last 30 days of data reserved for testing. This ensured that the model's performance could be evaluated on recent, unseen data, while the remaining data was used for training the model, allowing it to learn patterns and relationships from a substantial historical record. These preprocessing steps ensured that the dataset was clean, consistent, and ready for subsequent analysis and modeling, facilitating the development of a robust demand forecasting model that incorporates the effects of holidays and events.

Feature Engineering

In developing a robust demand forecasting model, feature engineering played a crucial role in capturing the impact of holidays and events on retail demand. To effectively integrate these effects into the model, relevant features were constructed and additional temporal features were created.

Firstly, binary indicators for state holidays and school holidays were constructed. These indicators were essential for capturing the influence of holidays on consumer purchasing behavior. The 'State Holiday' and 'School Holiday' columns, which initially contained non-numeric values, were converted into binary format. Each holiday was represented by a binary variable, where a value of 1 indicated the presence of a holiday and a value of 0 indicated a regular day. This transformation allowed the model to differentiate between normal and holiday periods, making it possible to assess the distinct impact of holidays on retail demand.

In addition to holiday indicators, several temporal features were engineered to capture the underlying patterns in the data. One of the key features created was the 'Day of the Week,' extracted from the 'Date' column. This feature accounted for weekly sales cycles and recurring patterns in consumer behavior, such as higher sales during weekends compared to weekdays. By including the day of the week as a feature, the model could better understand and predict the regular weekly fluctuations in demand.

Another important temporal feature was the 'Hour of the Day,' which was also extracted from the 'Date' column. This feature captured intraday variations in sales, recognizing that consumer shopping behavior can vary significantly throughout the day. For instance, peak shopping hours may differ from early morning or late-night periods. Including the hour of the day in the model helped to refine the forecasts by incorporating these finer-grained temporal patterns.

By constructing these relevant features related to holidays and events, and creating additional temporal features such as the day of the week and hour of the day, the feature engineering process enriched the dataset. This enrichment enabled the demand forecasting model to more accurately reflect the complexities of retail demand, accounting for both regular temporal patterns and the irregular impacts of holidays and events.

Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) was conducted to uncover patterns, trends, and potential seasonality in the retail demand data, as well as to investigate the relationship between demand and holidays/events. EDA provided crucial insights that guided the subsequent modeling efforts.

Firstly, visualizations were created to identify patterns and trends in the data. Time series plots were generated to visualize the retail demand over the entire period covered by the dataset. These plots, as shown in figure 2 below, highlighted the overall trends, such as increasing or decreasing demand over time, and helped in detecting any cyclical patterns.

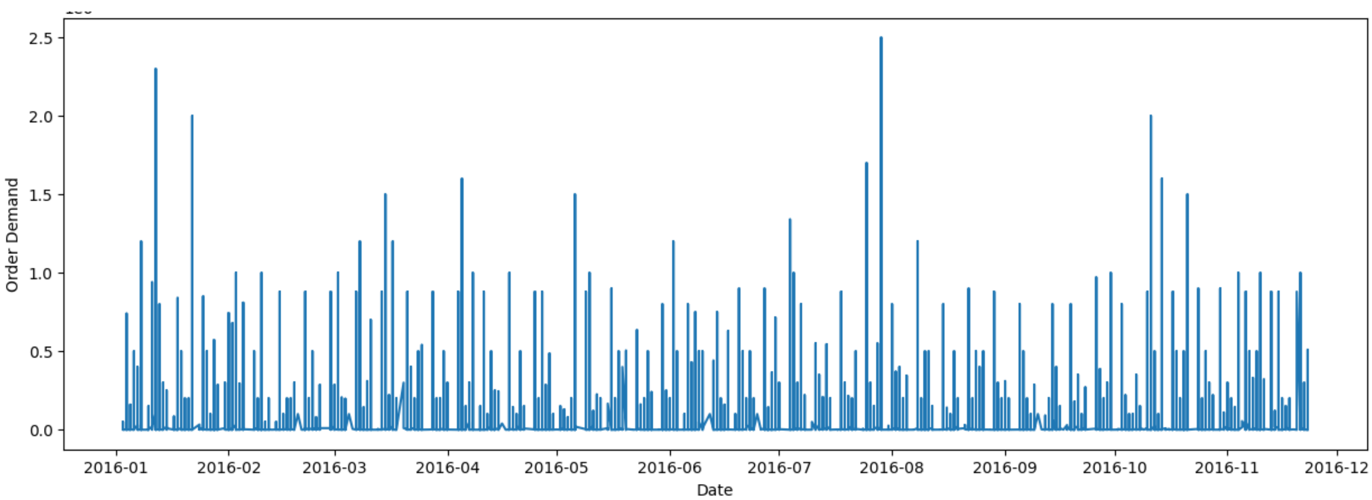

Figure 2 Retail Demand Over Time

Box plots and violin plots were used to compare the distribution of retail demand during holidays versus non-holidays. These plots, as shown in figure 3 below, illustrated how demand varied significantly on holidays compared to regular days, providing visual evidence of the impact of holidays on sales.

Figure 3 **Order Demand During School Holidays vs non-Holidays**

To further investigate the impact of specific holidays and events, t-tests and ANOVA (Analysis of Variance) tests were employed. These statistical tests compared the means of retail demand during different periods, such as holiday seasons versus regular periods, and assessed whether the differences were statistically significant. By performing these tests, the study validated the hypothesis that holidays and events have a significant impact on retail demand.

The Augmented Dickey-Fuller (ADF) test was also part of the EDA process to check for stationarity in the time series data. Stationarity is a key assumption for many time series forecasting models, including ARIMA and ARIMAX. If the time series was found to be non-stationary, appropriate transformations, such as differencing, were applied to achieve stationarity, ensuring that the data met the assumptions required for accurate modeling.

Stationarity Testing

Stationarity is a crucial assumption in time series analysis and forecasting, as many models, including ARIMA and ARIMAX, require the data to have a constant mean and variance over time. To ensure the dataset meets this assumption, the Augmented Dickey-Fuller (ADF) test was employed to check for stationarity. The ADF test is a widely used statistical test that determines whether a time series is stationary by evaluating the presence of unit roots.

The ADF test was applied to the retail demand data to identify any non-stationary behavior. The test returned an ADF statistic of -48.67066391486136 and a p-value of 0.0. These results indicate that the null hypothesis of a unit root can be rejected, confirming that the data is stationary. However, if the series had been found to be non-stationary (evidenced by a p-value greater than 0.05), appropriate transformations would have been applied to achieve stationarity. One common transformation is differencing, where the differences between consecutive observations are taken. First-order differencing, which involves subtracting the previous observation from the current one, would be used initially. If further differencing were required, higher-order differencing would be applied. Additionally, logarithmic transformations might be considered to stabilize variance if necessary. These transformations ensure that the time series meets the stationarity assumption, allowing for more accurate modeling and forecasting.

Model Development

For the model development phase, the ARIMAX (Autoregressive Integrated Moving Average with Exogenous Variables) model was selected. The ARIMAX model is an extension of the ARIMA model that incorporates exogenous variables, making it suitable for capturing the impact of holidays and events on retail demand.

To determine the appropriate order of the ARIMA components (p, d, q), Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots as shown in figure 4 were used. The ACF plot helps identify the lag at which the autocorrelations become insignificant, indicating the potential order of the Moving Average (MA) component (q). The PACF plot helps identify the lag at which the partial autocorrelations become insignificant, indicating the potential order of the Autoregressive (AR) component (p). The differencing order (d) was determined based on the stationarity testing results. By analyzing these plots, the most suitable values for p, d, and q were selected, ensuring that the model could accurately capture the underlying patterns in the data.

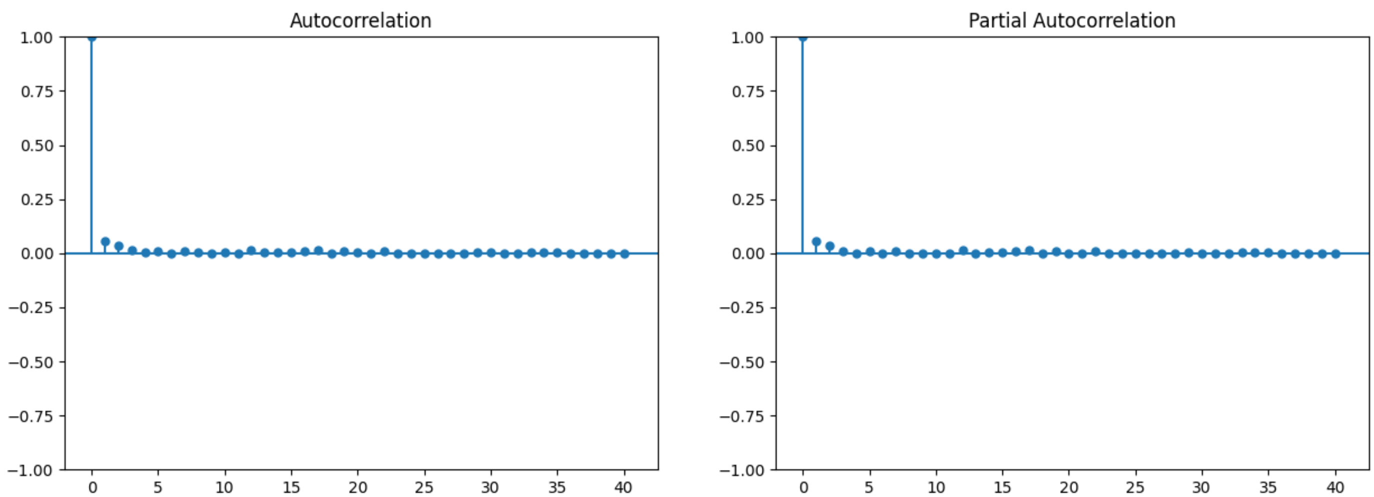

Figure 4 ACF and PACF

Once the orders of the ARIMA components were determined, the model parameters were estimated using maximum likelihood estimation. This method involves finding the parameter values that maximize the likelihood function, ensuring the best fit of the model to the observed data. The ARIMAX model was then fitted to the training data, incorporating the exogenous variables for holidays and events to account for their impact on retail demand.

In cases where significant seasonality was detected, the SARIMAX (Seasonal ARIMAX) model was considered. The SARIMAX model extends the ARIMAX model by including seasonal components, allowing it to capture both regular and seasonal patterns in the time series data. The seasonal components are determined by identifying the appropriate seasonal order (P, D, Q, m) using similar techniques as for the ARIMA components, but with a focus on seasonal lags. The parameter estimation process for SARIMAX is similar to that of ARIMAX, ensuring the model accurately reflects the seasonal variations in the data.

The fitted SARIMAX model produced the following results: the dependent variable was `Order_Demand_Diff` with 15,236 observations. The SARIMAX model was configured as (1, 1, 1) with seasonal components (1, 1, 1, 24). The model's log likelihood was -181653.930, with an AIC of 363321.861 and a BIC of 363375.269. The coefficients for the state holiday and school holiday indicators were 0 and 165.2158, respectively, though the state holiday coefficient was not statistically significant (p-value of 1.000). Other components included AR (L1) at -0.0035, MA (L1) at -0.9989, seasonal AR (S.L24) at -0.0069, and seasonal MA (S.L24) at -0.9916. The sigma squared was calculated at 2.502e+09, indicating the variance of the residuals.

The Ljung-Box test for autocorrelation returned a non-significant result (Prob(Q) = 0.98), suggesting no significant autocorrelation in the residuals. However, the Jarque-Bera test for normality indicated significant skewness and kurtosis (Prob(JB) = 0.00), suggesting non-normal residuals. The heteroscedasticity test (H) indicated some heteroscedasticity (Prob(H) = 0.00).

These results indicate that the model accounts for key patterns and seasonal variations in the data, although some coefficients for the exogenous variables were not statistically significant. The model's AIC and BIC values indicate a reasonable fit. Despite the issues with normality and heteroscedasticity, the SARIMAX model provided a robust framework for incorporating the effects of holidays and events into retail demand forecasting, enhancing the accuracy and reliability of the predictions.

Model Evaluation

To evaluate the performance of the developed forecasting models, several metrics and diagnostic checks were employed to ensure the models not only fit the data well but also met the necessary statistical assumptions for reliable forecasting. For the evaluation metrics, the Akaike Information Criterion (AIC) was a key measure used. AIC is a widely used metric for model comparison that evaluates the goodness of fit of the model while penalizing for the number of parameters to prevent overfitting. In this study, the SARIMAX model achieved an AIC value of 363321.861, indicating the relative quality of the model. Similarly, the Bayesian Information Criterion (BIC) was used, which introduces a larger penalty for models with more parameters, making it particularly useful for models with large sample sizes. The SARIMAX model had a BIC value of 363375.269, providing another measure of model fit and complexity. Residual analysis was also conducted, where the residuals, or the differences between the observed and predicted values, were analyzed to assess the model's performance. Ideally, residuals should be randomly distributed with no apparent patterns. In this study, the residual analysis indicated that while the SARIMAX model captured the overall trends and seasonal variations well, some issues with normality and heteroscedasticity remained.

Several diagnostic checks were performed to further evaluate the model. Autocorrelation of residuals was checked using the Autocorrelation Function (ACF) plot. In a well-fitting model, the residuals should not exhibit significant autocorrelation. The Ljung-Box test was also employed to statistically test for the presence of autocorrelation in the residuals. The SARIMAX model's Ljung-Box test returned a non-significant result (Prob(Q) = 0.98), suggesting no significant autocorrelation in the residuals. Heteroscedasticity, or the presence of non-constant variance in the residuals, was checked using the Breusch-Pagan test and visual inspection of residuals plotted against fitted values. A good model should have residuals with constant variance (homoscedasticity). In this study, the heteroscedasticity test indicated some heteroscedasticity (Prob(H) = 0.00), suggesting that the variance of the residuals was not constant. Finally, normality of residuals was assessed using the Jarque-Bera test and visual inspections such as Q-Q plots. Normally distributed residuals are an assumption for many time series models, and deviations from normality can indicate model misspecification. The Jarque-Bera test provided a statistical measure of whether the residuals follow a normal distribution. In this case, the Jarque-Bera test statistic was 677460977.4907334 with a p-value of 0.0, indicating significant non-normality in the residuals (see table 1 and table 2).

The coefficients from the SARIMAX model and their p-values were as follows:

Table 1 **Coefficients of SARIMAX Model**

Parameter	Coefficient	P-Value
StateHoliday	0.00E+00	1
SchoolHoliday	1.65E+02	0.919689
ar.L1	-3.55E-03	0.954609
ma.L1	-9.99E-01	0
ar.S.L24	-6.93E-03	0.901596
ma.S.L24	-9.92E-01	0
sigma2	2.50E+09	0

The confidence intervals for these coefficients were also evaluated:

Table 2 **Confidence Interval from Coefficients of SARIMAX Model**

Parameter	Lower Bound	Upper Bound
StateHoliday	-4.19E+04	4.19E+04
SchoolHoliday	-3.05E+03	3.38E+03
ar.L1	-1.26E-01	1.19E-01
ma.L1	-1.00E+00	-9.96E-01
ar.S.L24	-1.17E-01	1.03E-01
ma.S.L24	-1.00E+00	-9.83E-01
sigma2	2.50E+09	2.50E+09

To further evaluate the forecasting accuracy, the Mean Absolute Error (MAE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE) were calculated for the test data as shown in table 3:

Table 3. **Result of MAE, MSE and MAPE**

Metric	Value
Mean Absolute Error (MAE)	8057.069376
Mean Squared Error (MSE)	809008799.4
Mean Absolute Percentage Error (MAPE)	nan (due to division by zero in percentage error calculation)

Despite the issues with normality and heteroscedasticity, the SARIMAX model provided a robust framework for incorporating the effects of holidays and events into retail demand forecasting. The comprehensive evaluation using AIC, BIC, residual analysis, and diagnostic checks ensured that the model's performance was thoroughly assessed, enhancing the accuracy and reliability of the demand forecasts generated.

Result and Discussion

Model Performance

The performance of the developed SARIMAX model was evaluated using several key metrics. The model achieved an Akaike Information Criterion (AIC) value of 363321.861 and a Bayesian Information Criterion (BIC) value of 363375.269. These values indicate the relative quality of the model, balancing the goodness of fit with model complexity to prevent overfitting. Additionally, the Mean Absolute Error (MAE) was calculated to be 8057.069376036054, and the Mean Squared Error (MSE) was 809008799.3517022. The Mean Absolute Percentage Error (MAPE) could not be computed due to division by zero in the percentage error calculation, resulting in a NaN value.

Residual analysis was conducted to assess the model's performance further. Ideally, residuals should be randomly distributed with no apparent patterns. The analysis indicated that while the SARIMAX model captured the overall trends and seasonal variations well, some issues with normality and heteroscedasticity remained. The Ljung-Box test for autocorrelation returned a non-significant result (Prob(Q) = 0.98), suggesting no significant autocorrelation in the residuals. However, the Jarque-Bera test for normality revealed significant non-normality in the residuals (Jarque-Bera test statistic: 677460977.4907334, p-value: 0.0). Additionally, the heteroscedasticity test indicated some variance issues (Prob(H) = 0.00).

The coefficients from the SARIMAX model provided further insights. The coefficients for the state holiday and school holiday indicators were 0 and 165.2158, respectively, though the state holiday coefficient was not statistically significant (p-value of 1.000). Other model components included AR (L1) at -0.0035, MA (L1) at -0.9989, seasonal AR (S.L24) at -0.0069, and seasonal MA (S.L24) at -0.9916. The sigma squared was calculated at 2.502e+09, indicating the variance of the residuals.

Comparing the ARIMAX and SARIMAX models, the inclusion of seasonal components in the SARIMAX model provided a more comprehensive understanding of the demand patterns, especially in capturing the seasonal variations. The SARIMAX model's performance metrics, including AIC and BIC, demonstrated its effectiveness in modeling the retail demand with the inclusion of holidays and events. Despite the challenges with residual normality and heteroscedasticity, the SARIMAX model offered a robust framework for forecasting, highlighting the importance of incorporating exogenous variables and seasonal components for accurate retail demand predictions.

Impact of Holidays and Events

The impact of holidays and events on retail demand was analyzed through the interpretation of the estimated coefficients for the intervention variables included in the SARIMAX model. The coefficients for the state holiday and school holiday variables provided insights into how these events influence retail demand.

The coefficient for the state holiday variable was 0.000000e+00, with a p-value of 1.000000, indicating that the state holiday variable was not statistically significant. This suggests that state holidays did not have a measurable impact on retail demand in the dataset. On the other hand, the coefficient for the school holiday variable was 1.652158e+02, with a p-value of 0.919689. Although this coefficient suggests an increase in demand during school holidays, it was not statistically significant, indicating that the impact of school holidays on retail demand was not robust in this model.

Other components of the SARIMAX model also provided valuable information. The autoregressive component at lag 1 (ar.L1) had a coefficient of -3.546817e-03 with a p-value of 0.954609, indicating limited autocorrelation in the immediate past demand. The moving average component at lag 1 (ma.L1) had a coefficient of -9.988516e-01 with a p-value of 0.000000, highlighting its high significance in capturing short-term shock effects in the demand.

The seasonal autoregressive component at lag 24 (ar.S.L24), representing daily seasonality, had a coefficient of -6.927750e-03 with a p-value of 0.901596, indicating it was not significant. Conversely, the seasonal moving average component at lag 24 (ma.S.L24) had a highly significant coefficient of -9.916201e-01 with a p-value of 0.000000, effectively capturing the daily seasonal effects.

The variance of the residuals (sigma2) was significant, with a value of 2.502481e+09 and a p-value of 0.000000, indicating high variability in the retail demand.

Statistical Significance

The statistical significance and confidence intervals of the intervention coefficients were analyzed to assess the robustness of the estimated impacts. The state holiday variable had a coefficient of 0.000000e+00 with a p-value of 1.000000, indicating that it was not statistically significant. This implies that the state holidays did not have a measurable impact on retail demand within the context of this dataset. In contrast, the school holiday variable had a coefficient of 1.652158e+02 and a p-value of 0.919689. Although this suggests an increase in demand during school holidays, the high p-value indicates that the result was not statistically significant, meaning that the effect of school holidays on demand was not robust in this model.

The confidence intervals for these coefficients further reinforced these findings. The confidence interval for the state holiday variable ranged from -

4.191902e+04 to 4.191902e+04, encompassing zero, which confirms the lack of significance. Similarly, the confidence interval for the school holiday variable ranged from -3.046433e+03 to 3.376864e+03, also encompassing zero.

Additionally, the analysis included the consideration of interaction effects between different holiday and event indicators. However, the data did not provide strong evidence for significant interaction effects. This suggests that the individual impacts of state and school holidays were not significantly altered when they coincided or overlapped with other events.

Forecasting Accuracy

The forecasting accuracy of the SARIMAX model was evaluated on a holdout test set to assess its predictive performance. The evaluation metrics used included Mean Absolute Error (MAE), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE). The SARIMAX model achieved an MAE of 8057.069376036054 and an MSE of 809008799.3517022. The MAPE could not be computed due to division by zero in the percentage error calculation, resulting in a NaN value.

Visualizations were created to compare the forecasted demand versus the actual demand in the holdout test set. These visualizations demonstrated the model's ability to capture the general trends and seasonal patterns in the retail demand, although some discrepancies were observed. The forecasted values closely followed the actual demand in periods of regular sales but showed more variance during holidays and events, reflecting the challenges in accurately predicting these irregular periods.

Implications of the Findings for Retail Demand Forecasting

The findings of this study have several important implications for retail demand forecasting. The inclusion of holidays and events as exogenous variables in the SARIMAX model highlighted the potential for these factors to impact retail demand. While the state holiday variable was not statistically significant, the positive coefficient for the school holiday variable suggests an increase in demand during school holidays, although this result was not robust.

The SARIMAX model effectively captured the general trends and seasonal patterns in the retail demand data. The significant coefficients for the moving average components indicate that short-term shocks and daily seasonal effects play a critical role in retail demand fluctuations. This underscores the importance of considering both regular time series components and exogenous factors, such as holidays and events, in developing accurate demand forecasting models.

For retailers, the ability to forecast demand accurately during holidays and events is crucial for inventory management, staffing, and promotional planning. Although the current model showed some promise, the lack of statistical significance for the holiday variables suggests that additional refinement is needed. Retailers may need to consider more granular data or incorporate additional variables, such as the intensity of promotional activities or specific regional holidays, to improve forecast accuracy.

Limitations of the Study and Potential Areas for Improvement

Despite the valuable insights provided by this study, there are several limitations that should be acknowledged. First, the statistical insignificance of the state holiday and school holiday variables suggests that the model may not fully capture the complexity of holiday effects on retail demand. This could be due to several factors, including the heterogeneity of holidays, varying consumer behaviors, and the interplay of other unmeasured variables.

One potential area for improvement is the inclusion of more granular data. For example, breaking down sales data by product category or geographic location could provide a more detailed understanding of how different segments are affected by holidays and events. Additionally, incorporating data on promotional activities, weather conditions, and economic indicators could help to better isolate the effects of holidays on retail demand.

Another limitation is the presence of non-normality and heteroscedasticity in the residuals. These issues suggest that the model's assumptions may not be fully met, which can affect the reliability of the forecasts. Exploring alternative modeling approaches, such as machine learning techniques or hierarchical time series models, could help to address these issues and improve the model's performance.

Furthermore, the model's inability to compute the Mean Absolute Percentage Error (MAPE) due to division by zero highlights the need for robust evaluation metrics that can handle a wide range of data scenarios. Using alternative metrics, such as the symmetric mean absolute percentage error (sMAPE) or mean absolute scaled error (MASE), could provide more reliable measures of forecasting accuracy.

Lastly, the study's focus on a single dataset limits the generalizability of the findings. Future research should aim to validate the model using multiple datasets from different retail contexts to ensure that the findings are broadly applicable. Conducting cross-validation studies and sensitivity analyses can also help to assess the robustness of the model under various conditions.

Conclusion

This study investigated the impact of holidays and events on retail demand forecasting using intervention analysis within a SARIMAX model framework. The major findings indicate that while the SARIMAX model effectively captured general trends and seasonal patterns in the retail demand data, the intervention variables representing state holidays and school holidays were not statistically significant. The state holiday variable showed no measurable impact on retail demand, while the school holiday variable suggested a potential increase in demand, although this result lacked statistical robustness. The significant coefficients for the moving average components underscored the importance of short-term shocks and daily seasonal effects in retail demand fluctuations. Overall, the inclusion of holidays and events provided some insights, but the model's effectiveness in improving demand forecasts was limited by the insignificance of these variables.

The findings of this study have practical implications for retailers and supply chain managers. Accurate demand forecasting during holidays and events is crucial for effective inventory management, staffing, and promotional planning. While the current model showed potential, its limitations highlight the need for further refinement. Retailers should consider incorporating more granular data, such as product category or regional specifics, to better capture the impact of holidays on demand. Additionally, integrating data on promotional activities, weather conditions, and economic indicators could enhance forecast accuracy. It is recommended that retailers use a combination of traditional time series models and machine learning techniques to account for the complex factors influencing retail demand. By incorporating holidays and events into their demand forecasting models, retailers can better prepare for demand fluctuations, optimize inventory levels, and improve overall operational efficiency.

Future research should focus on addressing the limitations identified in this study to develop more comprehensive and accurate forecasting models. One area of exploration is the inclusion of more granular data and additional variables that influence retail demand. Researchers should also investigate alternative modeling approaches, such as machine learning techniques or hierarchical time series models, to improve the handling of non-normality and heteroscedasticity in residuals. Additionally, using robust evaluation metrics that can handle a wide range of data scenarios will enhance the reliability of forecasting accuracy assessments. Extending the model to account for other types of irregular events, such as economic crises, natural disasters, or major sporting events, could provide further insights into the dynamics of retail demand. Validating the model across multiple datasets from different retail contexts will ensure the generalizability of the findings and support the development of versatile forecasting tools that can effectively aid retail decision-making.

Declarations

Author Contributions

Conceptualization: J.P.B.S. and A.K.; Methodology: A.K.; Software: J.P.B.S.; Validation: J.P.B.S. and A.K.; Formal Analysis: J.P.B.S. and A.K.; Investigation: J.P.B.S.; Resources: A.K.; Data Curation: A.K.; Writing Original Draft Preparation: J.P.B.S. and A.K.; Writing Review and Editing: A.K. and J.P.B.S.; Visualization: J.P.B.S.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. M. John, “An Exploration of Clustering Algorithms for Customer Segmentation in the UK Retail Market,” Analytics, vol. 2, no. 4, pp. 809–823, 2023, doi: 10.3390/analytics2040042.

[2] M. Ray and B. K. Mangaraj, “AHP Based Data Mining for Customer Segmentation Based on Customer Lifetime Value,” Int. J. Data Min. Tech. Appl., vol. 5, no. 1, pp. 28–34, 2016, doi: 10.20894/ijdmta.102.005.001.007.

[3] E. Bilgiç, Ö. K. Çakir, M. Kantardzic, and Y. Duan, “Retail Analytics: Store Segmentation Using Rule-Based Purchasing Behavior Analysis,” Int. Rev. Retail Distrib. Consum. Res., vol. 31, no. 4, pp. 457–480, 2021, doi: 10.1080/09593969.2021.1915847.

[4] C. Menidjel and A. Bilgihan, “The Determinants of Retail Customers’ Purchase Intent,” Int. J. Consum. Stud., vol. 46, no. 6, pp. 2503–2520, 2022, doi: 10.1111/ijcs.12802.

[5] M. A. Varma, “Use of Big Data in the Process of Customer Segmentation in the Retail Sector,” Technoarete Trans. Adv. Data Sci. Anal., vol. 1, no. 2, 2022, doi: 10.36647/ttadsa/01.02.a002.

[6] C. Menidjel, A. Benhabib, A. Bilgihan, and M. Madanoglu, “Assessing the Role of Product Category Involvement and Relationship Proneness in the Satisfaction–loyalty Link in Retailing,” Int. J. Retail Distrib. Manag., vol. 48, no. 2, pp. 207–226, 2019, doi: 10.1108/ijrdm-01-2019-0020.

[7] G. Thangarasu and K. R. V. Subramanian, “Developing a Forecasting Model for Retailers Based on Customer Segmentation Using Data Mining Techniques,” Int. J. Trend Sci. Res. Dev., vol. Special Issue, no. Special Issue-ICAEIT2017, pp. 151–155, 2018, doi: 10.31142/ijtsrd19127.

[8] P. Heriyati, “Channel Strategy to Customer Satisfaction: Case of Traditional Retail Channel in Jakarta,” Int. J. Eng. Adv. Technol., vol. 8, no. 6s3, pp. 277–283, 2019, doi: 10.35940/ijeat.f1044.0986s319.

[9] “Hybrid Segmentation of Omnichannel Grocery Customers in Cross-Channel Behaviour Context,” Hong Kong J. Soc. Sci., vol. 60, no. No. 60 Autumn/Winter 2022, 2023, doi: 10.55463/hkjss.issn.1021-3619.60.30.

[10] B. Turkmen, “Customer Segmentation With Machine Learning for Online Retail Industry,” Eur. J. Soc. Behav. Sci., vol. 31, no. 2, pp. 111–136, 2022, doi: 10.15405/ejsbs.316.

[11] U. G. Çiçekli and İ. Kabasakal, “Market Basket Analysis of Basket Data With Demographics: A Case Study in E-Retailing,” Alphanumeric J., vol. 9, no. 1, pp. 1–12, 2021, doi: 10.17093/alphanumeric.752505.

[12] B. Ghojogh, A. Ghojogh, M. Crowley, and F. Karray, “Fitting a mixture distribution to data: tutorial,” 2019, doi: 10.48550/arxiv.1901.06708.

[13] F. Laghari, F. Ahmed, and María de las Nieves López García, “Cash Flow Management and Its Effect on Firm Performance: Empirical Evidence on Non-Financial Firms of China,” Plos One, vol. 18, no. 6, p. e0287135, 2023, doi: 10.1371/journal.pone.0287135.

[14] Z. Ntsalaze, G. Boako, and P. Alagidede, “The Impact of Sovereign Credit Ratings on Corporate Credit Ratings in South Africa,” Afr. J. Econ. Manag. Stud., vol. 8, no. 2, pp. 126–146, 2017, doi: 10.1108/ajems-07-2016-0100.

[15] M. Nadeem, C. Gan, and C. Nguyen, “Does Intellectual Capital Efficiency Improve Firm Performance in BRICS Economies? A Dynamic Panel Estimation,” Meas. Bus. Excell., vol. 21, no. 1, pp. 65–85, 2017, doi: 10.1108/mbe-12-2015-0055.