**Submitted To:**

Dr. Dawit Zerom

**Submitted By:**

Ruchika Narang (893521344)

Sanchit Singh (893502922)

Abhinay Sariswal (893582536)

**Forecasting Titanium beer bottle sales from march-december 2004**

Contents

[**Executive Summary** 2](#_Toc529404118)

[**Objective** 3](#_Toc529404119)

[**Examining Data Patterns** 4](#_Toc529404120)

[Determining Patterns in time series using Autocorrelation Analysis 4](#_Toc529404121)

[**Model Selection** 7](#_Toc529404122)

[**Forecast for rest of 2004 - Period (March ‘04 to December ‘04)** 9](#_Toc529404123)

[**Conclusion and Recommendations** 10](#_Toc529404124)

# **Executive Summary**

In our project, we conducted a thorough analysis for the demand of beer bottles at Titanium Brewery, located in the southern Caribbean Island of Trinidad. To suffice our forecasting values, we collected the monthly historical data from January 1999 to February 2004 (in thousands of cases). Then, we have utilized the application “Forecast Pro” to get actionable insights for any possible trend, seasonal, or cyclical patterns.

Per requirement, we withheld 9 months of data, from June 2003 – Feb 2004, and performed the Exponential Smoothing method to produce a forecast. From our analysis, we have narrowed down some possible forecasting methods, and then we ranked each method based on their fit and accuracy measures. By observing MAPE for all the methods, a measure of predicting accuracy of a forecasting method, the Custom Exponential Smoothing method with no trend and an additive seasonality has the lowest value of 8.01%. Hence, the Custom Exponential Smoothing method with no trend and an additive seasonality appeared to be the most accurate and fits the data correctly.

To help Keating in his bottle purchase decision at Titanium Brewery, we recommend distributing his sales over the period of several months with the help of marketing rather than driving sales on Carnival. This way we can minimize year-end inventory, thus utilizing the limited storage capacity of empty bottles. The following report demonstrates our analysis in a detailed manner, explaining each step of our study and how we reached our conclusion.

# **Objective**

The forecasting problem for the Titanium Brewery is that the covered storage space for empty bottles is tight and a bottle design change is expected in 2005 and 2006. So, they want to forecast the number of beer bottles to order in the year 2004 which would accommodate the year’s supply comfortably along with minimizing the year end inventories. Both under-forecasting and over-forecasting are going to be expensive situations hence, careful forecast is needed as a strong base before placing the order of beer bottles. The cost of under production and losing sales is much higher than the cost of over stock. If sales exceed, Titanium Brewery must expand their storage space. They must also consider by how much the quarterly sales exceed, for how many quarters the sales will exceed, and finally whether exceeding the sales in a year will be a short-term situation or if it will be a permanent state. They should also record if sales increase in a particular season. For instance, generally beer sales are more in summer than in winter except during carnivals. These decisions could be taken after proper forecasting of Titanium brewery business.

***Below is a graphical framework that illustrates how our decision making is related to forecasting.***

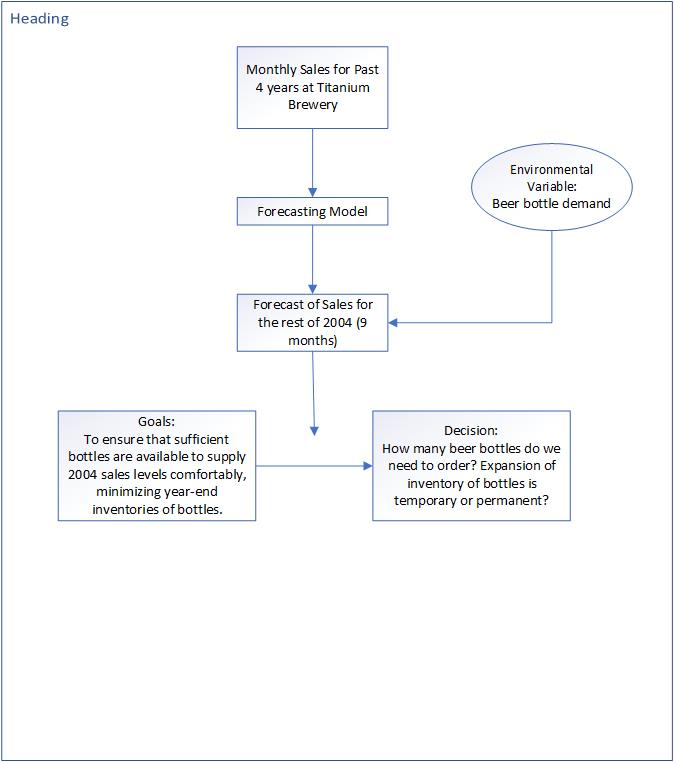


Figure 1: Relation between Forecasting and Decision

# **Examining Data Patterns**

Before identifying the number of bottles, we wanted to determine the patterns within our data such as seasonality, trend, stationarity, or irregularity. This would help us come up with better decision and recommendations.

If we visualize the below time series, we observe that our data has a trend resulting from long term increase in data. The series has a linear trend since we can draw a straight line along the increase. Now, the rises are not in fixed period. As a result, we can also infer that the series has a cyclic trend.

There is an increase in every February which is clear from consistent peaks during February repeated annually. The pattern of our time series also has seasonality, in fact, additive seasonality. We can say this from the seasonal variations which are independent of the level.

The time series lacks stationarity since mean and variance are not constant over time.

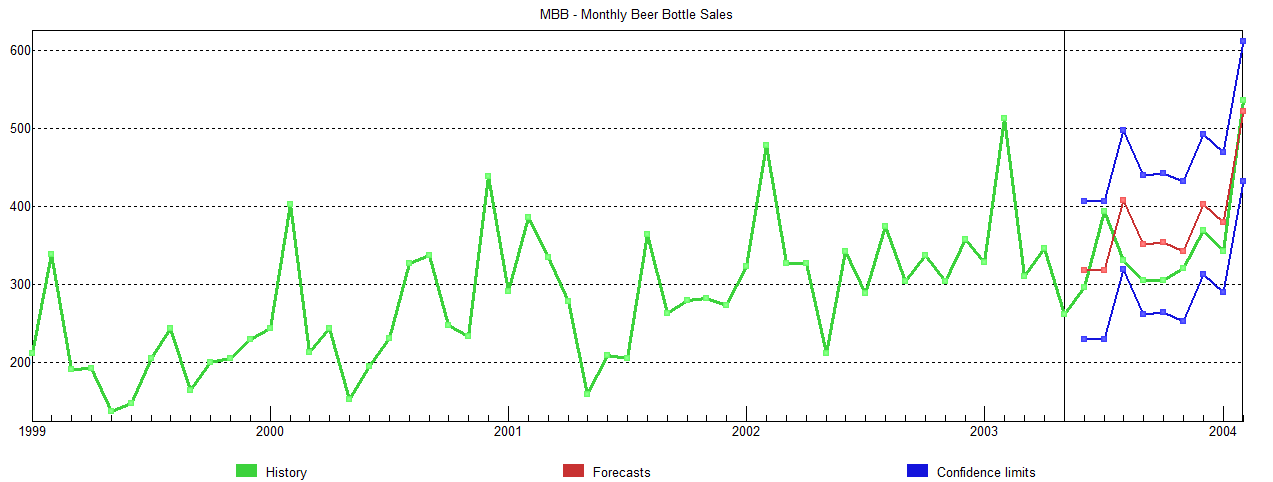


Figure 2: Time Series- Monthly Beer Bottle Sales

## Determining Patterns in time series using Autocorrelation Analysis

We performed autocorrelation to confirm on the above data patterns which we observed via visualizing the time series. We modified ‘Simple Differencing’ and ‘Seasonal differencing’ to determine the patterns in the data.

* **Patterns observed when no differencing was applied**

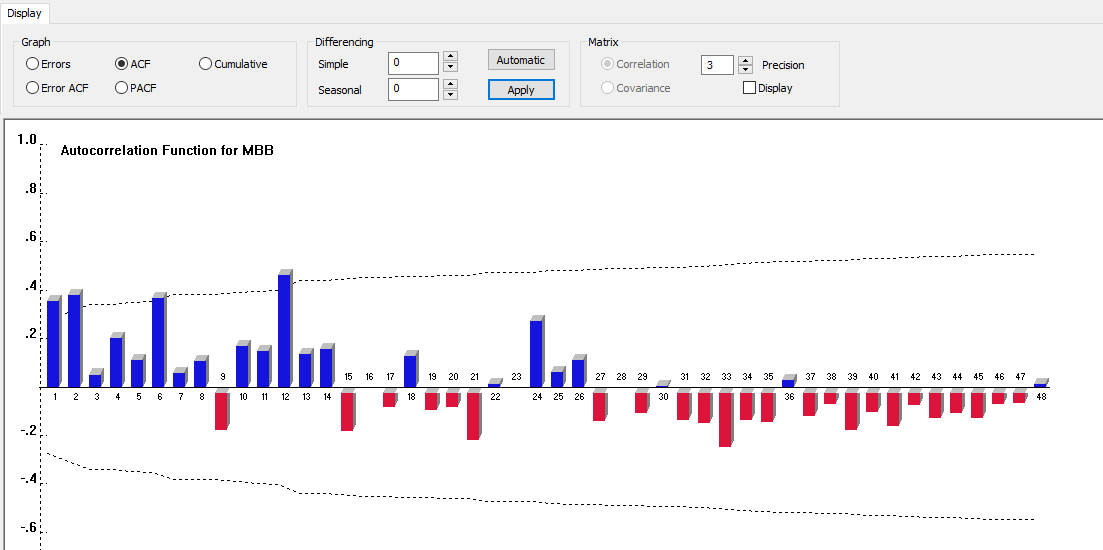


Figure 3: ACF with no differencing

From the figure above, we observe that the autocorrelation function gradually drops to 0. As a result, we can infer that the time series is non-stationary. Also, there is no trend within the series since there is no long term increase or decrease in data.

* **Patterns observed when first order simple differencing was applied**

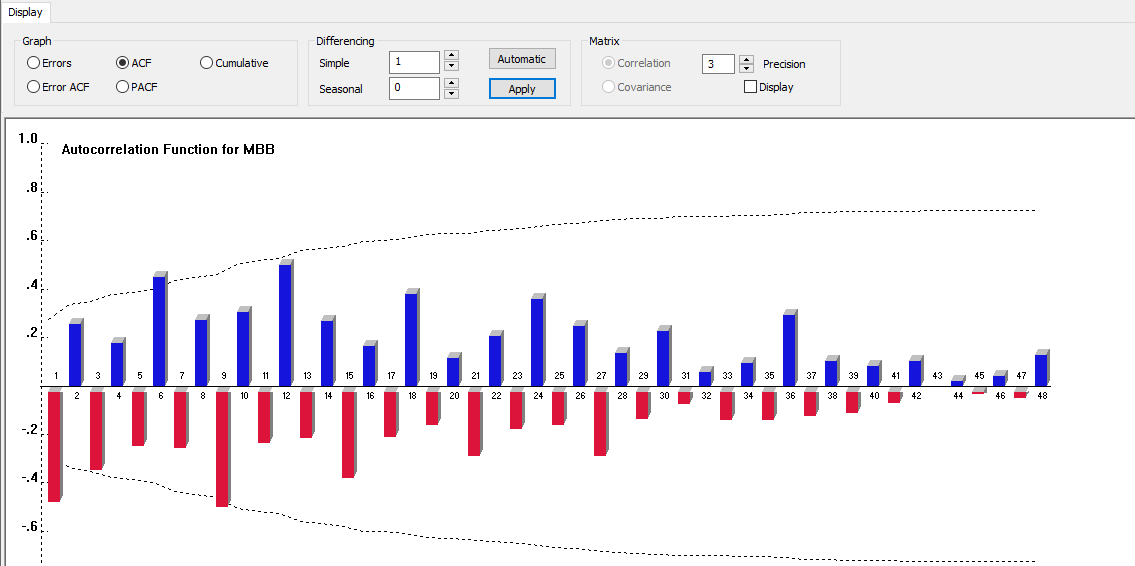


Figure 4: ACF with first order simple differencing

The figure shows that the series has seasonality which can be observed by increase of sales every February and also by a significant autocorrelation coefficient occurring in lags of 12 months.

* **Patterns observed when first order seasonal differencing was applied**

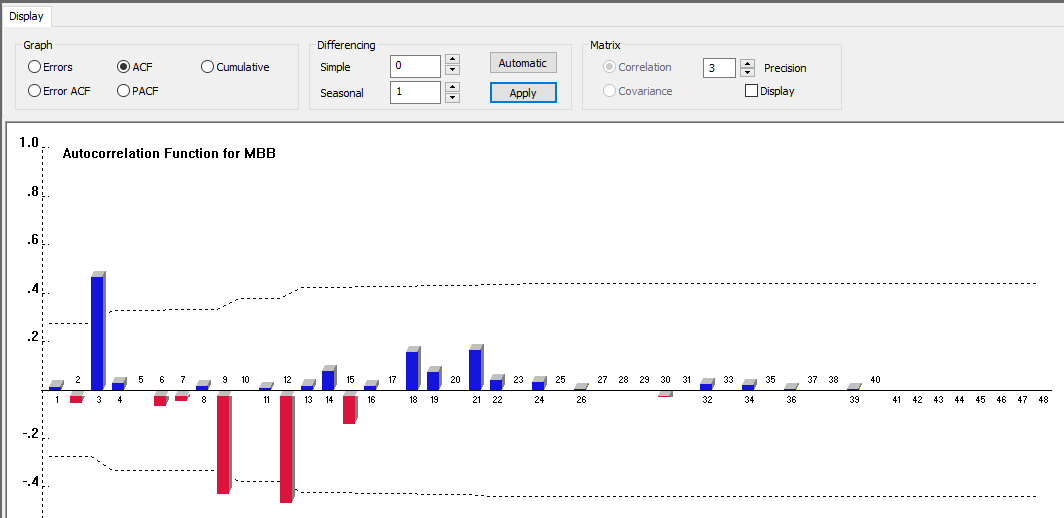


Figure 5: ACF with first order seasonal differencing

We see that the peak and depressions are not repeated after fixed periods. Therefore, the pattern within the series is cyclic.

* **Patterns observed when both first order seasonal and simple differencing were applied**

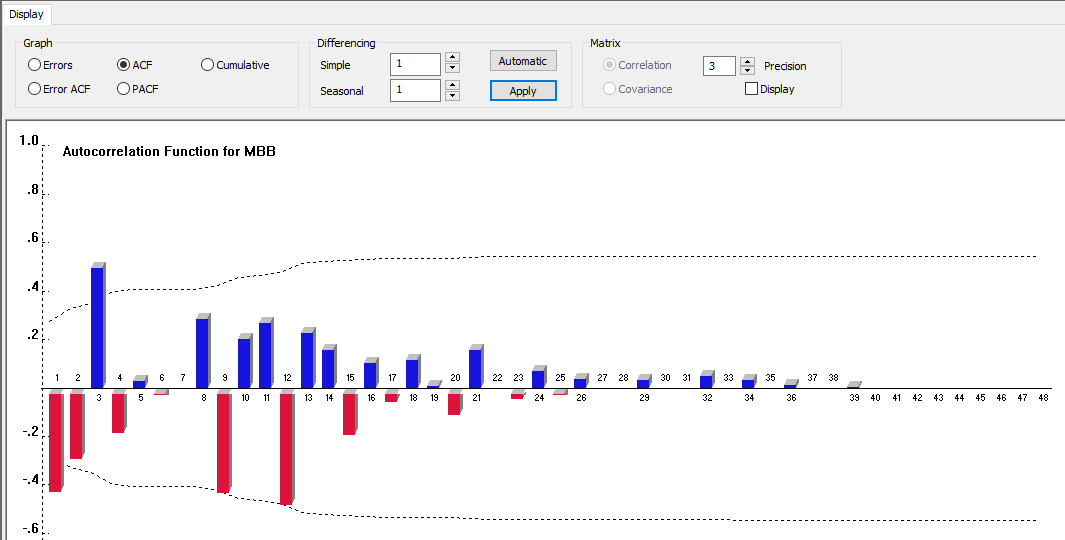


Figure 6: ACF with first order simple as well as seasonal differencing

Non-stationary data should be converted to stationary so that further statistical analysis can be done. So, we applied both first order Simple and Seasonal differencing in order to make the series stationary. We can observe that the series is moving towards stationary with constant mean and variance. Moreover, the autocorrelation immediately drops to 0 which happens when the series is stationary.

# **Model Selection**

As per the requirement, we need to use the methods from the class of exponential smoothing models to forecast the sales. The below table lists all the methods (including all custom smoothing methods). There are total 8 methods that we are considering. From the data pattern analysis, explained further in the report, we acknowledged that the data shows an additive seasonality and implies no trend. In light of this, a Custom Exponential Smoothing with a no trend and an additive seasonality should perform the best. Also, the time series does indicate a possible cyclical pattern.

The below table summarizes the calculated MAPE values comparing holdout forecasts (9 months) with actual data for the 8 methods that we are considering.

|  |  |  |
| --- | --- | --- |
| **Type of Exponential Smoothing** | **Method** | **MAPE** |
| Custom Exponential Smoothing | No Trend /Additive Seasonality | 8.01% |
| Custom Exponential Smoothing | No Trend /Multiplicative Seasonality | 8.03% |
| Custom Exponential Smoothing | Linear Trend / Additive Seasonality | 12.16% |
| Winters Exponential Smoothing | Linear Trend / Multiplicative Seasonality | 12.46% |
| Custom Exponential Smoothing | Damped Trend / Additive Seasonality | 9.76% |
| Custom Exponential Smoothing | Damped Trend / Multiplicative Seasonality | 12.35% |
| Custom Exponential Smoothing | Exponential Trend /Additive Seasonality | 12.16% |
| Custom Exponential Smoothing | Exponential Trend /Multiplicative Seasonality | 8.47% |

Table 1

The MAPE indicates the prediction accuracy of a forecasting method in statistics, for example in estimation of the best model. Thus, from the above chart we can see that custom exponential smoothing with no trend and an additive seasonality produced a considerably low MAPE value. Its MAPE is 8.01%, this means that the mean absolute error for the 9 months forecasted was 8.01% of the actual beer sales data.

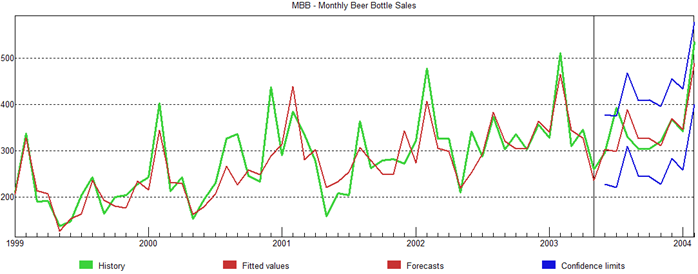


Figure 7: Graph after applying Exponential Smoothing

The above time series is for the custom exponential smoothing with no trend and an additive seasonality. We can see from the time series that the fitted value (red line) is a good fit for our history data (green line).

In the data patterns section, we applied simple differencing to the dataset in order to analyze the seasonality. We could see significant positive correlation in the ACF plot for the lags of 6, 12, 18, 24 and so on. This shows that there is a significant seasonality occurring in a period of 6 months. We also notice that the seasonality exhibited in the lag intervals of 12, 24, 36 are prominent whereas the seasonality at 6 and 12 further decreases in magnitude. We can see similar observation from the plots below. Our time series ends at February ‘04 and we can see that the most prominent periodic peaks are in February (lag=12) and there is a constant low in the months of July (lag=6).

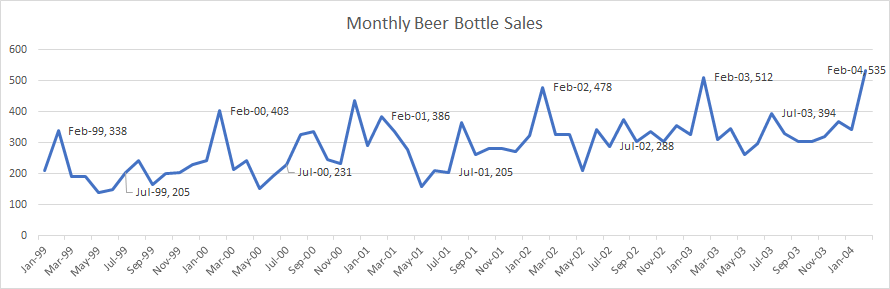
****

Figure 8: Monthly beer bottle sales

# **Forecast for rest of 2004 - Period (March ‘04 to December ‘04)**

Now, we move towards the final goal to predict the sales for the rest of the year 2004. Based on our analysis and model evaluation illustrated above, we have applied the Exponential smoothing method for No trend and an additive seasonality.

The forecasted values are given in the table below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Month** | **2.5% Lower** | **Forecast** | **Quarterly** | **Annual** | **97.5% Upper** |
| 2004-Mar | 255 | 346 | Q1, 1,223 | 4,205 | 436 |
| 2004-Apr | 258 | 350 | Q2, 916 | 442 |
| 2004-May | 163 | 256 | 350 |
| 2004-Jun | 214 | 309 | 405 |
| 2004-Jul | 236 | 333 | Q3, 1,039 | 430 |
| 2004-Aug | 279 | 378 | 477 |
| 2004-Sep | 228 | 328 | 429 |
| 2004-Oct | 227 | 329 | Q4, 1,028 | 431 |
| 2004-Nov | 219 | 322 | 426 |
| 2004-Dec | 271 | 377 | 482 |

Table 2

In our holdout analysis we got a MAPE error of 8.01% which is a crucial factor to take a call on the actual bottle levels to be maintained in the inventory. The MAPE indicates that on an average monthly forecast errors of the model are 8% of the monthly beer sales. The holdout analysis indicates that the forecast MAPE error is 8%, and the comparison of numbers from the holdout analysis in the table below shows that it would be ideal for the us to take the upper forecast limit of 97.5% into consideration. We can clearly see below that the actual values are majorly falling below the upper forecast limit of 97.5 %. This gives us a positive indication that the confidence limits can further be explored to optimize the space constraints as per the business decisions. A proposal can be made to discuss on the impact to decide the final order amount. We know that the beer sales are very low during the winter season (specially end year months - November, December). Hence, we can check the feasibility of choosing the ordering amount close to the point forecast values. One of the key business problems was to consider the fact that the we need to minimize the year-end inventories during the end of 2004. We can see from the table below that there is a very less difference between the actual and the forecasted value for 2003. We can say that the model is accurate to forecast the sales values for the months of November and December. Based on this, the management must make a decision on how much over forecasting cost can the company bear. They can calculate this cost based on the company’s stock keeping unit and the MAPE.

|  |  |  |  |
| --- | --- | --- | --- |
| **Month** | **Forecast** | **97.5% Upper** | **Actual** |
| Jun-03 | 304 | 393 | 296 |
| Jul-03 | 299 | 391 | 394 |
| Aug-03 | 390 | 484 | 331 |
| Sep-03 | 328 | 424 | 305 |
| Oct-03 | 328 | 426 | 305 |
| **Nov-03** | **312** | 413 | **321** |
| **Dec-03** | **370** | 473 | **369** |
| Jan-04 | 346 | 451 | 342 |
| Feb-04 | 490 | 596 | 535 |

Table 3

# **Conclusion and Recommendations**

Proper demand forecasting enables better planning and utilization of resources for business to be competitive. We have learnt through our analysis that to better understand which forecasting technique is best, we utilized the required error measurement such as MAPE and checked how each method perform in order to fit the prediction with the time series patterns. With our forecasting data and the confidence intervals, there is a less than 2.5% chance of the demand being under-forecasted. The upper and lower confidence limits represent a 95% confidence interval for each forecast. This means that we can say with a .95 probability that the actual sales in a month will be between their lower and upper limits. Also, it is noticeable that no monthly forecasted data exceeds the upper and the lower limits implying a credible forecast.

Our recommendation would be to create a demand system that considers all beverages simultaneously. Titanium Brewery could change their marketing strategy to influence customers who normally buy beer on Carnival to buy in different months of the year. This way we can minimize the year-end inventories of bottle by distributing the sales across the year and Keating will be able to better predict the sales. Since Titanium Brewery is driving the marketing strategies, it is easier to predict the sales based on the geographic and demographics of the target market.