# ISYE 6501
# Homework 4

September 18, 2019

## Question 7.1

We would apply exponential smoothing to a time series situation. One example is a series of stock price data, which can be very volatile. Let's say we want to assess if stocks with varying reporting periods (quarterly, semi-annually, or annually) have different rates of return. We would need daily stock price data at the close of every trading day for the past 20 years (for example), as well as the corresponding dates for each price. The exponential smoothing method would retain cyclicality for the different reporting periods, as well as the smoothed performance of each stock. Given the high volatility of stock data, we would expect the value of the smoothing factor $\alpha$ to be closer to 0.

## Question 7.2

In our prior homework assignment, we used the CUSUM method and found no compelling evidence of climate change (=longer summers) in Atlanta over the time-frame 1996-2015. In this homework assignment, we will use a slightly more sophisticated technique to take on the same question.
A quick aside before jumping into the analysis. As mentioned in one of the video, evidence of longer summers would not necessarily be evidence of a warmer climate, but rather evidence of (unspecified) climate change. While the two are often used interchangeably in the public debate, they are different concepts. More so, we as data scientists really want to abide by Popper's falsification principle and refuse to embrace over-arching theories following a single confirmatory observation.
That being said, we will use R's HoltWinters command to smooth our data and obtain parameter estimates relying on its built-in optimization routine. We will compare two different models:

- A double exponential smoothed model which does not assume a trend (and assumes seasonality);
- A triple exponential smoothed model which does assume a trend (and assumes seasonality);

In a nutshell, a necessary condition for arguing for climate change in Atlanta over the considered time-frame is that the second model fits the data better AND estimates a significant trend component ($\beta$). If the second model does not stack up any better against the first model, we would conclude that we do not have evidence supporting climate change in Atlanta.

Getting into the details, our first attempt at HoltWinters sets $\beta$ as FALSE as it does not assume an underlying trend in the time series. However, it does require an estimate of the seasonality parameter $\gamma$ (and sets it as "multiplicative") along with an estimate of the smoothing parameter $\alpha$. Thus, we exponential smooth our initial observations but we assume no upward/downward trend (=no climate change).

```
# reading data
data <- read.delim("temps.txt", header=TRUE)

# installing and loading the HoltWinters command
install.packages("forecast")
library(forecast)

-------------------- DATA PREPARATION --------------------
#Plotting time series

# ts() takes in vectors and returns a time series objects
# the data has to be prepared
# removing the first column and unlisting observations into a vector
# frequency is set to 123 as we are dealing with so many observations per year

str(data)
data_ts <- as.vector(unlist(data[,-1]))
data_ts <-ts(data_ts, start=1996, frequency=123)

-------------------- DOUBLE EXPONENTIAL SMOOTHING --------------------

model_double <- HoltWinters(data_ts, beta = FALSE, seasonal = "multiplicative")
model_double

##Holt-Winters exponential smoothing without trend
##and with multiplicative seasonal component.
##
##Call:
##HoltWinters(x = data_ts, beta = FALSE, seasonal = "multiplicative")
##
##Smoothing parameters:
## alpha: 0.6150515
## beta : FALSE
## gamma: 0.549585

# residual sum of squares
model_double$SSE
##[1] 68905.32
```

The parameter estimates are $\{\alpha, \gamma\} = \{0.62, 0.55\}$ and the residual sum of squares 68905.32. This last one figure is not meaningful per se but will help us out with the comparison we are up to next.

We run the HoltWinters command again requiring triple exponential smoothing (the seasonal component is still set to be "multiplicative"). The value of the estimated $\beta$ will inform us about possible trends. A minimum requirement for supporting the climate change hypothesis is that $\beta$ is non-zero.

```
model_triple <- HoltWinters(data_ts, seasonal = "multiplicative")
model_triple
##Holt-Winters exponential smoothing with trend
##and multiplicative seasonal component.
##
##Call:
```

```
##HoltWinters(x = data_ts, seasonal = "multiplicative")
##
##Smoothing parameters:
## alpha: 0.615003
## beta : 0
## gamma: 0.5495256

# residual sum of squares
model_triple$SSE
##68904.57
```

Astoundingly, the estimate for $\beta$ is zero, and the estimation performance of the triple exponential smoothed model $\{\alpha, \gamma, SSE\} = \{0.62, 0.55, 68904.57\}$ is almost indistinguishable from that of the prior model without a trend component. Based on this test alone, we cannot infer that climate change is real. However, as previously stated, we really want to be very cautious and stick to the some basic epistemology. Our dataset is too narrow and limited in scope to prove/disprove such an encompassing theory as that of climate change/global warming is.