MKT6971

Time Series Project

Name:

The objective of this project is for you to practice what you have learned about time series analysis and interpreting data. I suggest you use GRETL for this project.

1. Select a scientific, biomedical, business or other issue that appeals to you and go looking online for relevant time series data sets. The good news here is that there are tons of free and interesting time series data sets online. If you have problems locating them let me know and I will help. **Be sure that it looks like there is little or no seasonality to it.**
2. Plot out your time series variable. Tell me using your Mark I eyeball whether or not you think the time series data set is stationary in terms of **constant mean** and also **constant variance**.
3. Plot the ACF for the time series data set. Looking at ACF, does it look like there may be a trend or non-constant mean for each time series?
4. Now let’s examine the time series data set using unit root tests. First use the KPSS test for the time series data set and tell me if the test suggests if there is a constant mean or not. Then see if you can confirm your KPSS evaluation using the Augmented Dickey Fuller (ADF) or the ADF-GLS test and tell me what the ADF test suggests is the case.
5. Summarize the results of steps 2 through 4 and tell what your decision is regarding constant mean in the time series data set.
6. Review the decision in step #5. If the test suggests that there is a non-constant mean then use differencing to create a new differenced variable for the time series data set. Otherwise proceed to steps a through c below:
   1. Plot out the data for the new differenced data set. Tell me if it looks like the differencing got rid of the trend or non-constant mean.
   2. Plot the ACF for the differenced time series. Tell me if this new ACF plot looks like there now is no trend.
   3. Apply the KPSS test and the ADF or ADF-GLS test to the differenced data – does the trend disappear?

**Note: From this point onward through step 9, if the time series was differenced, use the differenced time series data set for all the rest of the questions. Otherwise you can use the undifferenced data set.**

1. Plot the PACF for the time series data set. Using the combined information from the ACF you plotted earlier along with the information in the PACF, tell me if you see autoregressive and/or moving average processes in the data set. To help with interpretation you may want to refer to online resources – here is a decent resource from Duke University [**https://people.duke.edu/~rnau/411arim3.htm**](https://people.duke.edu/~rnau/411arim3.htm) or Penn State https://onlinecourses.science.psu.edu/stat510/node/64
2. For your time series data set, experiment with different ARIMA models for them. Try at least four models. As you try them, list out the results of the various models and
   1. Comment on how each one is working and compare it to the previous model using various metrics such as AIC, BIC (Schwartz) and Box Leung Most students end up creating a small table with these statistics across the models tried so it is easy to compare them.
   2. Plot the observed versus fitted data for the time series data set **for each model** and comment on how well the model seems to be working
   3. Pick one of the models as your favorite and tell me why you like that one the best.
   4. Use the ARCH test to tell me if your best model has any issues with autoregressive conditional heteroscedasticity.