**Math339** **Fall 2022**

**Final Project**- *due 11:59pm 12/08/2022*

For the selected time series dataset of your interest, perform the following tasks (for univariate time series).

1. Provide the background information from which the data was collected and the problem that the to-be-conducted analysis intends to answer. Explain the potential value of the analysis.
2. Explain the specification of the data, i.e. the time interval over which the data was collected, the type of data (daily, weekly, monthly, or ...).
3. Generate a time-series plot that shows clearly the date on the $x$-axis with the major/minor ticks deemed appropriate. Add appropriate label, title and caption on the graph. If multiple series are plotted on the same plot, different colors or line types should be used for distinction and provide also the legend that explains and matches the aesthetics.
4. (if missing values are present) Perform appropriate imputation (considering the nature and the attributes of the data)
5. Evaluate the stationarity of the series, showing ACF plot, the result from Augmented Dicky-Fuller's test, PP test.
6. (if deemed not stationary in either mean or variance) Perform appropriate transformation and re-evaluate the stationarity of the transformed data.
7. Decompose the time series into time-cycle component, seasonal component and irregular components by X11 procedure.
8. Split the data into training and test set in 70%-30% manner if there is no seasonality (test set is the last 30% of the series). If there seems to be seasonality, the test set expects to contain at least one cycle.
9. Find the best two ARIMA model(s) (could be a mixture of seasonal and nonseasonal ARIMA) for the training series.
10. Perform the diagnostics on the two competing models, evaluating whether the residual series behave like white noise series by Ljung-Box test and the white-noise test from spectral analysis.
11. Apply the two competing models on the test series to compare their forecast error. What is your best final model?
12. Apply Exponential Weighted Moving Average smoother to the data (if no concern on the seasonality; use Holt-Winter's otherwise) in similar setup done above and compare the result to the best model you found in 11.
13. Re-do the literature search, how does your finding compare to others'?

Note that

- Prepare your report in a single PDF file. The number of pages is at least 10 and no more than 20, excluding cover page and appendix.
- All the tables and figures need to be explained, numbered and captioned. Any wandering figures/tables without explanation will cause deduction of the points.
- Reference section is mandatory.
- Place R/SAS/Python-code in the appendix.
- No handwriting!