

Forecasting Beer Demand at Anadolu Efes: A Predictive Analytics Case Report

August 3, 2024

Hunter Berberich

Executive Summary

Efes Beverage Group, a prominent player in Turkey's beer market, aims to improve its demand forecasting methods to enhance operational efficiency and strategic planning (Köksalan, Özpeynirci, & Süral, 2010). Historically, the company relied on subjective sales input but now seeks a more systematic approach by considering prices, tourism, and seasonality. Analyzing data from 1987 to 1993, significant trends in monthly demand were identified. Explanatory regression models were developed to understand the key factors influencing beer demand, including stepwise and polynomial regression.

Multiple linear regression and exponential smoothing models were employed for prediction purposes. The time-lagged multiple linear regression model used historical data and lagged variables to predict future demand, significantly improving accuracy with an R-squared value of 0.9818 and a root mean square error of 2.94 million liters. The Holt-Winters exponential smoothing model addressed level, trend, and seasonality components, providing reliable forecasts.

Implementing these methods will allow Efes Beverage Group to improve inventory management, reduce stockouts and overstock situations, and enhance customer satisfaction and financial performance (Kros & Keller, 2010). The report outlines the steps taken in the analysis, the rationale behind choosing these models, and the anticipated benefits of adopting these forecasting methods. These steps are anticipated to improve forecasting accuracy and marketing effectiveness significantly, optimizing operational outcomes and competitive positioning in the market.

Contents

Introduction.....	3
Data Exploration	3
Explanatory Regression Model Development and Variable Selection	4
Predictive Model Development and Variable Selection	8
References.....	12
Appendix.....	13

List of Tables

Table 1. Model Comparison.....	5
Table 2. Variable Significance of Final Model	6
Table 3. Performance Metrics Comparison.....	9

List of Figures

Figure 1. Monthly Trend Analysis	3
Figure 2. Average Beer Price by Season	4
Figure 3. Assumptions for Final Model	7
Figure 4. Final Comparison Plot.....	10
Figure 5. Percent Error Plot	10

Introduction

This section examines beer demand for Efes Beverage Group by developing robust explanatory and predictive models. We aim to identify critical factors influencing beer consumption and provide reliable forecasts for future demand. Using multiple regression techniques, we explore the relationships between beer demand and predictors like prices, tourist numbers, and seasonal effects. Additionally, we evaluate advanced time series forecasting methods, including Multiple Linear Regression with time-lagged data and Holt-Winters exponential smoothing, to determine the most effective approach. We aim to offer actionable insights to improve Efes's demand forecasting, inventory management, and strategic decision-making.

Data Exploration

The following visualizations comprehensively analyze beer consumption trends, prices, and tourist activity from 1987 to 1993. The combined monthly trend plot (Figure 1) cohesively shows the Monthly Beer Demand, Monthly Beer Price, and Monthly Total Tourists.

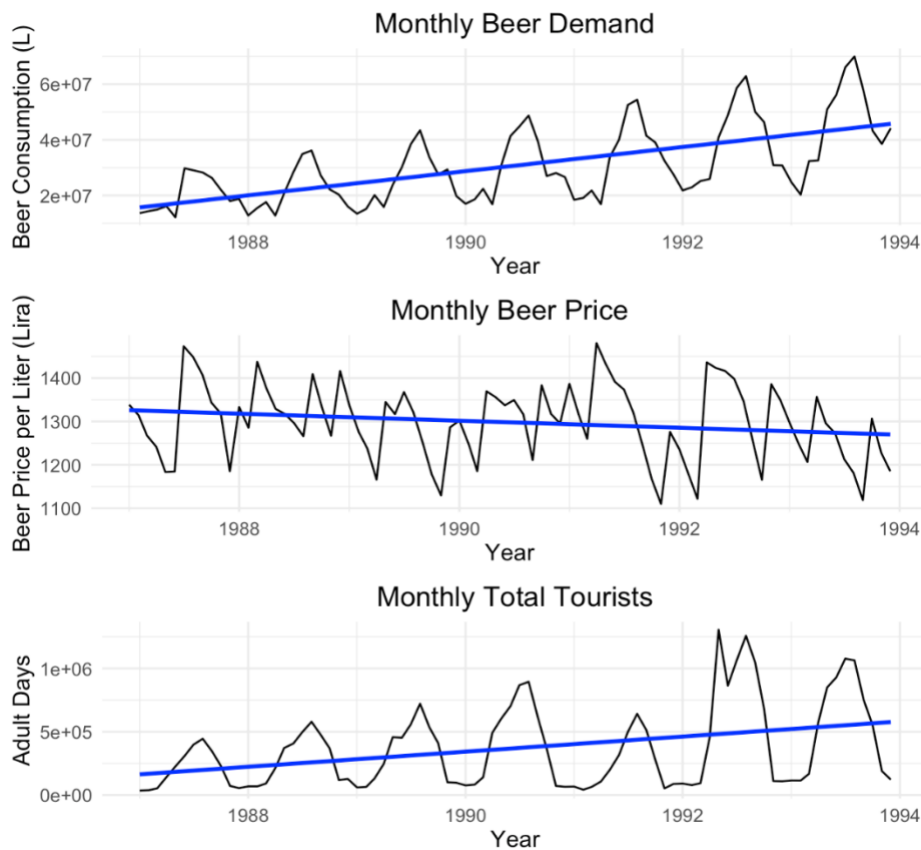


Figure 1. Monthly Trend Analysis

Despite seasonal fluctuations, this plot shows a clear upward trend in beer consumption over the years, with a slight downward trend in beer prices, indicating potential competitive pricing or economic factors at play. Additionally, the cyclical nature of tourist activity is evident, with peaks during certain months each year. The second plot (Figure 2), Average Beer Consumption

by Season, reveals that beer consumption is highest in the summer and fall, which aligns with peak tourist seasons.

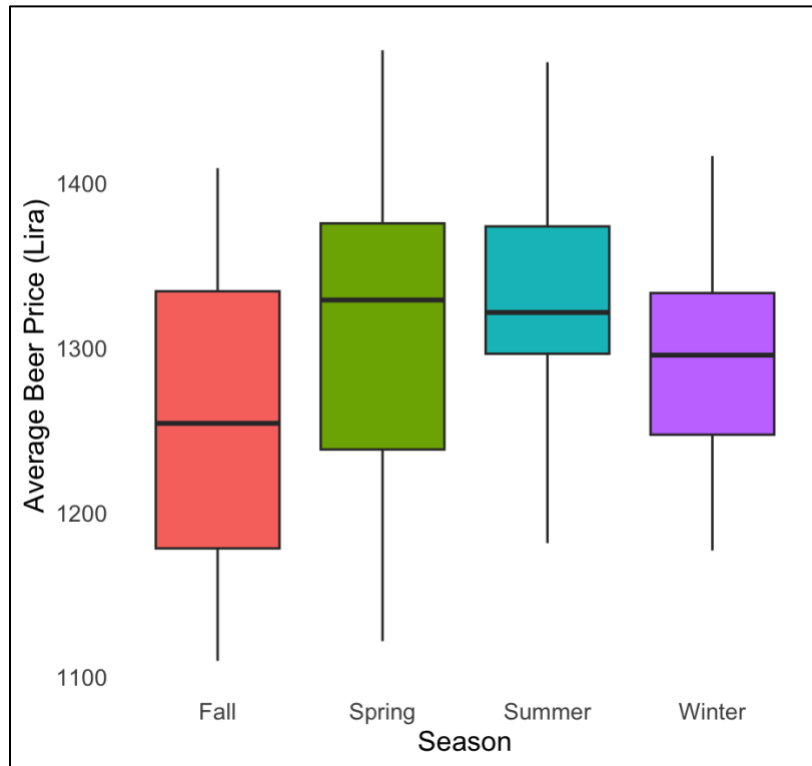


Figure 2. Average Beer Price by Season

These insights suggest a correlation between increased tourist activity and higher beer consumption while considering the impact of pricing trends. Such detailed analysis is crucial for Anadolu Efes to effectively strategize its production, marketing, and pricing decisions to optimize sales and cater to peak tourist seasons.

Explanatory Regression Model Development and Variable Selection

Simple Regression Model

A multiple linear regression model was developed to explain beer demand. A simple model was built using Total Tourist and Average Beer Price as predictors. This basic model had a high R-squared value, indicating strong explanatory power. However, the residual analysis indicated some non-random patterns, suggesting the need for a more comprehensive model.

Basic Model with Seasons

A model that includes seasonal effects was developed to capture potential seasonal effects and refine the initial model. This model included all variables and seasons.

Stepwise Linear Regression

Stepwise regression using the stepAIC function from the MASS package was applied to refine the basic model. This method adds and removes predictors based on the Akaike Information Criterion (AIC), balancing model fit and complexity. Both forward selection and backward elimination were used to ensure that only the most significant variables were included. The final model included Date, Average Beer Price, Average Canned Beer Price, tourist numbers from Czechoslovakia, the United Kingdom, and France, the 'Others Total' variable, and seasonal effects. This refinement improved simplicity, interpretability, and adjusted R-squared value while lowering the residual standard error, indicating a better fit.

Polynomial Regression

A polynomial regression model was tested to capture potential non-linear relationships. This model included squared terms for predictors like Average Canned Soft Drink Price and Czechoslovakia. The polynomial model achieved a higher R-squared value, indicating it could explain more variability in the data. However, it exhibited significant multicollinearity issues, as evidenced by high Variance Inflation Factors (VIFs), particularly for the squared terms. This redundancy among predictors compromised the model's reliability and increased its complexity without providing a proportional improvement in performance.

Model Comparison and Performance

The performance of each regression model was evaluated based on R-squared, Adjusted R-squared, and Variance Influence Factors. Performance metrics are shown in Table 1.

Table 1. Model Comparison

<i>Model</i>	<i>R-Squared</i>	<i>Adjusted R-Squared</i>	<i>VIF</i>
<i>Basic</i>	0.9389	0.9254	Moderate
<i>Stepwise</i>	0.9359	0.9271	Low to Moderate
<i>Polynomial</i>	0.9504	0.9395	High

Final Model Selection

The stepwise regression model was selected as the final model due to its balance between simplicity and explanatory power. It has fewer multicollinearity issues than the polynomial model. This model effectively captures the key factors influencing beer demand, making it robust for explanatory purposes.

The final equation is represented below (Equation 1), followed by a summary of the predictive factors (Jaggia et al., 2023):

Equation 1. Final regression model

<p>Beer Consumption</p> $= -29,260,000 + 9,000(\text{Date}) - 24,180(\text{Average Beer Price})$ $+ 5,802(\text{Average Canned Beer Price}) + 33.43(\text{Czechoslovakia})$ $+ 12.58(\text{United Kingdom}) - 45.68(\text{France}) + 26.75(\text{Others Total})$ $+ 653,800(\text{Season Spring}) + 10,260,000(\text{Season Summer})$ $+ 2,799,000(\text{Season Fall})$
--

Table 2. Variable Significance of Final Model

<i>Variable</i>	<i>Impact</i>	<i>P-value</i>
<i>Date</i>	Positive	2.36e-08 ***
<i>Average Beer Price</i>	Negative	0.000115 ***
<i>Average Canned Beer Price</i>	Positive	0.058809
<i>Czechoslovakia</i>	Positive	0.010945 *
<i>United Kingdom</i>	Positive	0.080404
<i>France</i>	Negative	0.000161
<i>Others</i>	Positive	5.10e-10 ***
<i>Season (Summer)</i>	Highest Positive Impact	5.09e-07 ***
<i>Season (Spring)</i>	Negative	0.157220
<i>Season (Winter)</i>	Negative	0.065644

Checking for Assumptions

To ensure the validity of the regression models, we checked several key assumptions across all four models (Simple, Basic with Seasons, Stepwise, and Polynomial). The linearity assumption was met as indicated by the residuals versus fitted values plots (Figure 3), which showed no clear pattern. The independence of residuals was confirmed by the lack of significant autocorrelation in the residual plots over time. Q-Q plots demonstrated that the residuals were normally distributed, meeting the normality assumption. The assumption of homoscedasticity was satisfied as scale-location plots showed no significant pattern. Finally, Cook's distance plots indicated a few potential outliers but did not significantly impact the models.

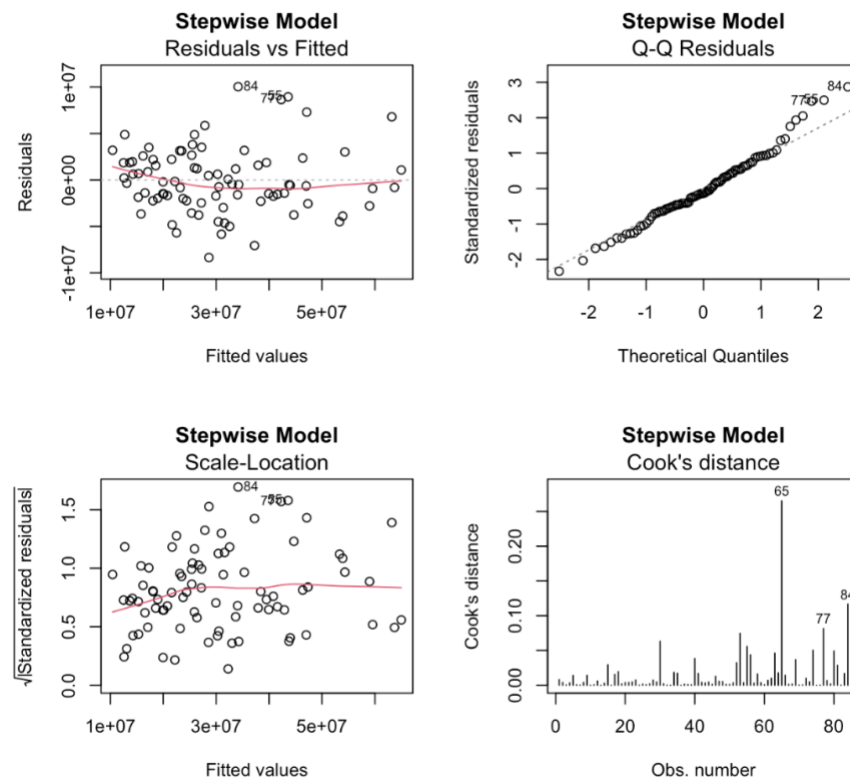


Figure 3. Assumptions for Final Model

Explanatory Regression Model Insights

The positive coefficients for variables such as Average Beer Price and tourist numbers from specific countries highlight their significant impact on beer demand. The negative impact of variables such as France and the Average Canned Soft Drink Price indicates areas that might require strategic adjustments.

Knowing these results can help Efes Beverage Group make better business decisions, such as optimizing pricing strategies, targeting specific tourist markets, and planning for seasonal fluctuations in demand. By understanding the factors that drive beer demand, Efes can improve inventory management, marketing effectiveness, and overall operational efficiency.

Recommendations:

- **Leverage Seasonal Trends:** Focus marketing and sales efforts during the summer when beer demand is highest.
- **Targeted Marketing:** Tailor marketing strategies to attract tourists from countries with a positive impact on beer demand.
- **Dynamic Pricing:** Adjust beer prices to optimize sales while considering the negative impact of price increases on demand.
- **Inventory Management:** Use predictive insights to manage inventory levels better, reducing the risk of stockouts or overstock situations.

By implementing these recommendations, Efes Beverage Group can significantly enhance forecasting accuracy and strategic planning, improving business outcomes.

Predictive Model Development and Variable Selection

The insights gained from the explanatory regression model informed the development of the three predictive model candidates using two approaches: Multiple-Linear Regression and Holt-Winters Exponential Smoothing. All models were trained on data from 1987-1992 and evaluated using out-of-sample testing against 1993 data. Each model's performance was assessed using Mean Absolute Percent Error (MAPE) and Mean Percent Error (MPE) to assess bias and overall accuracy.

Time-lag Predictive Regression Model

The selected variables included beer consumption for periods ($t-4$, $t-7$, $t-9$, $t-11$, $t-12$); canned beer price for periods ($t-10$, $t-11$, $t-12$); and total tourists for periods ($t-3$, $t-5$, $t-6$, $t-7$, $t-9$). The R-squared was 0.9755 for the training data and 0.9818 for the testing data. Compared to the baseline model, adding time-lagged variables improved the test set R-squared by 0.1046 (12%) and the RMSE (Root Mean Square Error) by 7.4 million liters (74%).

The regression model highlighted key factors influencing beer demand, such as historical consumption, prices, and tourist numbers. Including time-lagged variables significantly improved forecasting accuracy, providing Efes Beverage Group with a robust strategic planning and inventory management tool.

Triple Exponential Smoothing

For our predictive exponential smoothing model development for Anadolu Efes, the Triple Exponential Smoothing method, also known as the Holt-Winters model, was utilized due to its efficacy in capturing the trend and seasonality components inherent in monthly beer demand time series. The trend and seasonality components were identified by plotting the beer consumption time series from 1987 - 1993 (see Figure 1). The model was developed and implemented to forecast the monthly beer demand by addressing three key components: level, trend, and seasonality.

The Holt-Winters model operates using three equations:

- Level Equation: This smooths the series to estimate the average value at time t .
- Trend Equation: This estimates the trend in the series at time t , smoothing out irregularities.
- Seasonal Equation: This captures the repeating patterns or seasonal effects observed across the same periods in each cycle.

The predictive performance of the model was assessed using the Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) and results of these metrics were included in a comparative table with metrics from other models (see Table 3). The resulting MAPE value of 15.71% suggests that the model's forecasts were, on average, within 15.71% of the actual values, indicating a reasonably good fit for practical forecasting purposes.

Insights from Exponential Smoothing

The Holt-Winters model has proven effective in capturing the underlying trends and seasonal patterns in Anadolu Efes' beer demand data. The Triple Exponential Smoothing method's ability to account for level, trend, and seasonality makes it a robust tool for short-term demand forecasting. Key insights from this model include:

- **Improved Forecast Accuracy:** The model's accuracy metrics (e.g., M.E., RMSE, MAE, and MAPE) demonstrate its strength in providing reliable forecasts, which is crucial for strategic planning and inventory management.
- **Timely Trend Detection:** By assigning exponentially decreasing weights to past observations, the model gives more importance to recent data, ensuring timely trend detection.

Performance Evaluation and Model Selection

The Holt-Winters model's performance was evaluated alongside other predictive models, including baseline multiple linear regression (MLR) and Time-lagged MLR models. The evaluation metrics used were Mean Error (M.E.), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Percentage Error (MPE), and Mean Absolute Percentage Error (MAPE); all performance metrics were included in a comparative table for analysis (see Table 3).

Table 3. Performance Metrics Comparison

	<i>ME</i>	<i>RMSE</i>	<i>MAE</i>	<i>MPE</i>	<i>MAPE</i>
<i>Baseline MLR</i>	8,699,444	10,341,263	9,132,046	18.16	20.30
<i>Time-Lag MLR</i>	-620,083	2,939,571	2,293,019	-0.32	4.76
<i>Holt-Winters Forecast</i>	5,746,580	8,622,889	7,216,495	8.75	15.72

The comparison revealed the following:

- Baseline MLR had an RMSE of 10.34 million liters and a MAPE of 20.30%, indicating moderate predictive accuracy.
- Time-lagged MLR significantly improved accuracy by incorporating time-lagged variables, achieving an RMSE of 2.94 million liters and a MAPE of 4.76%.
- Holt-Winters Model achieved an RMSE of 8.62 million liters and a MAPE of 15.71%, demonstrating its effectiveness in capturing seasonal patterns and trends.

The Time-lagged MLR model performed best with the lowest RMSE and MAPE values, indicating superior predictive accuracy. However, the Holt-Winters model also provided valuable insights into seasonal trends and was effective for short-term forecasts. In addition to the performance metrics, the three models were plotted to show predicted vs actual beer consumption values (see Figure 4). Also, the error percentage of beer consumption predictions for all models was plotted (see Figure 5). As shown in figures 4 and 5, the Time-lagged model is more accurate in predicting beer consumption given that its predictions are closer to the actual beer consumption values.

Therefore, combining the Holt-Winters model for capturing seasonality and the Time-lagged MLR model for overall accuracy can provide a comprehensive approach to demand forecasting

for Anadolu Efes. This dual approach enables better inventory management, strategic planning, and operational efficiency, ultimately enhancing the company's competitive position in the market (Jaggia et al., 2023).

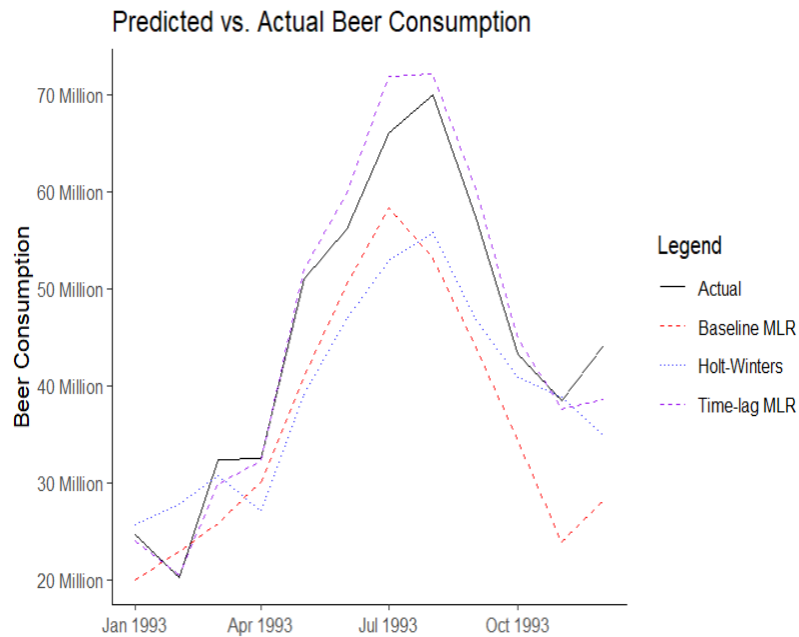


Figure 4. Final Comparison Plot

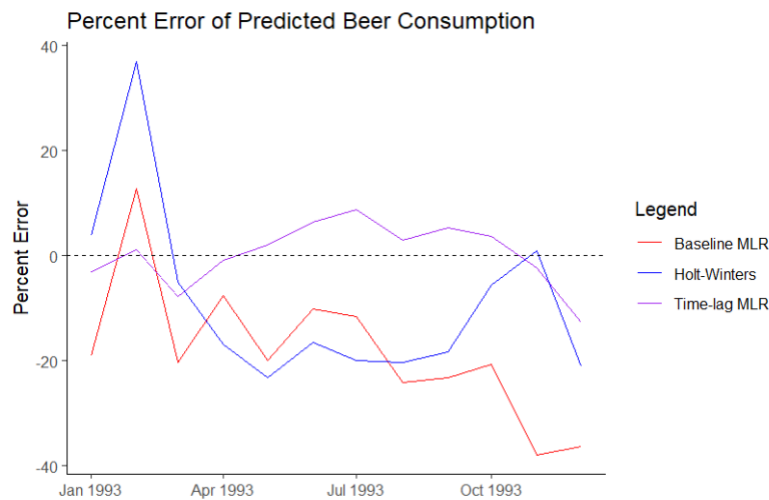


Figure 5. Percent Error Plot

Strategic Insights from Regression Model

Price Impact: A decrease in beer price is associated with increased beer demand.

Tourism Influence: The number of tourists significantly impacts beer demand, with higher tourist numbers leading to higher demand.

While these insights are crucial for strategic planning, the primary goal of Efes Beverage Group is to improve demand forecasting accuracy to enhance inventory management and operational efficiency.

Conclusion

This case study investigated multiple predictive models to forecast beer demand for Anadolu Efes, utilizing multiple linear regression and exponential smoothing techniques to enhance forecasting accuracy and reliability. Our analysis involved developing various regression models, including simple and stepwise regression and the Holt-Winters Triple Smoothing method. Each model's performance was assessed based on RMSE, MAE, and MAPE. Due to its simplicity and strong explanatory power, we selected the stepwise regression model as the final explanatory model. For prediction, the Time-lagged Multiple Linear Regression model outperformed others, achieving the lowest RMSE and MAPE values. Additionally, the Holt-Winters model effectively captured seasonal trends, making it particularly useful for short-term forecasting.

In general, forecast modeling offers significant benefits for Efes Beverage Group (Kros & Keller, 2010):

- **Improved Inventory Management:** Accurate forecasts optimize inventory levels, reducing stockouts and excess inventory, leading to cost savings.
- **Enhanced Customer Satisfaction:** Reliable forecasts ensure popular items are in stock, improving customer satisfaction and loyalty.
- **Financial Performance:** Precise demand forecasts enable better financial planning and budgeting, improving profitability.

By adopting advanced predictive modeling, Efes Beverage Group can improve operational efficiency, make informed strategic decisions, and enhance its market position.

References

1. Jaggia, S., Kelly, A., Lertwachara, K., & Chen, L. (2023). *Business Analytics: Communicating with Numbers* (2nd ed.). McGraw-Hill Higher Education. ISBN 978-1-264-30280-2.
2. Köksalan, M., Özpeynirci, S., & Süral, H. (2010). Case—Forecasting beer demand at Anadolu Efes. *INFORMS Transactions on Education*, 10(3), 142-145.
<https://doi.org/10.1287/ited.1100.0048>
3. Kros, J. F., & Keller, C. M. (2010). Seasonal regression forecasting in the U.S. beer import market. In K. D. Lawrence & R. K. Klimberg (Eds.), *Advances in Business and Management Forecasting* (Vol. 7, pp. 85-101). Emerald Group Publishing Limited.
[https://doi.org/10.1108/S1477-4070\(2010\)0000007007](https://doi.org/10.1108/S1477-4070(2010)0000007007)
4. RStudio Team. (2024). RStudio: Integrated Development Environment for R (Version 2024.04.1+748) [Computer software]. RStudio, PBC. <http://www.rstudio.com/>

Appendix

List of Tables

Table 4a.

Baseline Predictive Model Summary

<i>Parameter</i>	<i>Coefficient</i>
<i>Intercept</i>	1.562e+07
<i>Canned Soft Drink Price</i>	1.063e+04
<i>Canned Beer Price</i>	-7.310e+03
<i>Czechoslovakia</i>	6.953e+01
<i>Germany</i>	6.901e+00
<i>U.K.</i>	2.292e+01
<i>USA</i>	-5.420e+01
<i>France</i>	-6.512e+01
<i>Other</i>	3.612e+01

Description: This table provides the coefficients for the parameters in the Baseline Predictive Model, showing their impact on beer demand.

Table 4b.

Train Data Fit for Basic Predictive Model

<i>Metric</i>	<i>Value</i>
<i>Multiple R²</i>	0.8682
<i>Adjusted R²</i>	0.8515

Description: This table presents the fit metrics for the training data using the Baseline Predictive Model.

Table 4c.

Test Data Fit for Baseline Predictive Model

<i>Metric</i>	<i>Value</i>
<i>RMSE</i>	1.034e+07
<i>MAE</i>	9.132e+06
<i>R²</i>	0.8772

Description: This table presents the fit metrics for the test data using the Baseline Predictive Model.

Table 5a.

Time Lag Predictive Model Summary

<i>Parameter</i>	<i>Coefficient</i>
<i>Intercept</i>	1.254e+06
<i>Beer Price</i>	-1.949e+04
<i>Raki Price</i>	1.969e+03
<i>Canned Soft Drink Price</i>	3.993e+03
<i>Draft Beer Price</i>	-2.135e+04
<i>Beer Consumption t_{.4}</i>	-1.499e-01
<i>Beer Consumption t_{.7}</i>	1.195e-01
<i>Beer Consumption t_{.9}</i>	1.193e-01
<i>Beer Consumption t_{.11}</i>	6.251e-01
<i>Beer Consumption t_{.12}</i>	6.568e-01
<i>Canned Beer Price t_{.10}</i>	-5.365e+03
<i>Canned Beer Price t_{.11}</i>	4.295e+03
<i>Canned Beer Price t_{.12}</i>	6.100e+03
<i>Total Tourists t_{.5}</i>	1.875e+00
<i>Total Tourists t_{.6}</i>	2.407e+00
<i>Total Tourists t_{.7}</i>	-2.429e+00
<i>Total Tourists t_{.9}</i>	3.247e+00
<i>Total Tourists t_{.3}</i>	-3.656e+00

Description: This table provides the coefficients for the parameters in the Time Lag Predictive Model, showing their impact on beer demand.

Table 5b.

Train Data Fit for Time Lag Predictive Model

<i>Metric</i>	<i>Value</i>
<i>Multiple R²</i>	0.9755
<i>Adjusted R²</i>	0.9656
<i>F-statistic</i>	98.37
<i>p-value</i>	< 2.2e-16

Description: This table presents the fit metrics for the training data using the Time Lag Predictive Model.

Table 5c.

Test Data Fit for Time Lag Predictive Model

<i>Metric</i>	<i>Value</i>
<i>RMSE</i>	2.940e+06
<i>MAE</i>	2.293e+06
<i>R²</i>	0.9818

Description: This table presents the performance metrics for the test data using the Time Lag Predictive Model.