Forecasting Insights: A Statistical Analysis of Time Series Data

CONTENTS:

Effective Summary:

This study uses manual modeling, ETS, ARIMA, and a hybrid selection strategy, with a focus on Series M510, to solve the difficulties associated with time series forecasting in the glass container business. Reference points are provided via benchmarking techniques (Naive, Mean Forecasting, Holt-Winters). The Model Selection Strategy outperforms benchmarks and independent models with continual improvements in accuracy and adaptability. The disadvantages between the model's complexity and simplicity are brought to light by the critical discourse. Following thorough validation, the hybrid selection strategy—which blends ARIMA and ETS—becomes an effective and adaptable forecasting tool required for making informed decisions in quickly changing sectors.

INTRODUCTION

In the dynamic environment of the food glass container industry, understanding and forecasting trends in the supply of certain products is of utmost importance. This report begins with an in-depth review of forecasting methods, focusing on competitive M3 data. The research involves two main areas: stroke prediction, which involves an in-depth analysis of a hundred different series, and manual modelling, which focuses on a specific series called M3[[1911]]. The main objective is to provide accurate demand forecasts that enable better service delivery, optimized production schedules, targeted marketing initiatives and effective inventory management.

A rigorous effort to understand the complexities of forecast dynamics is highlighted by the choice to evaluate group forecasts using manual modelling. By exploring the nuances of a given series, the report looks for hidden patterns and structures in the data using manual approaches such as data mining, linear regression and exponential smoothing. At the same time, automated techniques and customized model selection procedures are used to evaluate the impact forecasts of several hundred series.

The ultimate objective is to provide a comprehensive understanding of forecasting methods that include the advantages of both automated and human-centred approaches. By using these two approaches, we are able to gain deeper insights that apply to a wider range of situations while also improving our comprehension of the chosen series. The report's goal is to provide readers a sophisticated grasp of time series forecasting as the narrative develops, empowering them to make wise choices with intricate and dynamic data. Achieving competitive advantage, deftly controlling prices, inventory, and service levels, and keeping up with changing consumer trends all depend on accurate forecasting, which increases demand. Accurate forecasting becomes essential for more seamless supply chains and adaptable operational planning.

By developing the newly-emerging area of decision-making procedures, the report seeks to fulfil these objectives. Combining both automatic and human forecasting, it is meant to be a helpful tool for business professionals and decision makers who are struggling with forecasting issues in a corporate environment that is changing quickly. Targeted programs that boost the impact of advertising are made possible by accurate forecasting in marketing. Furthermore, in order to survive in dynamic marketplaces, one must be able to modify action plans and make wise judgments, both of which are backed by a thorough understanding of time series dynamics.

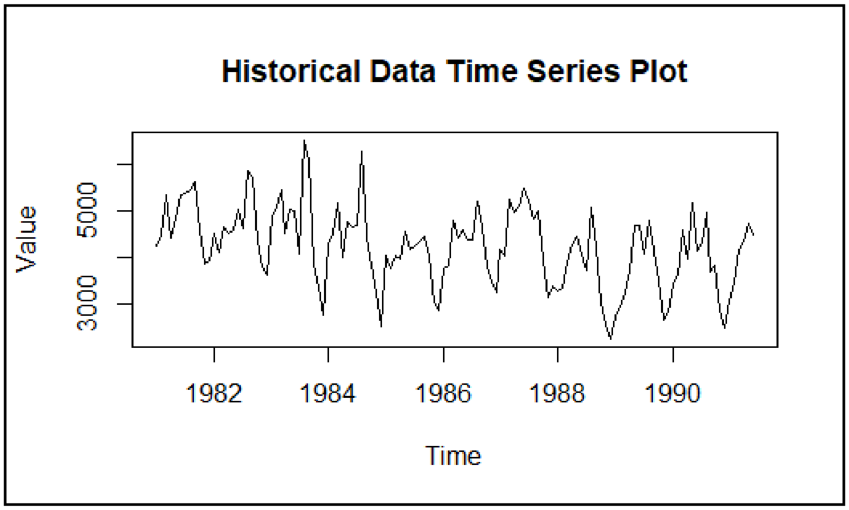
In the present unstable economic climate, competitive advantage, efficiency, and flexibility are all dependent on this forecasting process.

MANUAL MODELLING

Data Visualization and Exploration:

The investigation began with a detailed examination of the historical time series data, in particular Series M510, a monthly time series relevant to the industry. This series deals with the shipping of glass food containers with narrow necks. To give an overview of the variability and central patterns in the data, descriptive statistics were produced. A minimum production of 2266 units, a median production of 4271 units, and a maximum production of 6494 units are shown by the summary data. During this time, the monthly output average was about 4236.27 units. The middle 50% of production values are represented by the interquartile range, which ranges from 4563 to 5190. The variety and primary tendencies that are crucial for forecasting and strategic decision-making in the glass container business are shown in this review of data distribution. The investigation began with a detailed examination of the historical time series data, in particular Series M510, a monthly time series relevant to the industry. This series deals with the shipping of glass food containers with narrow necks.

Time Series Plot: For better understanding of the shipping dynamics, a time series plot provided an analysis of the behavior of the data over time. This graphic provides insight on the general temporal trends pertaining to the shipping of glass food containers. The historical time series plot provides insights into patterns and trends within the dataset by graphically illustrating the variation in values over time.



Augmented Dickey-Fuller Test:. The stationarity of the time series is evaluated using the augmented Dickey-Fuller test. The key findings from the test are as follows:

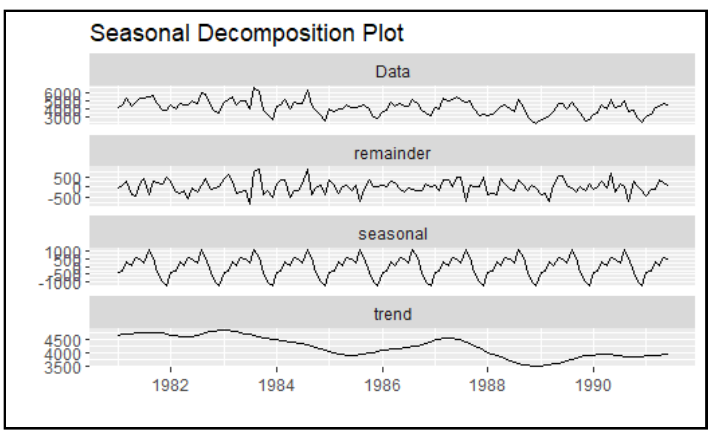
- Dickey-Fuller Test Statistic: -5.1342
- Lag Order: 4

- P-Value: 0.01

The time series is assumed to be non-stationary and to have a unit root in the ADF test's null hypothesis. According to the alternate hypothesis, the time series is stationary. We reject the null hypothesis with a p-value of 0.01, which is less than the standard significance level of 0.05. As a result, there is ample evidence for us to draw the conclusion that the historical time series is stationary. A key component of time series analysis is the time series' stationarity. Reliable forecasts may be made easier and the modeling process made simpler with the use of a stationary time series. This study establishes the groundwork for future time series modeling and forecasting while providing insightful information about the characteristics of the historical time series.

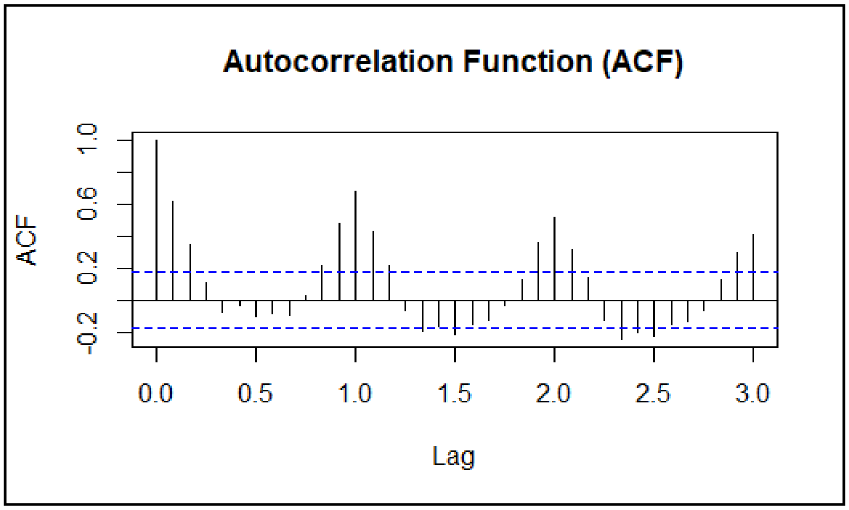
Unit Root/Cointegration Test (Phillips-Perron)
For time series data, stationarity is essential because it guarantees that the data's statistical characteristics—like mean and variance—remain consistent across time.
In the output, the value of the test statistic is: -5.9633, 11.9281, 17.8715. To assess the stationarity of the time series, this test statistic is compared against critical values at different significance levels (11.9281 and 17.8715 in this case). If the test statistic is more negative than the critical values, there is evidence to reject the null hypothesis of a unit root, suggesting that the time series is stationary.
The results of the Unit Root/Cointegration and Augmented Dickey-Fuller tests point to the stationary nature of the historical time series. For time series analysis and forecasting, this is a good result since stationary time series make model construction easier and increase prediction reliability. These results lay the groundwork for more research and modeling of the historical time series data.

Seasonal Disintegration: We conducted seasonal decomposition on Series M510 using the STL (Seasonal-Trend decomposition using Loess) technique. In order to provide a more nuanced view of the shipment patterns in the food (narrow-neck) glass container business, this phase attempted to break down the time series into trend and seasonal components.



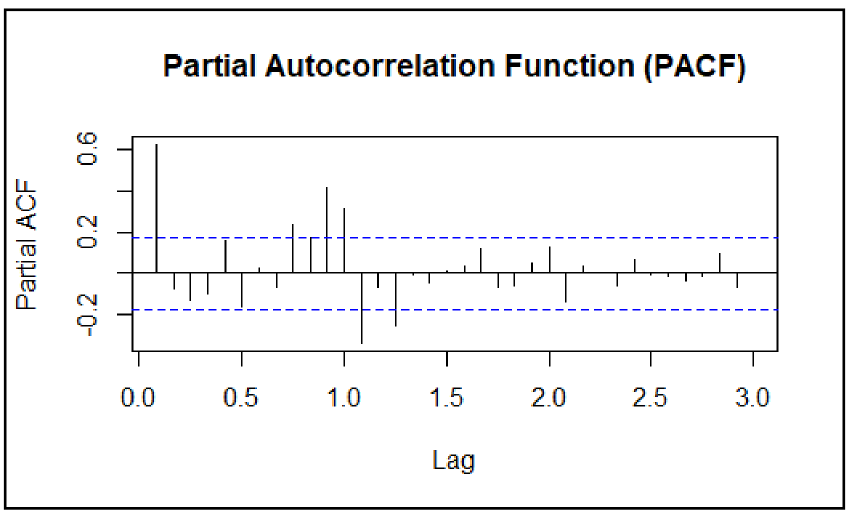
Analysis of Autocorrelation and Partial Autocorrelation: To determine Series M510's autocorrelation structure, ACF and PACF graphs were created. These charts, which highlight industry-specific temporal relationships, were extremely helpful in choosing the model.

Autocorrelation Function(ACF): Understanding the relationship between a time series and its lag values is possible with the use of the Autocorrelation Function (ACF) graphic. The correlation between the series at time t and the series at a certain lag is shown by each bar on the ACF plot. The ACF plot aids in revealing the temporal dependencies and linkages within the shipment data in the context of Series M510.



Extending beyond the confidence intervals, either positive or negative, indicates strong autocorrelation at certain delays. The ACF plot suggests a significant positive autocorrelation at lag 1.0, 2.0 and 3.0 while other lags show varying degrees of autocorrelation.

Partial Autocorrelation Function(PACF): When accounting for the intermediate lags, the PACF plot shows the relationship between a time series and its lag values.



.

Without accounting for the impact of the delays in between, each bar on the PACF plot shows the correlation between the series at time t and the series at a certain lag. The PACF plot indicates that there may be a considerable partial autocorrelation with positive and negative correlations at lags initially. After that the partial autocorrelation starts to decline and doesn't become statistically significant. Comprehending these patterns is essential for choosing a suitable time series model and deciphering the temporal correlations present in the data. To improve knowledge of the underlying time series structure, more analysis and model selection may need taking into account variables other than the information criterion.

Shapiro-Wilk Normality Test: The historical time series data were subjected to the Shapiro-Wilk normality test. The related p-value is 0.8115, and the test statistic (W) is computed to be 0.99327. This test determines if the distribution of the data is normal. The higher p-value of 0.8115 in this instance indicates that there is insufficient data to rule out the null hypothesis of normality. The historical time series data so seems to be about normally distributed based on this test.

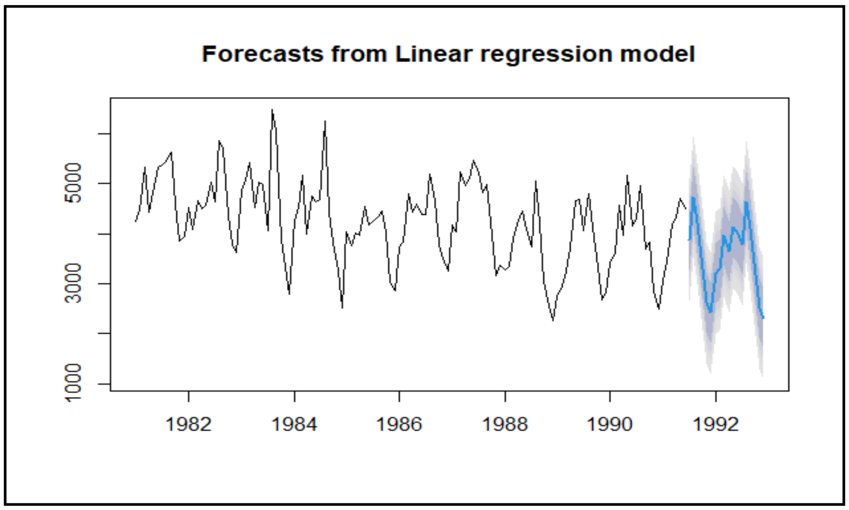
Regression Model:

From 1981 to 1992, monthly shipping data of narrow-neck food glass containers were analyzed using the time series linear regression model. With the addition of a trend component and seasonal dummy variables for every month, this all-inclusive model provided a thorough picture of the shipment patterns. A significant coefficient of -8.841, indicating a downward trend in shipments over time, is one of the noteworthy coefficients. Positive coefficients for seasonal variables (seasons 2 through 12) also show higher exports in particular months. The model is significant because it can forecast future shipments, detect broad trends, and record seasonality patterns. These abilities provide crucial information for strategic planning in the food glass container sector.

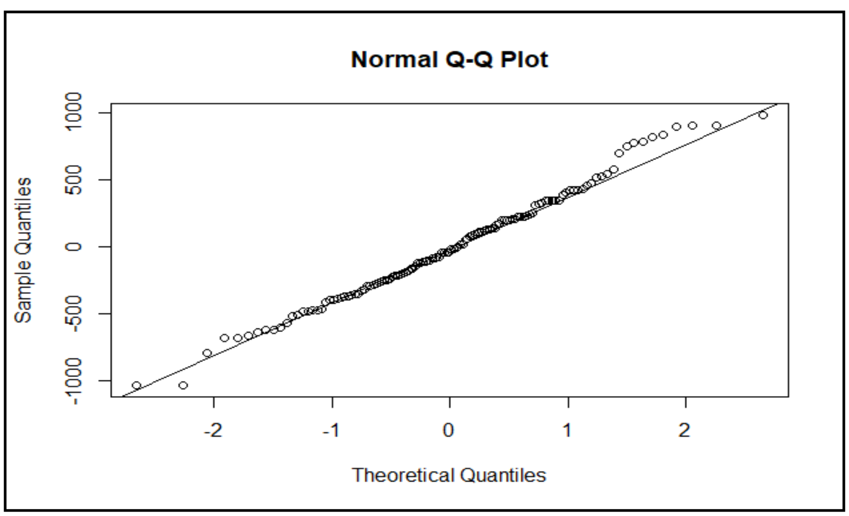
Despite its advantages, the model has several drawbacks that should be taken into account. A spread in residuals is shown by residual analysis, suggesting that the model is not explaining some of the variability. Important coefficients that provide information about the importance of each variable are the intercept (4389.663), trend (-8.841), and seasonal dummy variables. These coefficients also come with standard errors and p-values attached.

The predicted values for the time series data are shown in the forecast values plot, extending into the future. The point prediction is shown by the blue line, which gives an estimate of the expected values at each projected time point. Prediction intervals are indicated by the shaded areas surrounding the line, which provide a range of likely values. The fact that these gaps widened near the margins indicates that the prediction is more questionable. Having a visual depiction of the model's predictions and accompanying uncertainty makes it easier to plan and make decisions based on the predicted values.

	Point Forecast	Lo 80	Hi 80	Lo 90	Hi 90	Lo 95	Hi 95	Lo 99	Hi 99
Jul 1991	3884.695	3284.435	4484.956	3112.442	4656.948	2962.164	4807.227	2664.681	5104.709
Aug 1991	4740.495	4140.235	5340.756	3968.242	5512.748	3817.964	5663.027	3520.481	5960.509
Sep 1991	4213.295	3613.035	4813.556	3441.042	4985.548	3290.764	5135.827	2993.281	5433.309
Oct 1991	3302.295	2702.035	3902.556	2530.042	4074.548	2379.764	4224.827	2082.281	4522.309
Nov 1991	2620.295	2020.035	3220.556	1848.042	3392.548	1697.764	3542.827	1400.281	3840.309
Dec 1991	2412.295	1812.035	3012.556	1640.042	3184.548	1489.764	3334.827	1192.281	3632.309
Jan 1992	3213.813	2614.649	3812.976	2442.971	3984.655	2292.968	4134.658	1996.029	4431.597
Feb 1992	3288.904	2689.740	3888.067	2518.062	4059.745	2368.059	4209.749	2071.119	4506.688
Mar 1992	3958.540	3359.377	4557.704	3187.699	4729.382	3037.695	4879.386	2740.756	5176.325
Apr 1992	3644.722	3045.559	4243.886	2873.881	4415.564	2723.877	4565.567	2426.938	4862.506
May 1992	4128.358	3529.195	4727.522	3357.517	4899.200	3207.513	5049.204	2910.574	5346.143
Jun 1992	4039.631	3440.468	4638.795	3268.790	4810.473	3118.786	4960.476	2821.847	5257.416
Jul 1992	3778.603	3175.580	4381.626	3002.797	4554.410	2851.827	4705.380	2552.975	5004.232
Aug 1992	4634.403	4031.380	5237.426	3858.597	5410.210	3707.627	5561.180	3408.775	5860.032
Sep 1992	4107.203	3504.180	4710.226	3331.397	4883.010	3180.427	5033.980	2881.575	5332.832
Oct 1992	3196.203	2593.180	3799.226	2420.397	3972.010	2269.427	4122.980	1970.575	4421.832
Nov 1992	2514.203	1911.180	3117.226	1738.397	3290.010	1587.427	3440.980	1288.575	3739.832
Dec 1992	2306.203	1703.180	2909.226	1530.397	3082.010	1379.427	3232.980	1080.575	3531.832



Residual Analysis: A thorough residual analysis was carried out in order to ensure the linear regression model was reliable. The Q-Q plot demonstrates that the residuals are normally distributed.



The model's assumptions were evaluated using the following diagnostic tests, which also helped to pinpoint possible areas for development.

- **Studentized Breusch-Pagan Test:** With 12 degrees of freedom, the calculated test statistic (BP) is 10.865, resulting in a p-value of 0.5405. The non-significant p-value indicates that there is insufficient data to rule out the homoscedasticity null hypothesis. Stated differently, the model's stability is reinforced by the residuals' consistent fluctuation across a range of independent variable values.

- **Durbin-Watson Test:** The p-value that corresponds to the DW statistics of 1.186 is extremely low (5.519e-06). Given the modest p-value and the existence of autocorrelation in the residuals, it is suggested that the null hypothesis—that there is no autocorrelation—be rejected. This result is critical for upholding the independent residuals assumption, and more research or model modifications could be required.

- **Shapiro-Wilk Normality Test:** With a p-value of 0.4874, the calculated W statistic is 0.98989. The lack of substantial evidence to reject the null hypothesis of normalcy is indicated by the non-significant p-value. This implies that the residuals are roughly normally distributed, hence corroborating the validity of statistical conclusions derived from the model.

ETS Model:

Several Exponential Smoothing State Space (ETS) models are investigated in order to determine which forecasting model is optimal. Here's a condensed explanation:

Model 1: Exponential Smoothing

Detects a general trend without taking seasonal fluctuations into account.

Model 2: Holt's exponential smoothing model

Increases accuracy by adding error corrections to the trend.

Model 3: Damped Exponential Smoothing

Incorporates a trend that has steadied, offering estimates that are more accurate.

Model 4: Additive seasonality using Holt-Winter's exponential smoothing

Provides a full perspective by recognizing seasonal swings in addition to trends.

Model 5: The exponential smoothing of Holt-Winter (multiplicative seasonality

Combines seasonality and multiplicative trends, closely matching observed patterns.

Based on the analysis, Model 5 (ETS(M,Ad,M)) is the best choice since it strikes a good balance between model complexity and fit quality, based on evaluation metrics such as AIC, AICc, and BIC. Furthermore, this model has excellent performance on several training set error measures, including RMSE, MAE, MPE, MAPE, MASE, and ACF1. Particularly in view of the temporal variations and potential seasonality in the shipping data, Model 5's multiplicative seasonality and trend components demonstrate its applicability and superior fit statistics by fitting the dataset's natural patterns effectively.

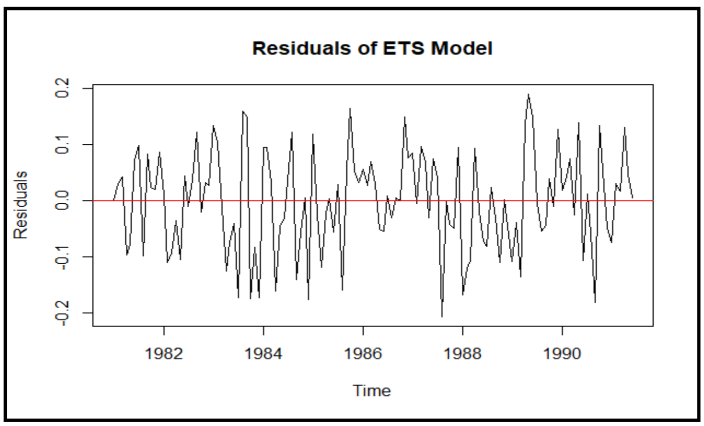
Column1	Column2	Column3
ETS Models	**MASE Values**	**MASE Values**
Set Type	**Training Set**	**Test Set**
Model 1	**1.1702**	**1.6444**
Model 2	**1.1834**	**1.7326**
Model 3	**1.1739**	**1.6442**
Model 4	**0.6479**	**0.9323**
Model 5	**0.6685**	**0.8698**

For forecasting to deliver precise forecasts and well-informed decisions in choosing the best model is crucial. A thorough examination demonstrates an intriguing difference between two popular models, Models 4 and 5, each of which offers unique benefits.

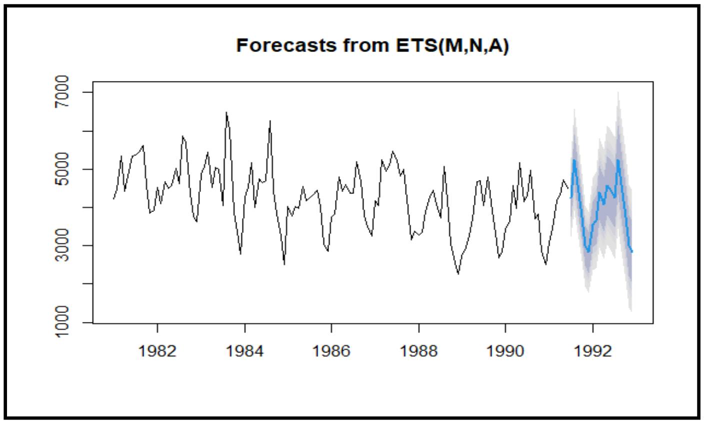
Model 5 was first in line with higher AIC and BIC values, demonstrating a preference for models that generate trustworthy results without irrelevant complexity. On the other hand, Model 5 demonstrated substantial MASE values for both the training and test datasets, but Model 4 was the clear winner when it came to predicting precision, as demonstrated by its higher Mean Absolute Scaled Error (MASE) scores for the training dataset.

In conclusion, Model 5 excels in simplicity and efficiency as demonstrated by its exceptional MASE values and lowered AIC and BIC values, while Model 4 excels in accuracy as demonstrated by its reasonable MASE scores. As a result, choosing between the two depends on specific needs: Those that place the most importance on straightforward yet precise forecasting are encouraged to choose Model 5.

A comprehensive residual analysis was carried out to evaluate the model's prediction accuracy. The residual analysis yielded insightful measures, including an average deviation that indicated a little negative bias in the model's predictions. The residuals obtained from the ETS (Error-Trend-Seasonal) time series model were analyzed and visualized using a residual plot.



Forecasting was conducted for the next 18 periods, providing point forecasts along with lower and upper bounds at confidence levels of 80%, 90%, 95%, and 99%.



Overall, the ETS model provides a strong foundation for projecting future shipments in the narrow-neck food glass container market by utilizing exponential smoothing approaches to capture both level and seasonality components in the time series. This graphical depiction provides the range of potential outcomes easier to see and helps to clarify the uncertainty around the projections.

Arima Model:

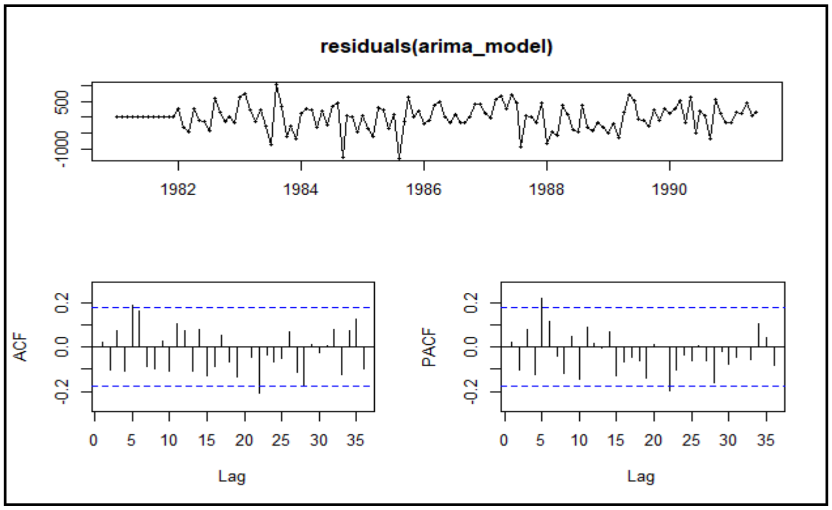
An effective time series forecasting approach for predicting future values based on previous data is the ARIMA (Autoregressive Integrated Moving Average) method. The following is a succinct summary:

Model Combinations:
- Combination 1: ARIMA(0, 1, 0) with Seasonal ARIMA(3, 1, 2) of period 12.
- Combination 2: ARIMA(0, 1, 0) with Seasonal ARIMA(2, 1, 0) of period 12.
- Combination 3: ARIMA(0, 1, 0) with Seasonal ARIMA(3, 1, 0) of period 12.

Model Analysis and Selection: The estimated coefficients, standard errors, residual variance ($sigma^2$), log likelihood, AIC, and several training set error measures (ME, RMSE, MAE, MPE, MAPE, MASE, ACF1) are the characteristics of each model. **After a thorough analysis, it can be concluded that Combination 1 is the best model because of its lowest AIC of 1735.991. A lower AIC denotes a better-fitting model.** The AIC acts as an exchange between model fit and complexity.

Residual Analysis: The residuals obtained from the Arima model were analyzed and visualized using a residual plot. The residuals appear to be white noise.

The ACF plot suggests that there is only 1 positive autocorrelation at lag 0.4 and rest all correlations within the threshold limits. Even if a single spike hitting the threshold is a sign, more research and diagnostics are necessary to confirm the importance of the pattern that was noticed.

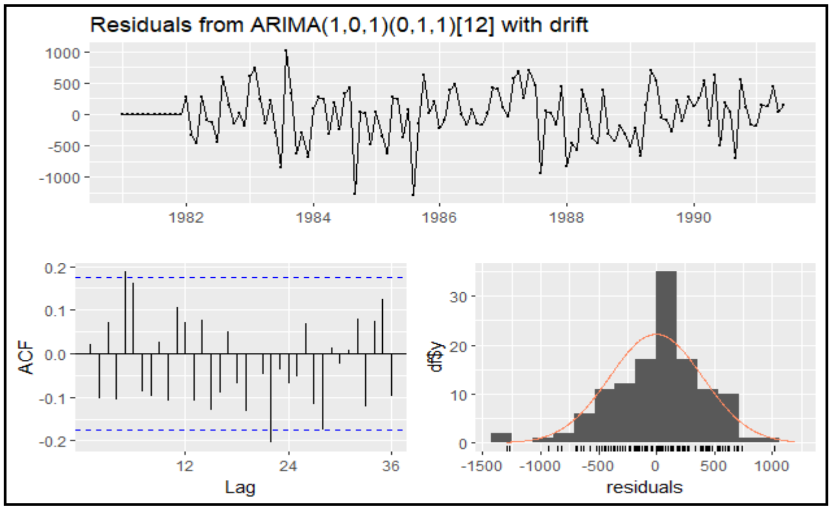


The PACF plot demonstrates that there is only spike at 0.4, indicating potential statistical importance. Additional analysis, including determining the significance of outliers or verifying normality, may shed more light on the model's operation.

Forecasted Values: In order to provide a comprehensive view of uncertainty, prediction intervals are included at different confidence levels (80%, 90%, 95%, 99%). The Point Forecast for July 1991, according to the ARIMA model projections, is 4366.362, with an 80% confidence range spanning from 3706.5454 to 5026.179. The 95% confidence interval covers the range 3357.25952 to 5375.465, while the 90% confidence interval covers the range 3519.4965 to 5213.228. The range for a greater degree of confidence, 99%, is 3040.17646 to 5692.548. These intervals give an idea of the forecast's level of uncertainty. Plotting the predicted values over time is done visually in the accompanying plot. Furthermore, for additional evaluation, the residuals—which show the prediction errors of the model—have been retrieved.

	Point Forecast	Lo 80	Hi 80	Lo 90	Hi 90	Lo 95	Hi 95	Lo 99	Hi 99
Jul 1991	4366.362	3706.5454	5026.179	3519.4965	5213.228	3357.25952	5375.465	3040.17646	5692.548
Aug 1991	4615.265	3682.1428	5548.387	3417.6157	5812.914	3188.17794	6042.351	2739.75478	6490.775
Sep 1991	4009.193	2866.3565	5152.029	2542.3783	5476.007	2261.37552	5757.010	1712.17155	6306.214
Oct 1991	3576.585	2256.9517	4896.219	1882.8539	5270.317	1558.37989	5594.791	924.21378	6228.957
Nov 1991	2686.870	1211.4747	4162.265	793.2207	4580.519	430.44764	4943.292	-278.57163	5652.312
Dec 1991	2636.155	1019.9405	4252.369	561.7661	4710.544	164.36817	5107.942	-612.32353	5884.633
Jan 1992	2942.547	1196.8354	4688.258	701.9506	5183.143	272.71174	5612.381	-566.21118	6451.304
Feb 1992	3172.867	1306.6229	5039.110	777.5687	5568.165	318.69313	6027.040	-578.15320	6923.887
Mar 1992	3870.130	1890.6795	5849.580	1329.5328	6410.727	842.82169	6897.438	-108.42749	7848.687
Apr 1992	3989.259	1902.7353	6075.783	1311.2348	6667.284	798.19627	7180.322	-204.50841	8183.027
May 1992	4410.134	2221.7691	6598.499	1601.3981	7218.870	1063.31880	7756.949	11.67326	8808.595
Jun 1992	4006.860	1721.1872	6292.532	1073.2308	6940.489	511.22526	7502.494	-587.18268	8600.902
Jul 1992	3844.189	1364.9228	6323.455	662.0853	7026.292	52.47862	7635.899	-1138.96310	8827.341
Aug 1992	4579.033	1920.2317	7237.833	1166.4986	7991.567	512.74758	8645.318	-764.97177	9923.037
Sep 1992	3773.353	946.3968	6600.310	144.9939	7401.713	-550.10360	8096.810	-1908.63226	9455.339
Oct 1992	3145.189	159.5329	6130.846	-686.8593	6977.238	-1420.97826	7711.357	-2855.77214	9146.151
Nov 1992	2397.556	-738.7804	5533.893	-1627.8882	6423.001	-2399.05664	7194.169	-3906.26162	8701.374
Dec 1992	2202.453	-1077.6484	5482.555	-2007.5117	6412.419	-2814.02949	7218.936	-4390.32268	8795.230

The ARIMA(1,0,1)(0,1,1)[12] model with drift has a Q* statistic of 36.045, 21 degrees of freedom, and a matching p-value of 0.02161, according to the **Ljung-Box test**. The purpose of this test is to determine if the residuals exhibit considerable autocorrelation and if there is white noise. The p-value in this instance is less than 0.05, indicating the presence of residual autocorrelation. It could be necessary to do further research or adjust the model in order to enhance its performance.



Shapiro-Wilk Test: The Shapiro-Wilk test on ARIMA residuals yielded a W statistic of 0.98123 and a p-value of 0.07766. The higher p-value suggests that the residuals approximate a normal distribution, supporting the validity of ARIMA model assumptions.

BATCH FORECASTING

In order to provide forecasts and insights for a range of data, the field of batch forecasting examines many time series simultaneously. This approach is especially useful when working with diverse datasets, such as the M-Competition dataset, which is made up of many time series reflecting the INDUSTRY(33), MACRO(30), and MICRO(37) categories. Having a complete understanding of the trends and patterns within each category through batch forecasting may help the shipping industry make better decisions.

The MAPE, MAE, and RMSE metrics are applied to each time series across a forecasting horizon of eighteen time periods in order to evaluate prediction accuracy. For comparison analysis, forecast values are included in the forecast_values array. Our goal is to divide the in-sample data into "training data" and "validation data" so that we can use validation to determine which of the models in our pool is the most suitable, the matrices are used to hold error data; and documentation is provided on frequency optimization.

Automatic Exponential Smoothing Model Selection: The ets() function in R's forecast package automatically selects the best Exponential Smoothing model based on in-sample data. This method considers various combinations of error, trend, and seasonality components to determine the most suitable model for each time series. Lower MAPE value explains better accuracy.

- **Average MAPE (Mean Absolute Percentage Error): 17.17%**
- **Average MAE (Mean Absolute Error): 605.02**

- **Average RMSE (Root Mean Squared Error): 733.55**

Automatic ARIMA Model Selection: The auto.arima() method is used to choose ARIMA models automatically. It methodically examines several ARIMA configurations to determine the best model for a given time series, much like ETS. The effectiveness of this method in managing various patterns in time series data is well known. The findings are:

- **Average MAPE (Mean Absolute Percentage Error): 19.97%**
- **Average MAE (Mean Absolute Error): 641.19**
- **Average RMSE (Root Mean Squared Error): 781.88**

Model Selection Strategy: To evaluate the performance of the ETS and (ARIMA) models on various time series data, the validation strategy is carefully used in our forecasting research. The Mean Absolute Percentage Error (MAPE) metrics were utilized to conduct a thorough model comparison for every series. The ETS and ARIMA models generated distinct forecasts for each dataset. When comparing the ETS and ARIMA models, MAPE is preferred above MAE (Mean Absolute Error) and RMSE in the model selection process. By providing a percentage representation of predicting accuracy, MAPE makes it easier to compare different time series with different sizes. Minimizing the MAPE led to the selection of the best accurate model for each series. By ensuring that our forecasting technique is precisely tailored to the distinct qualities of each time series, this validation procedure helps to improve the whole dataset's prediction accuracy.

- **Average MAPE (Mean Absolute Percentage Error): 14.50%**
- **Average MAE (Mean Absolute Error): 540.11**
- **Average RMSE (Root Mean Squared Error): 684.46**

The above findings show that, with a reduced Mean Absolute Percentage Error (MAPE) of 14.50%, the third model forecasts more accurately than the separate models. A more accurate forecast of future values is indicated by a lower MAPE number. Comparatively, the ETS and ARIMA models independently had higher MAPE values (17.17% and 19.97%, respectively). This demonstrates that a more reliable and accurate forecasting strategy for the provided time series data is produced by the hybrid technique, which combines the best features of the ARIMA and ETS models. The practical implication is that integrating several forecasting techniques can improve prediction performance overall and support more informed corporate decision-making.

Naive Benchmark: For the out-of-sample data, the Naive forecasting approach produced fewer accurate forecasts because it makes the assumption that future observations would be equal to the most recent previous observations. The Naive method's Average MAPE was 22.72%, suggesting a significant percentage discrepancy between the observed and predicted values. Additionally high at 872.13 and 1071.97, respectively, were the average MAE and RMSE. These findings show that more complex models are required and emphasize the drawbacks of predicting using only the most recent data.

Mean Forecasting Benchmark: The benchmark Mean Forecasting approach showed low forecasting performance as it predicted future values by averaging previous data. The approach produced an Average MAPE of 35.62%, which suggests a significant relative inaccuracy in the forecasts. The RMSE and Average MAE values were 1320.16 and 1183.43, respectively, which highlights the limits of the approach in identifying the underlying patterns in the time series data.

Holt-Winters Benchmark: Among the benchmarks, the Holt-Winters benchmark approach demonstrated comparatively higher predicting performance. It is an extension of exponential smoothing that accounts for trends and seasonality. Compared to the Naive and Mean Forecasting standards, Holt-Winters' Average MAPE was 18.92%, which indicates a smaller percentage discrepancy between projected and observed values. Improved accuracy was shown by the average MAE and RMSE of 698.80 and 835.71, respectively.

Analysis and Comparative Evaluation: Important insights into the relative performances of the forecasting models—ETS, ARIMA, the Model Selection Strategy and the three benchmark techniques (Naive, Mean Forecasting, and Holt-Winters) are obtained through their examination across a number of variables.

Performance Across All Series: When the models' accuracy is examined over all series, the Model selection strategy consistently performs better than any of the individual models or the benchmarks. By combining the best features of ARIMA and ETS, the hybrid approach forecasts more accurately.

Comparing Individual Models: When comparing ETS and ARIMA separately, it can be demonstrated that ETS is better at capturing time series patterns, whereas ARIMA shows somewhat greater errors. On the other hand, the selection technique reduces individual flaws and achieves higher accuracy by strategically utilizing both models.

Benchmark Methods: As crucial points of reference, benchmarking techniques like Naive, Mean Forecasting, and Holt-Winters are employed. Because they are so simple, naive and mean forecasting are not very good at managing intricate time series patterns. The more complex Holt-Winters approach performs worse than the integrated technique, yet it still gives an edge among benchmarks.

The trade-offs between forecasting model complexity and simplicity are emphasized in a critical discussion. Although benchmarks offer simple methods, it is apparent that they are not flexible enough to accommodate a wide range of patterns. Although they have advantages, ARIMA and ETS have drawbacks. The Model Selection Strategy combines the best attributes through a rigorous validation procedure, providing increased accuracy and flexibility.
The findings suggest that the Model Selection strategy is a powerful tool for decision-making. Its accuracy in a variety of circumstances and flexibility with regard to various data types establish it as a trustworthy forecasting method. This has significant implications for making well-informed decisions in shipping industries.

Conclusion:

The objective of this extensive study was to enhance time series forecasting in the glass container sector, specifically in the dynamic food glass container market. With a focus on Series M510 in particular, the study used hybrid model selection, ETS, ARIMA, and manual modeling. The conclusions included insightful information on the advantages and disadvantages of several forecasting techniques, along with benchmark assessments (Naive, Mean Forecasting, Holt-Winters).

Limitations:

- The primary focus of the study was Series M510; hence, care should be taken when extrapolating the results to other series.
- The models' effectiveness is predicated on an assumption that past trends will endure, which may not be true in markets that are undergoing rapid change.
- The models did not specifically account for external factors like changes in the economy, geopolitical events, or technology improvements, which might have introduced sources of mistake in predictions

Managerial Implications:

- Decision-makers in the glass container business are advised to use the Model Selection Strategy, which provides accurate and flexible predictions for strategic decision-making.
- Since the sector is dynamic and underlying trends are likely to change, it is essential to continuously analyze and validate forecasting models.
- It is important for decision-makers to understand the trade-offs between model simplicity and complexity. While more complicated models may improve accuracy, they may also make interpretation more difficult.

Appendix:

1. Model Selection Strategy Evaluation Metrics:

Accuracy Metrics:

- Mean Absolute Percentage Error (MAPE)
- Mean Absolute Error (MAE)
- Root Mean Squared Error (RMSE)

Validation Metrics:

- Akaike Information Criterion (AIC)
- Bayesian Information Criterion (BIC)
- Augmented Dickey-Fuller Test Statistic
- Phillips-Perron Unit Root/Cointegration Test Statistic

2. **Manual Modeling - Descriptive Statistics for Series M510:**
    - Minimum Production: 2266 units
    - Median Production: 4271 units
    - Maximum Production: 6494 units
    - Monthly Output Average: 4236.27 units
    - Interquartile Range: 4563 to 5190 units
3. **Time Series Plot and Stationarity Tests:**
    - Time Series Plot of Series M510
    - Augmented Dickey-Fuller Test Results:
    - Test Statistic: -5.1342
    - Lag Order: 4
    - P-Value: 0.01
    - Unit Root/Cointegration Test (Phillips-Perron):
    - Test Statistic Values: -5.9633, 11.9281, 17.8715
4. **Seasonal Disintegration and Autocorrelation Analysis:**
    - Seasonal Decomposition (STL) for Series M510
    - Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) Plots
5. **Regression Model Analysis for Series M510:**
    - Coefficients: Intercept, Trend, and Seasonal Dummy Variables (Seasons 2 through 12)
    - Residual Analysis:
        - Q-Q Plot for Normality
        - Studentized Breusch-Pagan Test
        - Durbin-Watson Test
        - Shapiro-Wilk Normality Test
7. **ETS Model Evaluation for Series M510:**
    - ETS Model Options: Exponential Smoothing, Holt's Exponential Smoothing, Damped Exponential Smoothing, Additive Seasonality, Multiplicative Seasonality
    - Performance Metrics (MASE) for Training and Test Sets
8. **ARIMA Model Analysis for Series M510:**
    - ARIMA Model Combinations: ARIMA(0, 1, 0) with Seasonal ARIMA(3, 1, 2) of period 12, ARIMA(0, .1, 0) with Seasonal ARIMA(2, 1, 0) of period 12, ARIMA(0, 1, 0) with Seasonal ARIMA(3, 1, 0) of period 12

- Model Evaluation Metrics: AIC, ME, RMSE, MAE, MPE, MAPE, MASE, ACF1
- Residual Analysis and Ljung-Box Test

9. Batch Forecasting:

- Evaluation Metrics (MAPE, MAE, RMSE) for Automatic Exponential Smoothing and Automatic ARIMA Models
- Model Selection Strategy Performance Metrics Across Multiple Time Series

10. Benchmarking Analysis:

- Performance Metrics for Naive, Mean Forecasting, and Holt-Winters Benchmarking Techniques