Stat 790 Fifth Meeting

DL

3/11/2022

Things to consider for next meeting

- Write out the model of ARIMA(0,0,3) with interpretations

This model was returned using the search method from the as the best model, in terms of lowest RMSE for the Bronx total waste values

$$\hat{y} = 41207.0951 - 0.334e_{t-1} - 0.3269e_{t-2} - 0.2368e_{t-3}$$

This tells us that this model uses 3 previous observations to calculate current observations.

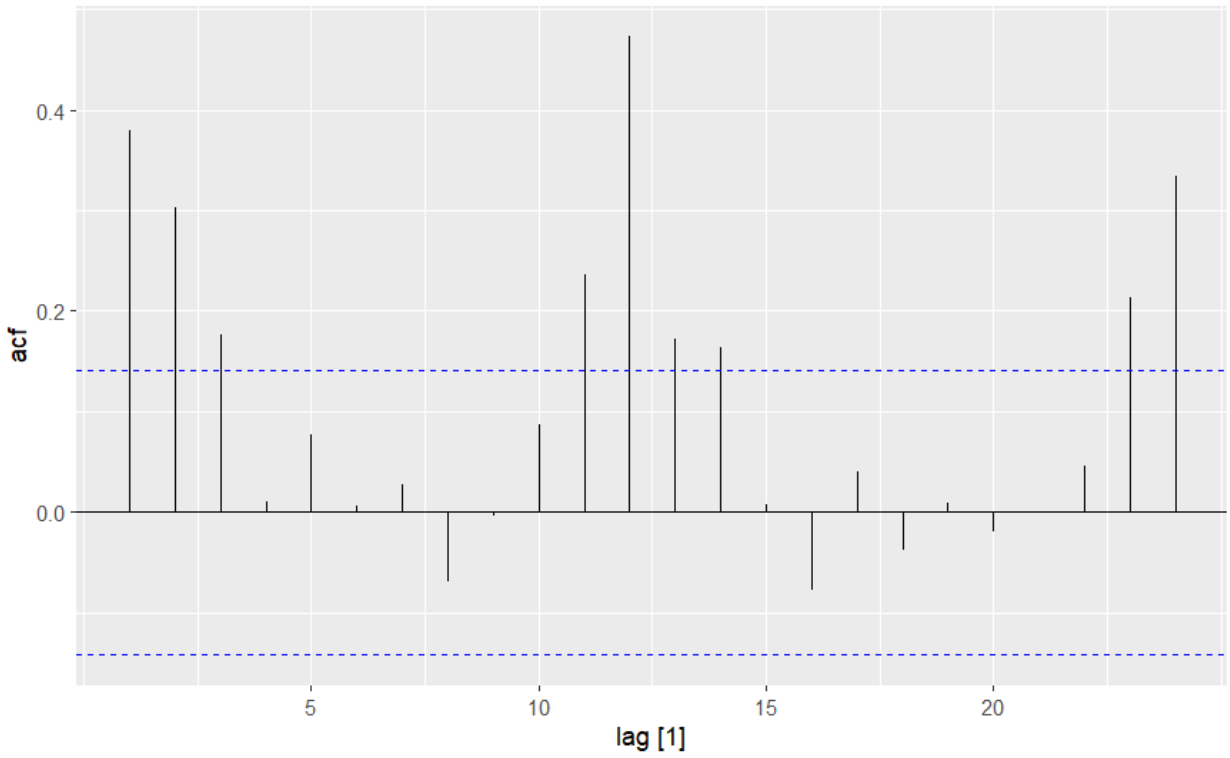

Figure 1: ACF Plot

We can see a downward trend from lags 1-3. And an upward trend from lags 10 to 12. I have learned that if the series has positive autocorrelations out to a high number of lags, then it probably needs a higher order of differencing. This is not the case with our ACF plot, as the ACF values quickly go to zero around lag 4. I

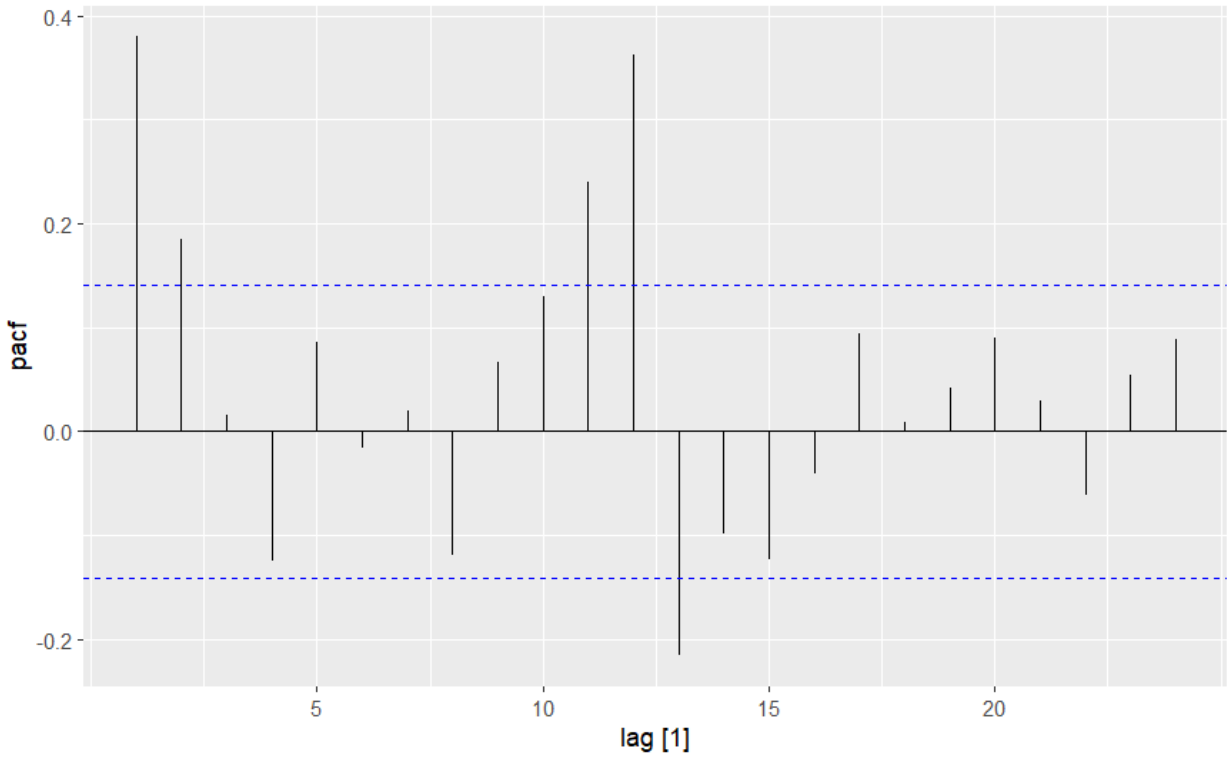

Figure 2: PACF Plot

also don't believe any AR terms should be the only terms included, this would imply that AR forecasts are a linear function of the coefficients as well as a linear function of past data.

Interpreting the PACF plot is a bit more complicated. But the PACF values contain smaller negative values, when compared to the ACF plot.

- Identify the the PDQ arguments to account for seasonality
  - Attempts at adding a seasonal component with the PDQ() argument included "1+ pdq(0,0,3) + PDQ(0,0,1)", "1+ pdq(0,0,3) + PDQ(0,0,12)" and "pdq(0,0,3) + PDQ(0,0,12)". The first attempt was to see how what kind of coefficients these PDQ arguments would return. The second attempt was intentional, where 12 represents the total number of months per calendar year. The third attempt was used to see if there was any difference was made removing the constant. But in all, these attempts at adding a seasonal component returned the same model as the ARIMA(0,0,3), in terms of coefficients and accuracy scores.

References:

- Read the Duke website of ARIMA tutorials
  - Use articles from "5. ARIMA models for time series forecasting," to identify the pdq arguments manually
  - use the residuals plots, along with ACF plots to help you with the choice of pdq arguments
- Continue reading this article which details the differences of the fpp2, fpp3 and modeltime & modeltk packages using data from the fpp3 package and textbook.