Hi Sunny –

For this next project, I would like to try an ARIMA and SARIMA model on some data related to my work for my actual job for PeerLogix. We track television and movie viewership data, and I’d like to try and predict future viewership of a show based on historical viewership data.

I propose we try a ARIMA or SARIMA model instead of one of the VAR models because no two shows should be related to each other, so I think it makes sense to study each individually as opposed to inputting data as a vector where we assume each time series may be related to each other (like we did for stock data).

I’ve included viewership data for a few TV shows going back for about 3 years. Can we try applying the models to each and see how the results look? I think one question I’m also trying to figure out is if its best to aggregate the data by daily values (as they are now) or to instead aggregate by weekly data. I also understand that we may be limited here by the amount of data history we have (only 1000 rows in most cases), so not sure aggregating is possible but it is something I wanted to raise that I think would be interesting to look at.