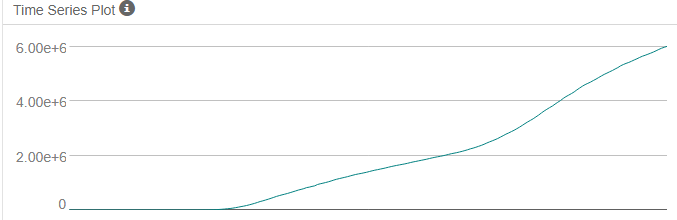
The goal is to find a model to forecast the next 30 days of total COVID cases. We will do so for the United States. I will compare between ETS and ARIMA models.  
  
The characteristics of a time series dataset consists of:

* Continuous data over a long period of time
* The data is in sequential order
* Every consecutive pair of points are one day apart from each other
* There is at most one value per date listed.

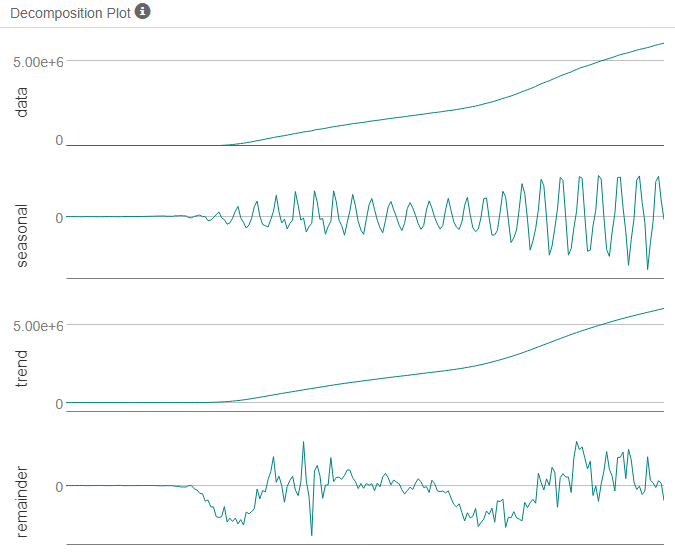
We are also attempting to provide a forecast for the following 30 days; hence we will hold out the last samples.

If we look at the time series plot below, we can see how there is an upward trend occurring. We cannot see if there is seasonal pattern from the plot below, but we will look more into this in the decomposition plot. There does not appear to be any cyclical pattern occurring in the data.



Below, the decomposition plot confirms the upward trend. There also appears to be a seasonal pattern within the graph. Given our seasonal findings when using an ARIMA model we should find the seasonal difference. When using an ETS model, we can see that the magnitude changes for the seasonal component, hence I will consider using a multiplicative method, but will still compare to the additive method.

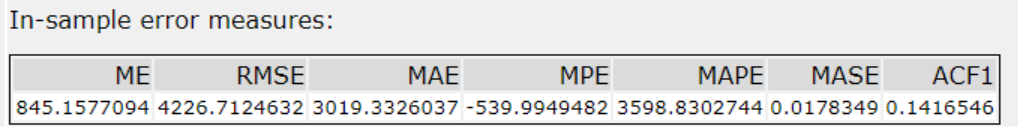
Finally, when looking at the error plot, the error does not stay consistent throughout the time series plot. It would be best to apply error with a multiplicative method when using the ETS model but will still compare to the additive method.

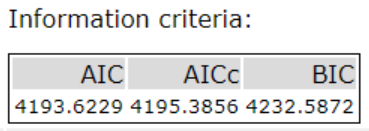


## ETS MODEL

Earlier we mentioned how we were considering multiplicative methods for error and seasonality with an additive method for the trend. We end up with an extremely high error, so I compared to model with all additive methods.

This results in an ETS(A, A, A) model.





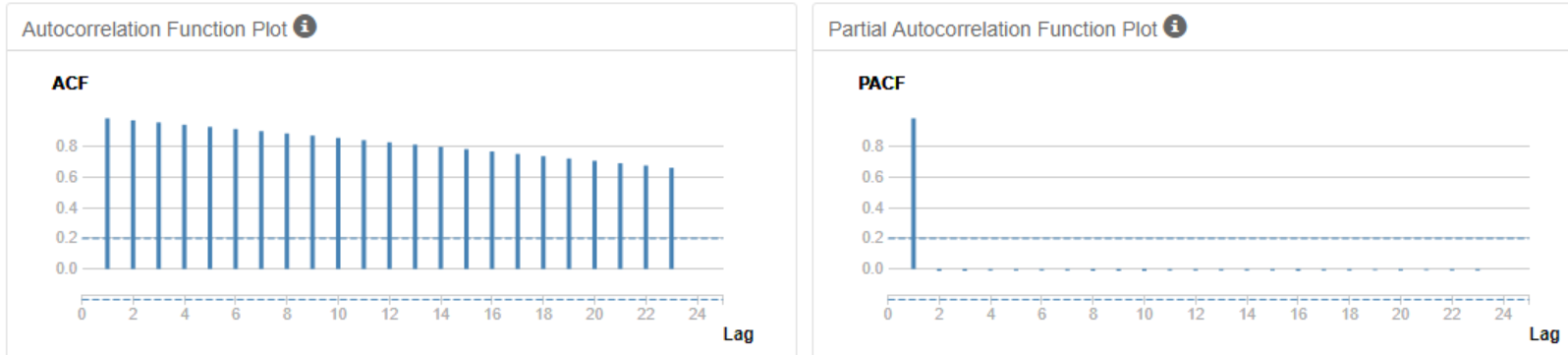
The model results with a RMSE value of 4,227 units around the mean. The MAE is 3,019 units around the mean. We can also see the values for the AIC and BIC are 4194 and 4233, respectively.

## ARIMA MODEL

From our previous analysis we will use an ARIMA(p, d, q)(P, D, Q)S model to forecast.

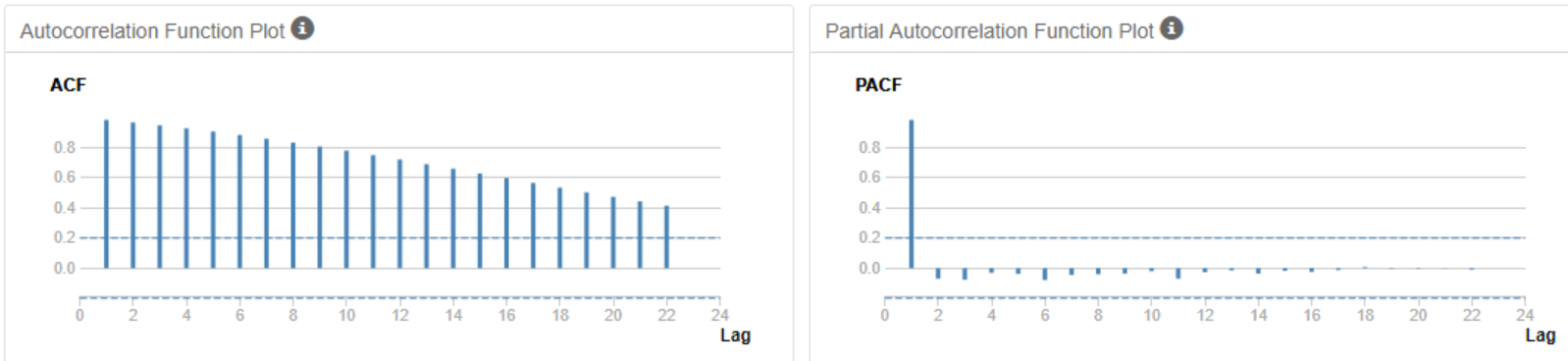
Time Series ACF and PACF:

From the ACF we can see how the data is decreasing the more we proceed. It would be wise to the seasonal difference in the series.



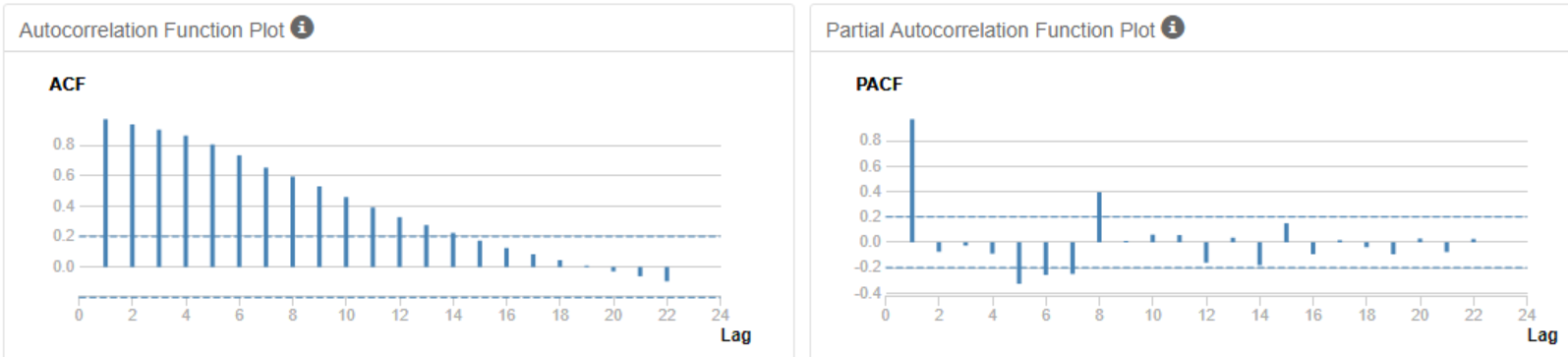
Seasonal Difference ACF and PACF:

We can see similar results to the ACF and PACF from the initial plots without differencing. The only difference is that the correlation decreased. We will take another difference to remove correlation.



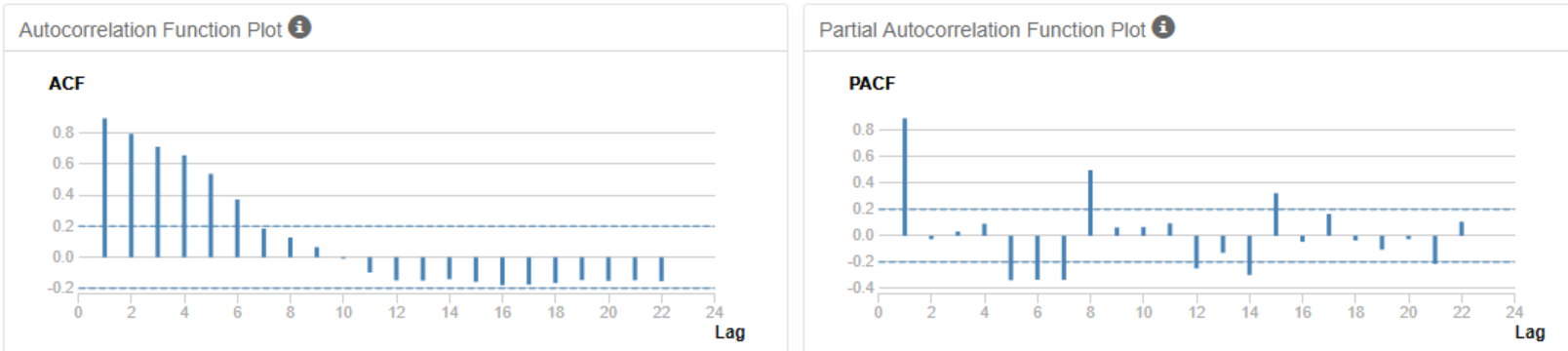
Seasonal First Difference ACF and PACF:

We can see that the results for the ACF and PACF decreased a lot more and its actually decaying towards 0. We will take another difference.



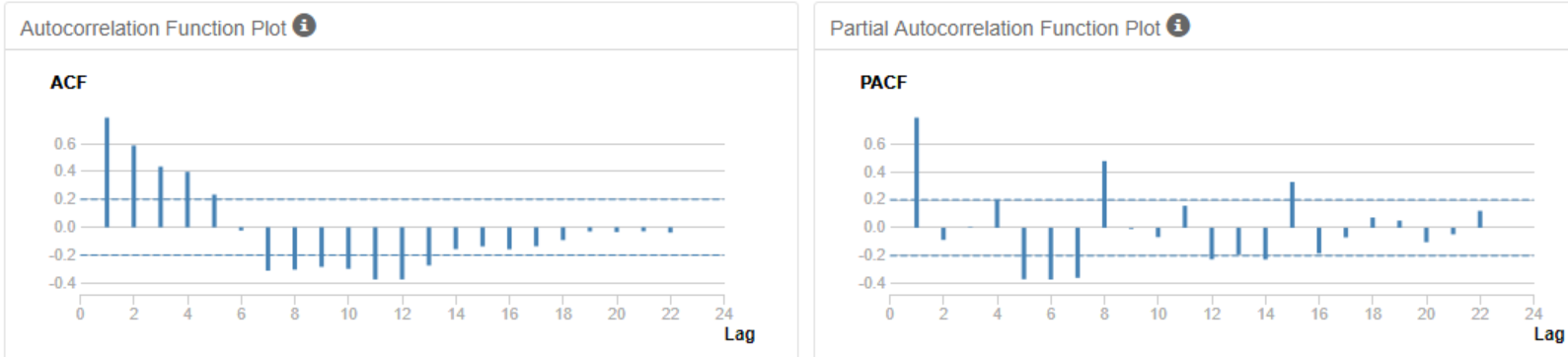
Seasonal Second Difference ACF and PACF:

The correlation continues to decay more, hence taking another difference would be wise.



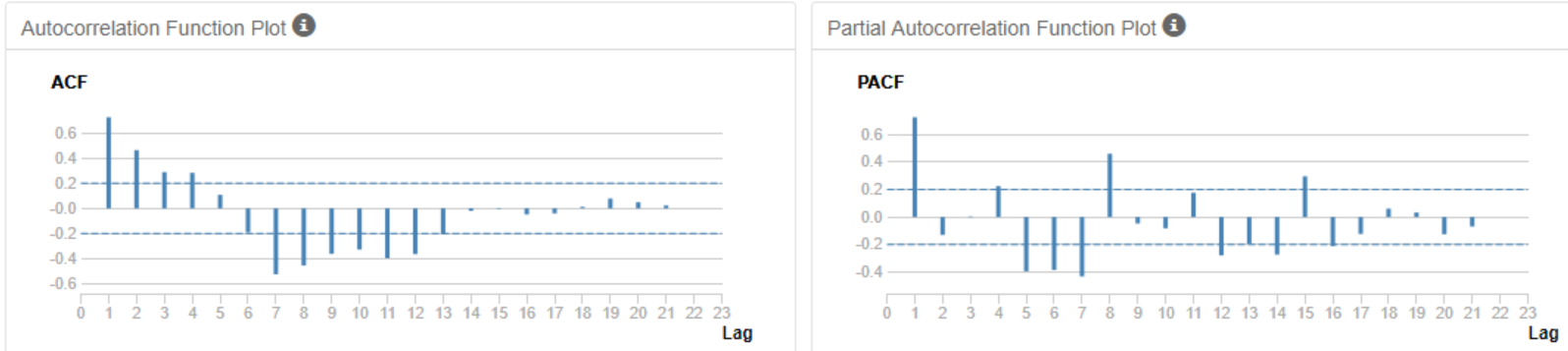
Seasonal Third Difference ACF and PACF:

The correlation continues to decay; hence I will consider taking one more difference.



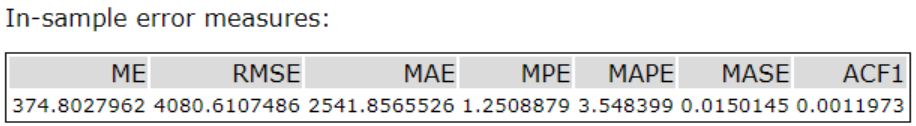
Seasonal Fourth Difference ACF and PACF:

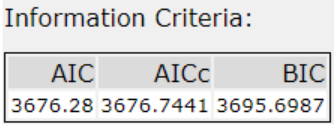
Although the correlation was decreasing, we can see how it also started increasing again towards the center of the ACF plot.



Given that we could not ultimately decide what terms to use for the ARIMA model, I went ahead and allowed the program to decide what parameters to use. This resulted in: ARIMA(0, 2, 1)(0, 0, 4)[7]

Now we will look at the in-sample errors to provide a closer look at the model accuracy.

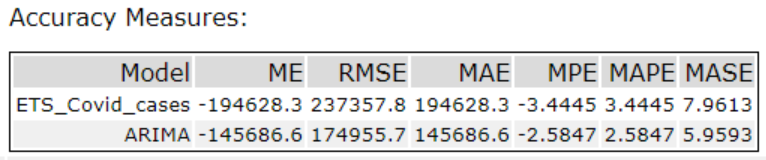




The model results with a RMSE value of 4,081 units around the mean. The MAE is 2,541 units around the mean. We can also see the values for the AIC and BIC are 3676 and 3696, respectively.

### MODEL COMPARISON:

Referring to our in-sample errors we can see how the RMSE, MAE, AIC, and BIC values are all smaller for the ARIMA model. Below we can further compare and find the same results where the error is smaller for the ARIMA model.



Therefore, we will use the ARIMA model for the forecasting.

### FORECAST

Below we have our forecasted values for the next 30 days.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| date | forecasts | forecasts\_high\_95 | forecasts\_high\_80 | forecasts\_low\_80 | forecasts\_low\_95 |
| 9/2/2020 | 6068387.33 | 6076754.08 | 6073858.05 | 6062916.61 | 6060020.58 |
| 9/3/2020 | 6107255.49 | 6121644.41 | 6116663.90 | 6097847.07 | 6092866.56 |
| 9/4/2020 | 6147072.66 | 6167891.45 | 6160685.33 | 6133459.99 | 6126253.88 |
| 9/5/2020 | 6190981.58 | 6218756.57 | 6209142.67 | 6172820.49 | 6163206.59 |
| 9/6/2020 | 6230906.46 | 6266168.79 | 6253963.26 | 6207849.65 | 6195644.12 |
| 9/7/2020 | 6267100.21 | 6310363.06 | 6295388.27 | 6238812.15 | 6223837.36 |
| 9/8/2020 | 6302579.84 | 6354333.89 | 6336420.00 | 6268739.69 | 6250825.79 |
| 9/9/2020 | 6339508.48 | 6401638.29 | 6380132.98 | 6298883.98 | 6277378.67 |
| 9/10/2020 | 6378247.62 | 6451478.82 | 6426130.93 | 6330364.31 | 6305016.42 |
| 9/11/2020 | 6416843.43 | 6501844.06 | 6472422.35 | 6361264.51 | 6331842.81 |
| 9/12/2020 | 6458610.50 | 6556004.26 | 6522292.86 | 6394928.14 | 6361216.73 |
| 9/13/2020 | 6497512.54 | 6607887.66 | 6569682.95 | 6425342.13 | 6387137.42 |
| 9/14/2020 | 6533102.98 | 6657018.31 | 6614126.86 | 6452079.10 | 6409187.65 |
| 9/15/2020 | 6568101.94 | 6706091.48 | 6658328.45 | 6477875.43 | 6430112.40 |
| 9/16/2020 | 6605127.36 | 6758965.84 | 6705716.94 | 6504537.78 | 6451288.88 |
| 9/17/2020 | 6642551.43 | 6813071.42 | 6754048.46 | 6531054.40 | 6472031.44 |
| 9/18/2020 | 6680655.43 | 6868631.68 | 6803566.49 | 6557744.36 | 6492679.18 |
| 9/19/2020 | 6719336.66 | 6925496.52 | 6854137.34 | 6584535.98 | 6513176.81 |
| 9/20/2020 | 6757640.92 | 6982672.26 | 6904781.01 | 6610500.84 | 6532609.59 |
| 9/21/2020 | 6793426.43 | 7037983.82 | 6953333.91 | 6633518.94 | 6548869.04 |
| 9/22/2020 | 6829562.63 | 7094272.14 | 7002646.87 | 6656478.39 | 6564853.12 |
| 9/23/2020 | 6866261.00 | 7152703.07 | 7053555.40 | 6678966.60 | 6579818.93 |
| 9/24/2020 | 6903258.97 | 7212307.86 | 7105335.18 | 6701182.76 | 6594210.08 |
| 9/25/2020 | 6940843.43 | 7273325.42 | 7158241.72 | 6723445.15 | 6608361.45 |
| 9/26/2020 | 6978580.41 | 7335280.40 | 7211814.00 | 6745346.83 | 6621880.43 |
| 9/27/2020 | 7016089.49 | 7397756.50 | 7265648.14 | 6766530.84 | 6634422.47 |
| 9/28/2020 | 7052685.83 | 7460037.56 | 7319038.81 | 6786332.84 | 6645334.09 |
| 9/29/2020 | 7089006.66 | 7522733.21 | 7372605.21 | 6805408.12 | 6655280.12 |
| 9/30/2020 | 7125793.83 | 7587075.94 | 7427409.99 | 6824177.67 | 6664511.72 |
| 10/1/2020 | 7162580.99 | 7652227.21 | 7482743.44 | 6842418.54 | 6672934.78 |

