Time Series Forecasting for Alcohol Revenues in the United States

By: Ethan Biddle

MA 585

2 May 2023

I. Introduction

Since its reintroduction to the USA in 1933 [1], alcohol has become a mainstay of many people's daily lives. But now with the introduction of several new brands and different forms of alcohol unlike the typical beer and wine, drinking alcohol has become less of an event, and more of a casual activity done alongside other plans. This is even more evident in my personal daily life, where all of my friends are now of proper age and can legally partake in this lifestyle. This notion is supported by the data, where we can see a 5% increase in alcohol usage over the last century, as well as a huge spike in revenue for these alcohol brands. In response, I've generated both a seasonal ARIMA and a seasonal Holt-Winters model to forecast what future alcohol revenue will look like over the next 12 months.

II. Data Set

This data set [2], was retrieved from the Federal Reserve of Economic Data and goes back as far as 1992. But for the purposes of this project, I only looked at the last five years given that the recent rise in new alcohol brands has caused a massive, sudden spike in alcohol sales that largely differed from what sales were in the 1990s. After some basic filtering and cleaning to make the data more presentable, I ran a box plot [1] to check for possible outliers, of which there were none.

III. Model Selection

I first began by simply plotting my data to see the overall trend, which is clearly positive as seen in Figure 1. However, from this figure, we can also see that the data is not stationary, as our variance is not consistent. Suspecting there to be a seasonal trend, I then produced a decomposition chart, Figure 2, supporting my preconceived claim. Given this, I proceeded to apply a difference order of 12 given the data is monthly, but to no avail, my data was still non-stationary. I then applied another difference order, which finally got my data to be stationary, as seen in Figure 3.

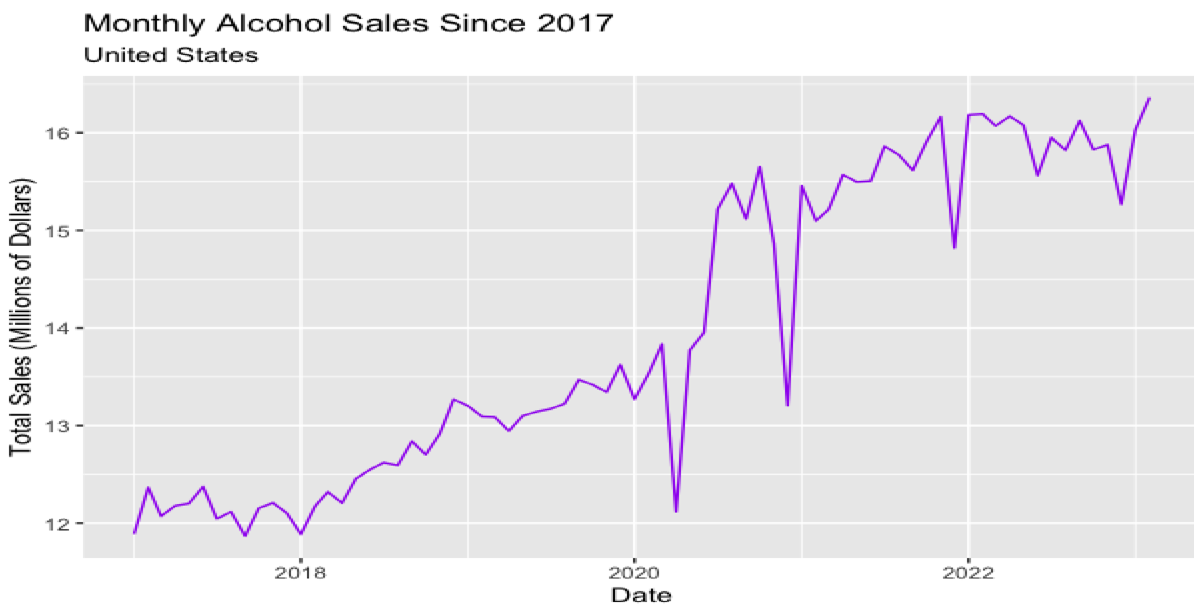

Figure 1.

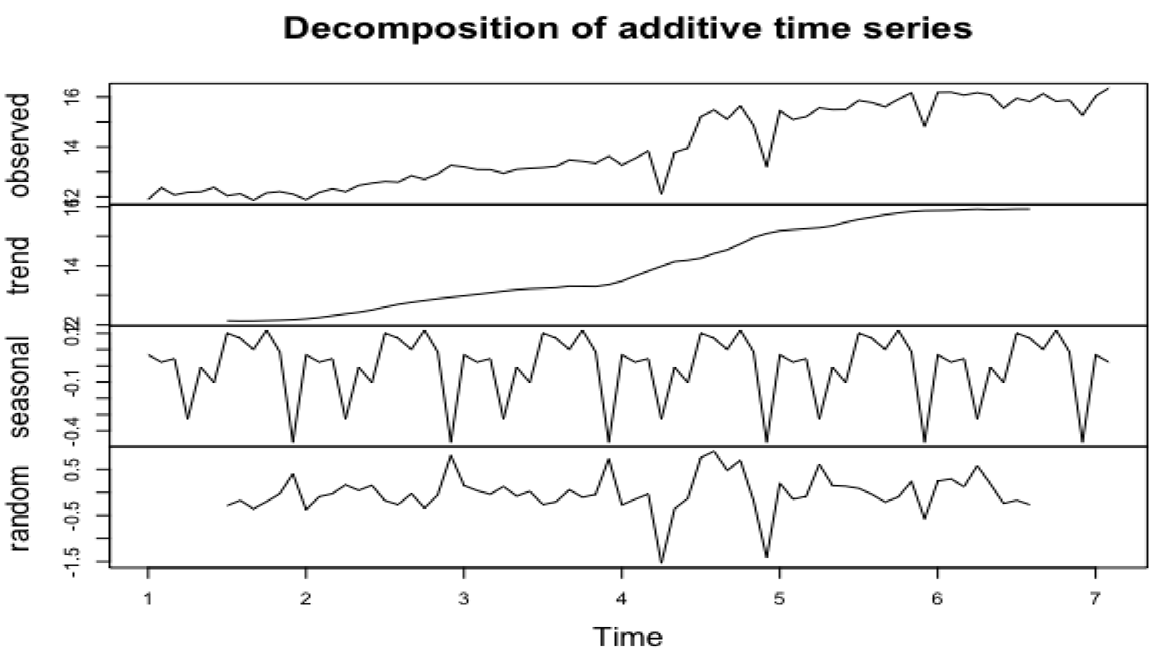

Figure 2.

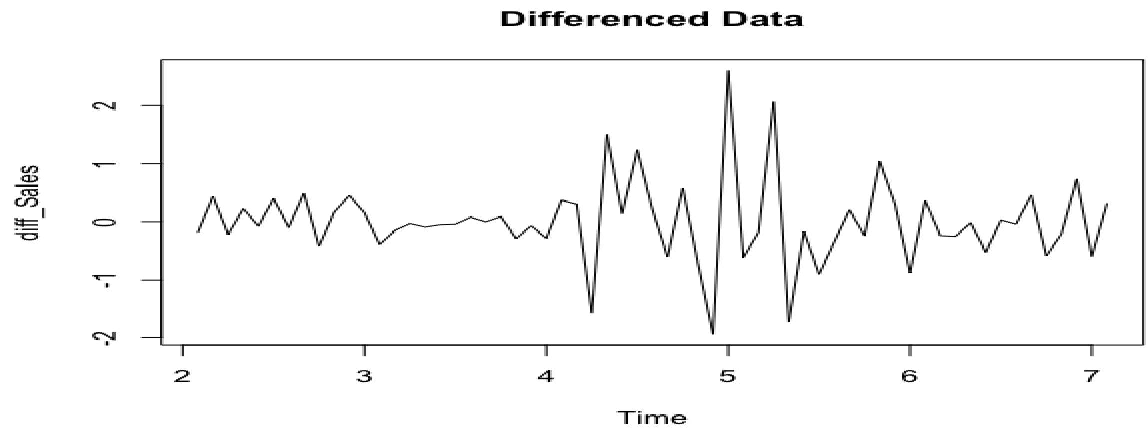

Figure 3.

I then began working on my ARIMA model, which based on the produced ACF and PACF plots seen in Figure 4, was likely to consist of an AR(1) aspect due to the alternating ACF pattern, and some sort of MA component as well. With large lags appearing further down at lag 15, I opted for an ARMA(1,15) model, which would follow with me testing different seasonal components, as seen in Figure 5.

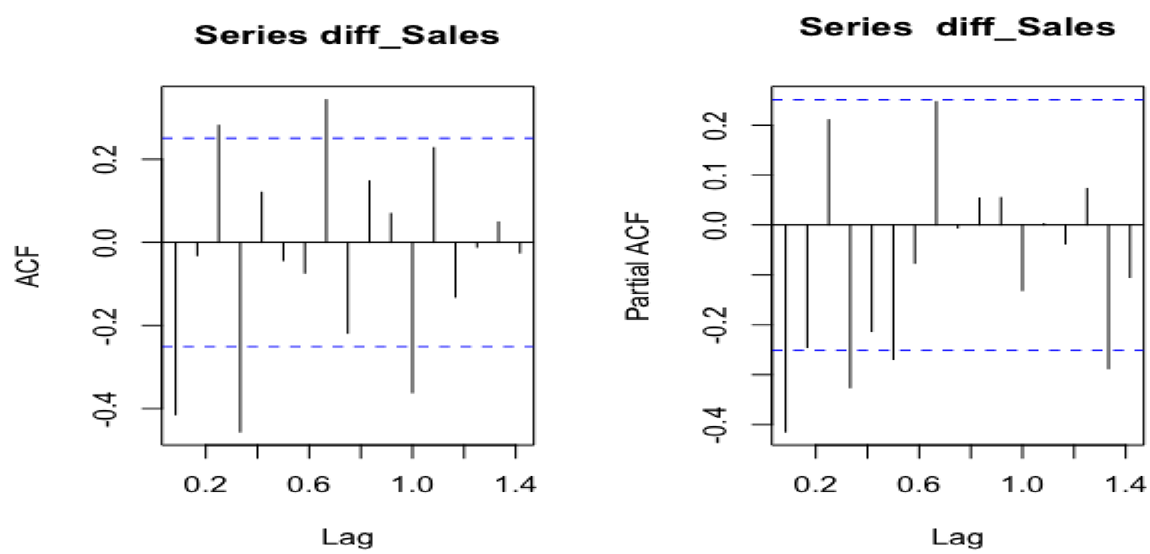

Figure 4.

AICc Values by Model

Model	AICc
(1,0,15) x (1,1,1)[12]	142.27

(1,0,15) x (1,2,1)[12]	166.43
(1,0,15) x (1,1,2)[12]	166.43
(1,0,15) x (1,0,1)[12]	141.65
(1,0,15) x (1,0,0)[12]	134.12

Figure 5.

It appeared that based on our AICc values, our best model is (1,0,15) x (1,0,0)[12], however after using the "auto.arima" function in R, I was also able to find an even lower AICc value of 108.84 with the model (0,1,1)x(1,0,0)[12].

IV. Forecasts

Beginning with my first ARIMA model, we can see in Figure 6 that the forecast sees a downward trend in alcohol sales over the next year, perhaps influenced by the downward trends seen earlier in the graph and that the current peak is just that, a peak.

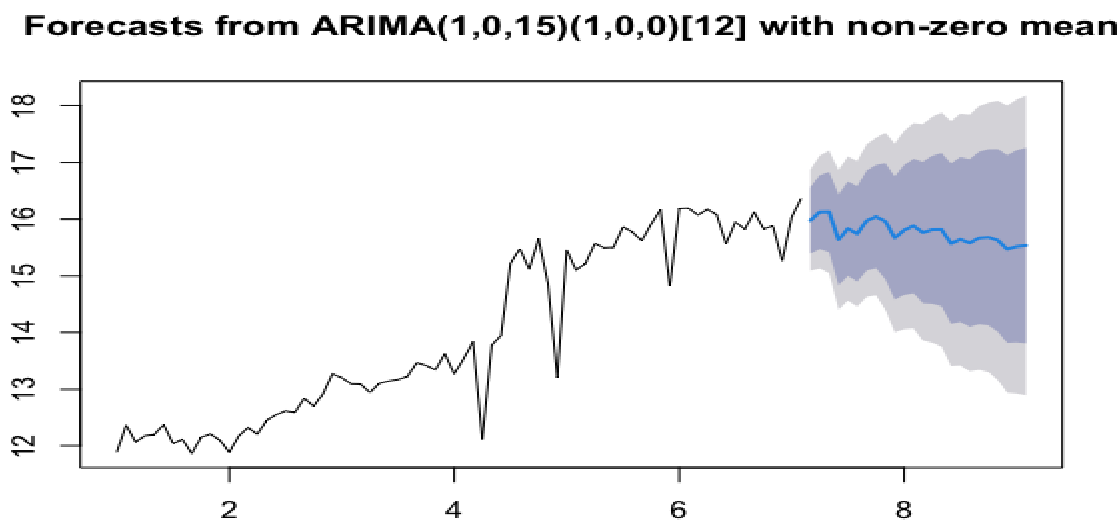

Figure 6.

Now for the ARIMA model generated by R, we get a much more optimistic forecast, where Figure 7 shows a clear upwards trend for alcohol revenue, proceeding with the idea that the current state of alcohol sales is only just the beginning of what will continue to be a fruitful business.

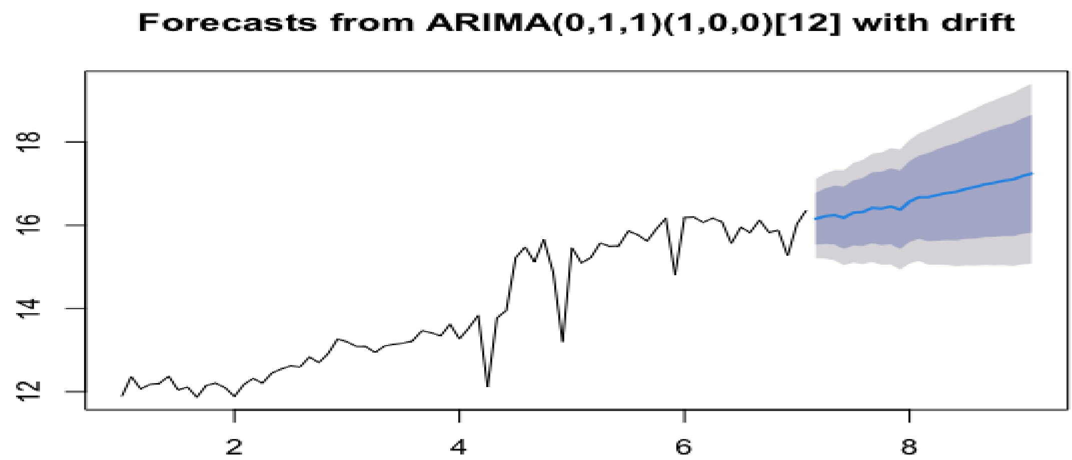

Figure 7.

Lastly, I conducted a Holt-Winters forecast to see an alternate perspective on what the forecast might look like. This forecast was a bit in the middle of the first two, as while it was optimistic about alcohol's future revenue, they also recognize a trend of sudden dips in revenue, which has been accounted for in the forecast, as seen in Figure 8.

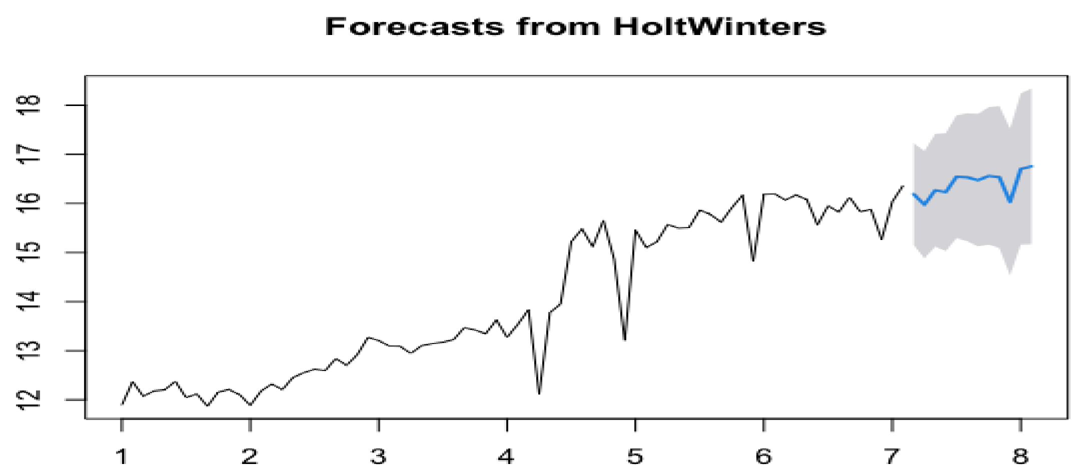

Figure 8.

V. Conclusion

In an effort to see how accurate my forecasts actually were, I proceeded to calculate root mean squared error (RMSE) and mean absolute percentage error (MAPE) to see how close my models were to the actual data. Given my lack of future data, I used the last year of data (Feb. 2022 - Feb 2023) as test data for this procedure, allowing me to see if these models' forecasts accurately lined up with what the revenues actually were for alcohol over the last year. And based on our results in Figure 9, we can see that all three of the models were actually quite good, with the Auto ARIMA Model being the best, which makes sense given that it was the most optimistic of the models, and 2022 was quite a good year for alcohol revenues.

	ARMA Model	Auto ARIMA Model	HW Model
RMSE	0.5104988	0.2156558	0.5516259
MAPE	0.02613329	0.01059114	0.03055769

Figure 9.

In all, I am very happy with the final output of this project. My main goal was to produce an accurate forecast for alcohol revenues for the next 12 months, and while the three models do slightly differ, they all are accurate and provide an interesting perspective on where alcohol revenue might go. The industry, while obviously super profitable, is also highly competitive, making it very difficult for newer brands to make a splash. However, as we've seen over the past five years with the emergence of several seltzer brands, there are unexplored niches within the industry, so it's far from likely this revenue is nearing any sort of plateau.

Moving forward, I'd love to try a similar experiment with data from a different country with a different drinking age. The United Kingdom is one that could be quite interesting given the decently sized population as well as how commonplace drinking is in their daily lives. I'd be curious to see if we'd see a similar trend, or if it's on less of an upswing due to the lack of public marketing of newer brands seen across the Atlantic.

VI. References

[1]
https://news.gallup.com/poll/467507/percentage-americans-drink-alcohol.aspx#:~:text=Percentage%20of%20U.S.%20Adults%20Who%20Drink%2C%20Trend%20Since%201939&text=Line%20graph%20showing%20Americans'%20use,reading%2C%20from%202021%2D2022.

[2] https://fred.stlouisfed.org/series/S4248SM144SCEN

VII. Appendix

[1]