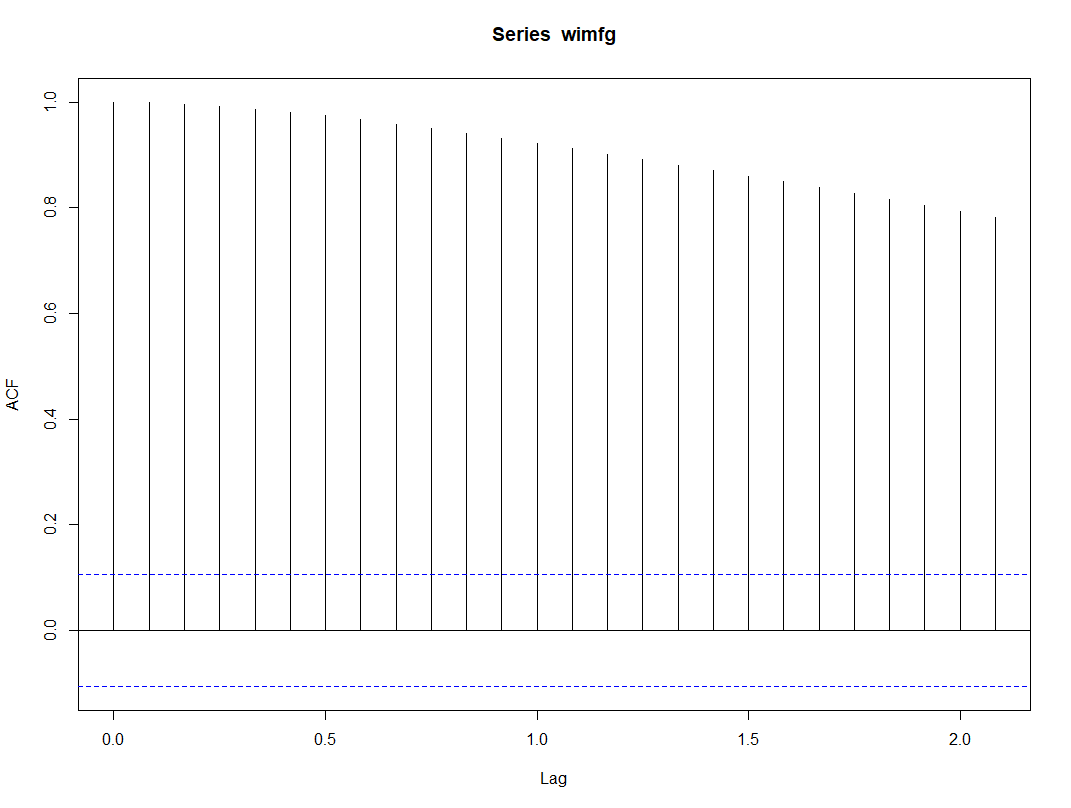
Time Series Analysis on Wisconsin Manufacturing Employment

In this project we are working with a dataset populated with manufacturing employment data in Wisconsin. Using this data, we were tasked with performing time series analysis and ultimately create a forecast which could accurately predict where Wisconsin manufacturing employment would be at a given time. To accomplish this goal, we will try fitting a few different models to our data and use measurements such as the RMSE and MAPE to decide which one is best. Once we find the model that best fits a training dataset, we will fit that model to the full dataset and create a forecast we can use to predict future values of Wisconsin Manufacturing.

To begin, we fit an ordinary time series to the dataset we were given and then plotted the time series and the ACF of the time series. These graphs are provided below.

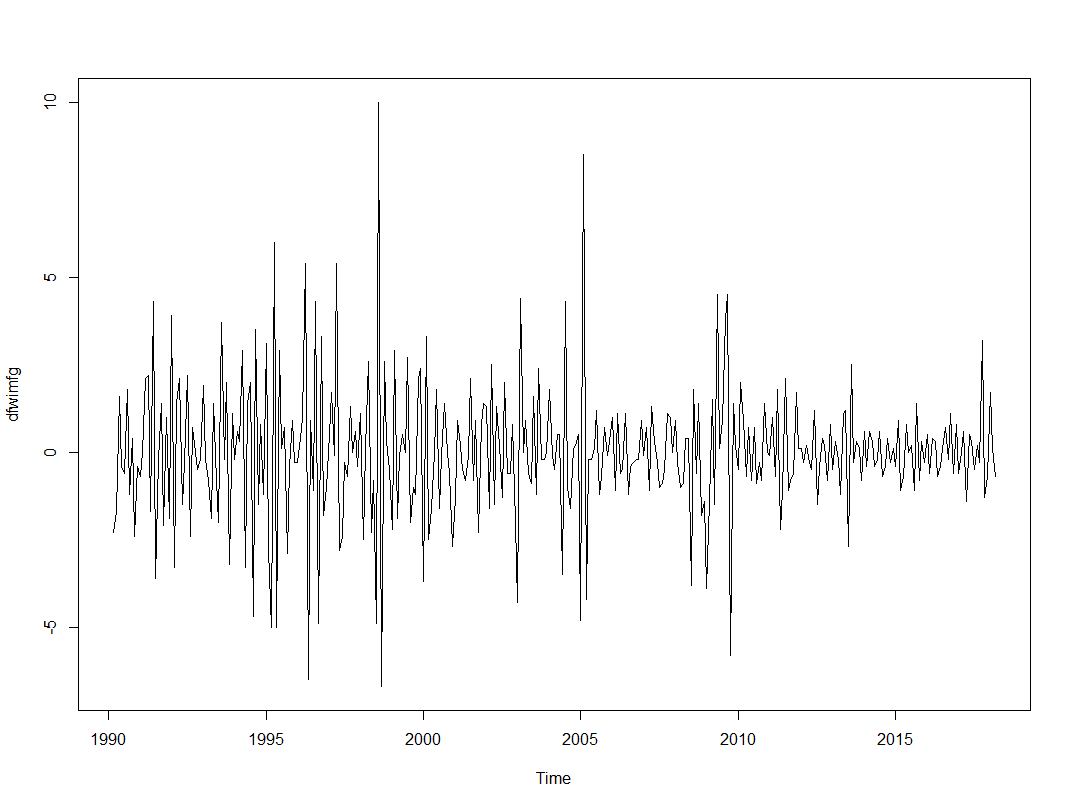
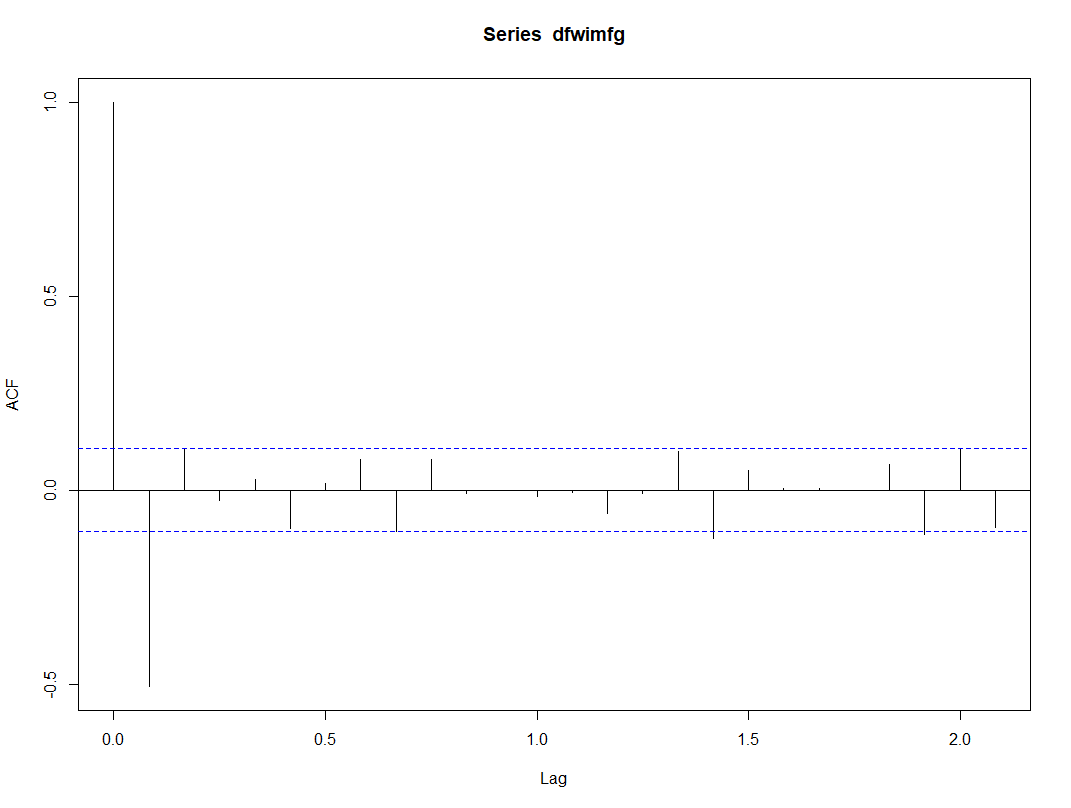


By just taking a glance at the plot of our time series, it appears that Wisconsin manufacturing employment may be in a downtrend. This gives us suspicion to believe that this time series is not stationary but to really know we must also analyze the ACF plot of the series.

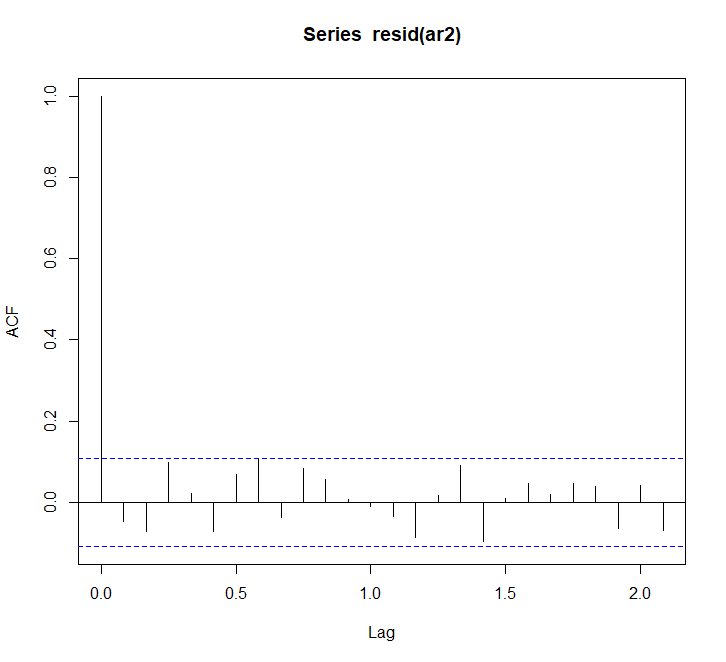


The above ACF plot gives us even more of a reason to believe that the time series we have fit to the data is not stationary. We know this because none of the ACF values plotted fall within our 95% limit (dotted blue lines). Thus, we are going to have to transform our time series to make it as close to stationary as possible before we move on.

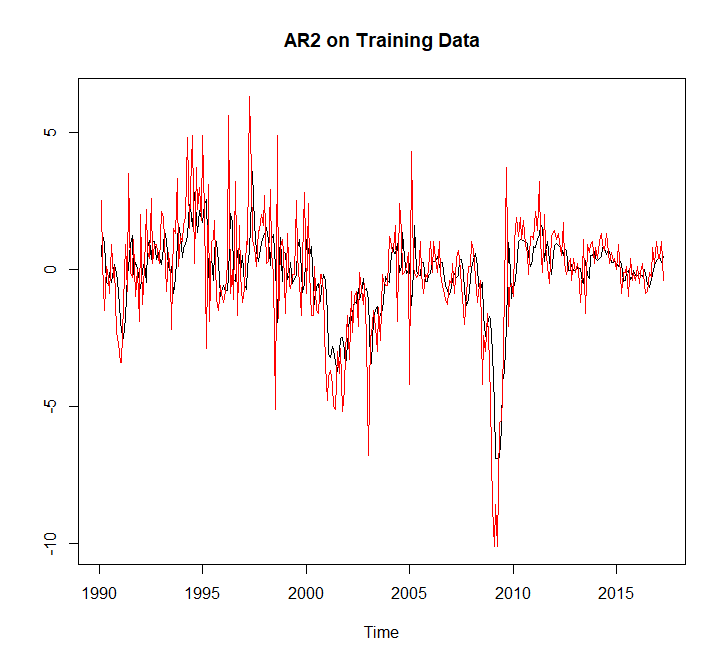
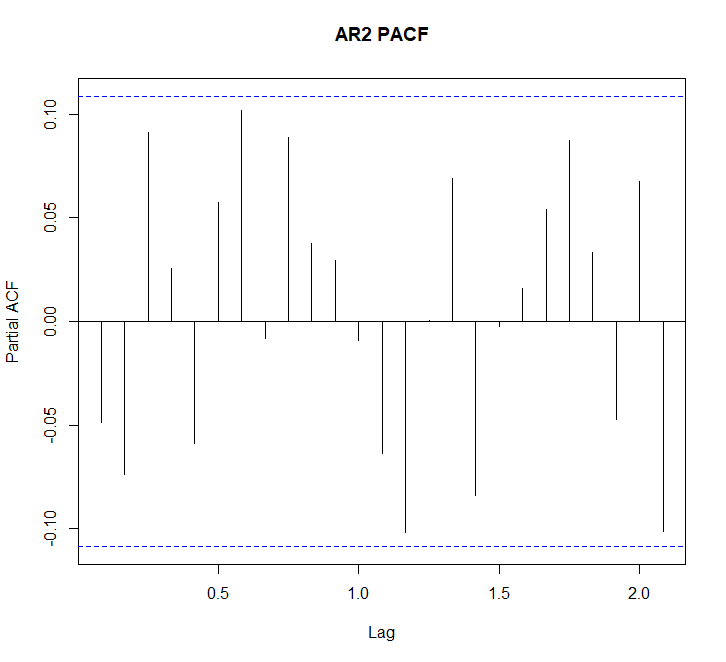
To transform the data we decided to take the second differences series. We chose to use second differences because it is the transformation that has seemed to work best for us in the past, and after transforming the data it was evident that we had found a transformation that would give us something very close to a stationary series. We know this because in the graphs below there is no identifiable trend in the plot of the time series and furthermore, roughly 95% of our ACF values fall within our 95% limits. Also P(0) = 1.

After taking a good look at the ACF plot above, we felt it was appropriate that we fit an AR2 model to the training data since the ACF at P(2) was the only value plotted that exceeded our 95% limit. After fitting the AR2 model to our data, we found that both AR1 and AR2 were significant and p-values of 3.780e-12 and 2.377e-13, respectively. The ACF plot of the residuals for our AR2 model shows that P(0) = 1, and all the other plotted ACF values fall within our 95% limit. This plot can be seen below.

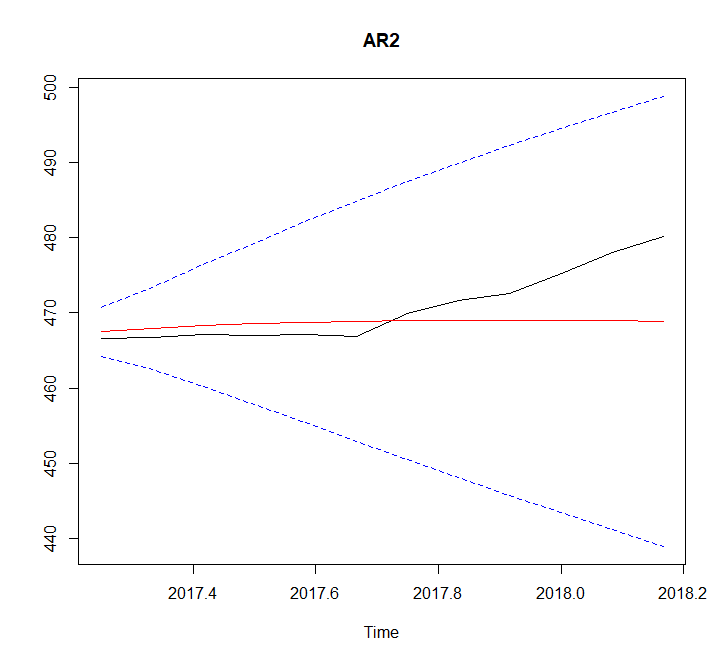


Furthermore, every plotted value on our PACF looks consistent since each falls within the set limit, and when we create a time series plot of the residuals of our AR2 model there is no identifiable trend and our AR2 model (red line) seems to fit our data reasonably well with the only exception being at the extremities of the original data.

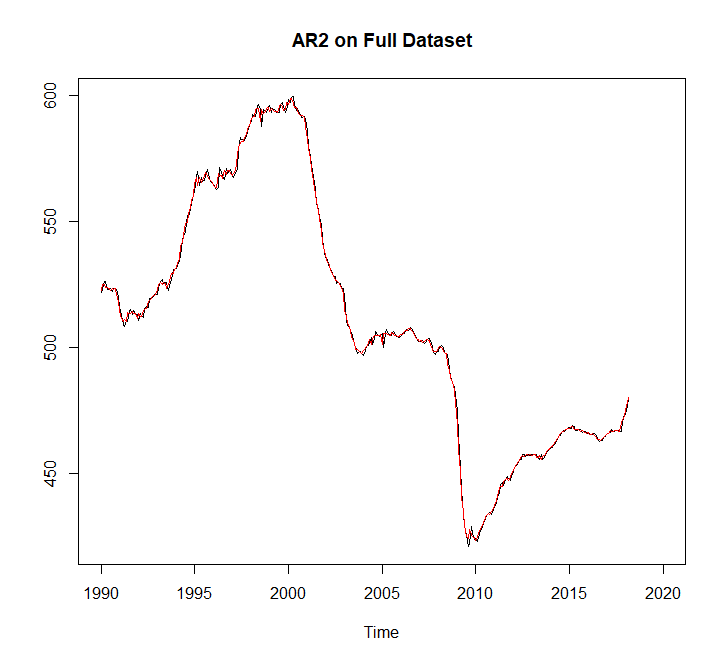
 

We also checked to see if an AR3 model would be a good fit on the training data and found that although AR3 was significant, its magnitude of significance was nearly nothing compared to AR1 and AR2. Although the AR3 model had a slightly better AIC and BIC we still decided that we would continue on with creating a forecast for our AR2 model, since it was found to be the most significant.

After we created a forecast for our AR2 Model on the hold-out data, we found the RMSE on our test data to be about 4.88 and the MAPE on our test data to be about 0.75. A plot of our forecast against the fitted values of our model can be seen below.



When we proceeded with fitting our AR2 model to the full dataset and created a forecast for it, the following graph was produced (red line is forecast, black is original data).



As we can see above, creating a forecast for an AR2 model fit to our full dataset seems to fit very well. Seeing this gave us confidence that we would be able to use this forecast to accurately predict the value of Wisconsin’s manufacturing employment in April and May of 2018. We found that our forecast predicted the manufacturing employment would be roughly between 485 and 486 in April and May of 2018. Our graph supporting this conclusion can be found below (red line is forecast).

