For this project you will fit a forecasting model with trend, seasonal dummies, and cycles. You may choose any data of your choice (except any data used in class and/or previous homework assignments) provided your time-series data suggest the presence of all three components. To make sure you capture any dynamics related to cycles, you will need a large time horizon, therefore, you might want to make sure that your observations span at least $\sim$ 10 years. In addition, since you will also try to fit a VAR model to your data, you will need to have at least two time-series variables, e.g., S&P500 returns from 1985-present and Apple returns from 1985-present with the same frequency.

The assignment that you will submit, consists of a written report which includes answers to the respective questions (including plots), and source code (must be your own) you wrote to execute the computations.

Your report needs to be typed (no limit on the number of pages) and will consist of 5 parts:

I. (5%) Introduction (describe the data, provide some background on the topic, etc.).

II. (80%) Results (answers and plots).

III. (5%) Conclusions and Future Work.

IV. (5%) References (include the source of your data and any other resources).

V. (5%) R Source code. Although the code is only worth 5%, if you do not submit your code, you will not receive credit for the assignment.

1. Modeling and Forecasting Trend, Seasonality, and Cycles

For each of your time-series variables, compute the following:

1. Produce a time-series plot of your data including the respective ACF and PACF plots.
2. As a baseline model, fit an ARIMA model to each series and comment on the fit. For the next questions, you will instead use the model estimated in (3) for their respective answers.
3. Fit a model that includes, trend, seasonality and cyclical components. Make sure to discuss your model in detail.
4. Plot the respective residuals vs. fitted values and discuss your observations.
5. Plot the ACF and PACF of the respective residuals and interpret the plots.
6. Plot the respective CUSUM and interpret the plot.
7. Plot the respective Recursive Residuals and interpret the plot.
8. For your model, discuss the associated diagnostic statistics.
9. Use your model to forecast 12-steps ahead. Your forecast should include the respective error bands.
10. Fit an appropriate VAR model using your two variables. Make sure to show the relevant plots and discuss your results from the fit.
11. Compute, plot, and interpret the respective impulse response functions.
12. Perform a Granger-Causality test on your variables and discuss your results from the test.
13. Use your VAR model to forecast 12-steps ahead. Your forecast should include the respective error bands. Comment on the differences between the two forecasts (VAR vs. ARIMA).
14. Backtest your ARIMA model. Begin by partitioning your data set into an estimation set and a prediction set.
    (a) Use a recursive backtesting scheme, and forecast 12-steps ahead at each iteration. Compute the mean absolute percentage error at each step. Provide a plot showing the MAPE over each iteration.
    (b) Shorten your forecast horizon to only 1-step ahead. Compute the absolute percentage error at each iteration, and plot.
    (c) Based on your findings above, does your model perform better at longer or shorter horizon forecasts?
    (d) Now test your model using a moving window backtesting scheme. Forecast out 12-steps ahead at each iteration, and plot the forecast errors observed at each iteration. Repeat for a 1-step ahead forecast horizon. Provide plots of both.
    (e) How do the errors found using a recursive backtesting scheme compare with the errors observed using a moving average backtesting scheme? Which scheme showed higher errors overall, and what does that tell you about your model?