

# MSDS 6306: Doing Data Science – Time Series

## Live session Unit 12 assignment

**Due: 1 hour before your 14<sup>th</sup> live session**

### Submission

**ALL MATERIAL MUST BE KNITTED INTO A SINGLE, LEGIBLE, AND DOCUMENTED HTML DOCUMENT.** Use RMarkdown to create this file.

Formatting can be basic, but it should be easily human-readable. Unless otherwise stated, **please enable `{r, echo=TRUE}` so your code is visible.**

### Questions

#### **Warm Up: Brief Financial Data (20%)**

- a. Natively in R, you have access to sample data sets of prominent stocks over time. We'll be using EuStockMarkets for this question. Type `help(EuStockMarkets)` to learn more. From these data, pull specifically the DAX index. For all questions in this assignment, you're welcome to normalize (or don't!) how you see fit, but, if you choose to, please document what you're doing and why for the grader. It's not necessary for the purpose of this assignment.
- b. These are annual European Stock Data from 1990 onward. Create a rudimentary plot of the data. Make the line blue. Give an informative title. Label the axes accurately. In 1997, an event happened you want to indicate; add a vertical red line to your plot which divides pre-1997 and post-1997 information.
- c. Decompose the time series into its components (i.e., trend, seasonality, random). Keep in mind that this is a multiplicative model you want. Create a plot of all decomposed components. As before, make all lines blue and have a vertical divider at the year 1997.

#### **2. Temperature Data (40%)**

- a. Using the maxtemp dataset granted by loading fpp2, there are maximum annual temperature data in Celsius. For more information, use `help(maxtemp)`. To see what you're looking at, execute the command in 'Examples' in the help document.
- b. We are only concerned with information after 1990. Please eliminate unwanted information or subset information we care about.
- c. Utilize SES to predict the next five years of maximum temperatures in Melbourne. Plot this information, including the prior information and the forecast. Add the predicted value line across 1990-present as a separate line, preferably blue. So, to review, you should have your fit, the predicted value line overlaying it, and a forecast through 2021, all on one axis. Find the AICc of this fitted model. You will use that information later.

- d. Now use a **damped** Holt's linear trend to also predict out five years. Make sure `initial="optimal."` As above, create a similar plot to 1C, but use the Holt fit instead.
- e. Compare the AICc of the `ses()` and `holt()` models. Which model is better here?

### 3. The Wands Choose the Wizard (40%)

- a. Utilize the `dygraphs` library. Read in both `Unit12TimeSeries_Ollivander` and `_Gregorovitch.csv` as two different data frames. They do not have headers, so make sure you account for that. This is a time series of Wands sold over years.
- b. You don't have your information in the proper format! In both data sets, you'll need to first convert the date-like variable to an actual `Date` class.
- c. Use the library `xts` (and the `xts()` function in it) to make each data frame an `xts` object (effectively, a time series). You'll want to `order.by` the `Date` variable.
- d. Bind the two `xts` objects together and create a `dygraph` from it. Utilize the `help()` index if you're stuck.
  - Give an effective title and x/y axes.
  - Label each Series (via `dySeries`) to be the appropriate wand-maker. So, one line should create a label for Ollivander and the other for Gregorovitch.
  - Stack this graph and modify the two lines to be different colors (and not the default ones!) Any colors are fine, but make sure they're visible and that Ollivander is a different color than Gregorovitch.
  - Activate a range selector and make it big enough to view.
  - Use `dyShading` to illuminate approximately when Voldemort was revived and at-large: between 1995 to 1999.
  - Enable Highlighting on the graph, so mousing over a line bolds it.

### Reminder

To complete this assignment, please submit **one** RMarkdown and matching HTML file at least one hour before your live session. Please submit all files at the same time; only one submission is granted.

Good luck!