Question 7.1
Describe a situation or problem from your job, everyday life, current events, etc., for which exponential smoothing would be appropriate. What data would you need? Would you expect the value of α (the first smoothing parameter) to be closer to 0 or 1, and why?

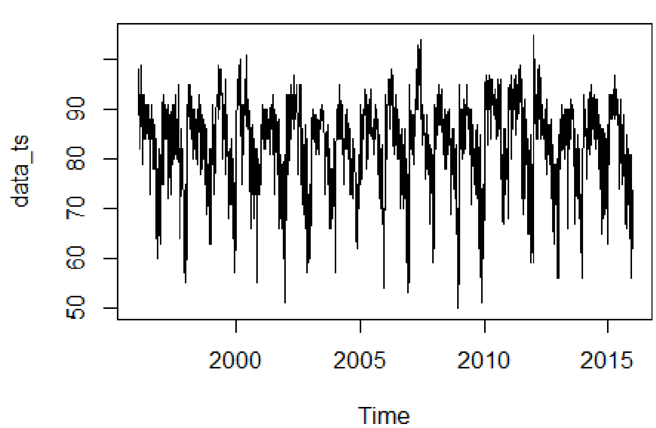
I think an exponential smoothing model could be quite valuable in projecting sales for businesses. As the exponential smoothing model assigns a greater weight to more recent data than older data, it could be used to get smooth data which more clearly shows the trends present.

In terms of implementing such a model, it might be necessary to have sales revenue data for the previous few years. An exponential smoothing model might also be quite valuable because I can look into seasonality with it (e.g., does the business cycle through lulls in revenue?), and the trends would be valuable in showing how revenue changes with time. Also, since there might be variability when it comes to providing services (e.g., if something goes wrong in sourcing parts or supplies), I think using an alpha value closer to 1 is not the best idea. However, since it shouldn't be completely random either, I think using a middle ground value or a value near there like 0.5-0.7 might be good for alpha.

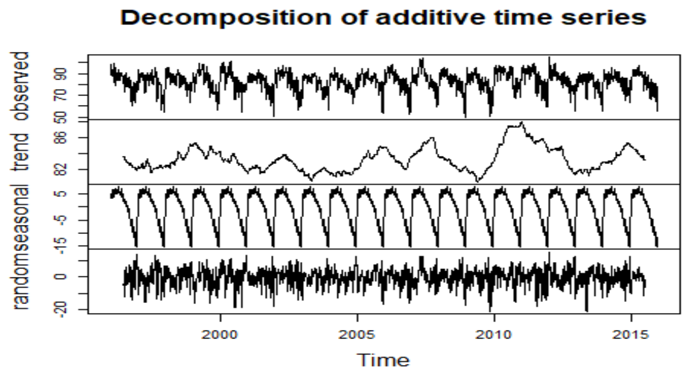
Question 7.2
Using the 20 years of daily high temperature data for Atlanta (July through October) from Question 6.2 (file temps.txt), build and use an exponential smoothing model to help make a judgment of whether the unofficial end of summer has gotten later over the 20 years. (Part of the point of this assignment is for you to think about how you might use exponential smoothing to answer this question. Feel free to combine it with other models if you'd like to. There's certainly more than one reasonable approach.)

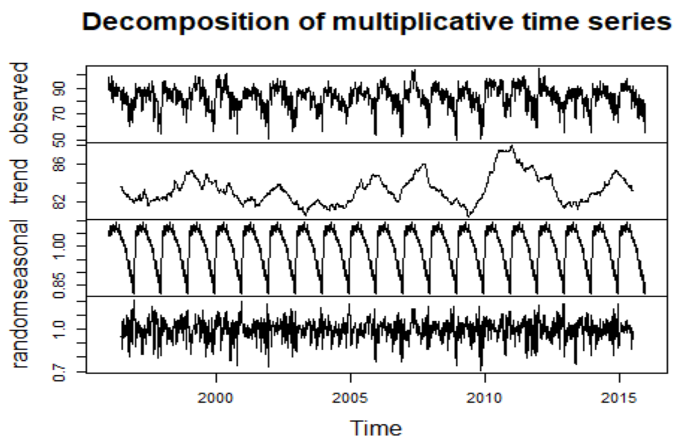
Note: in R, you can use either HoltWinters (simpler to use) or the smooth package's es function (harder to use, but more general). If you use es, the Holt-Winters model uses model="AAM" in the function call (the first and second constants are used "A"dditively, and the third (seasonality) is used "M"ultiplicatively; the documentation doesn't make that clear).

The code for this approach can be found in appendix 7.2.1. In order to build and use an exponential smoothing model, I first began by creating a timeseries for the data and then plotting it:

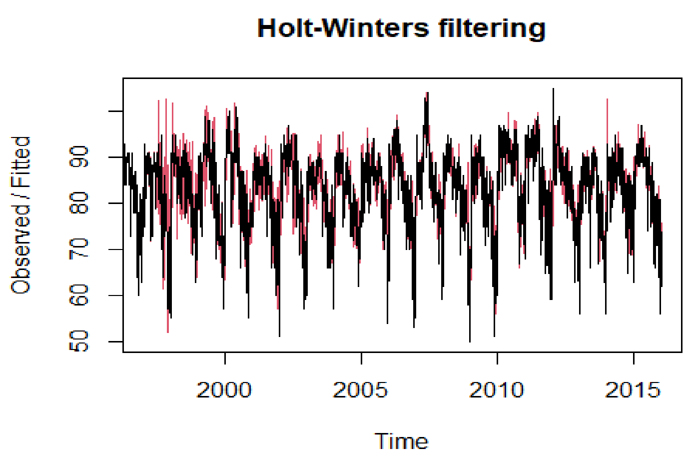


Once I had finished that, I performed an additive and multiplicative decomposition:

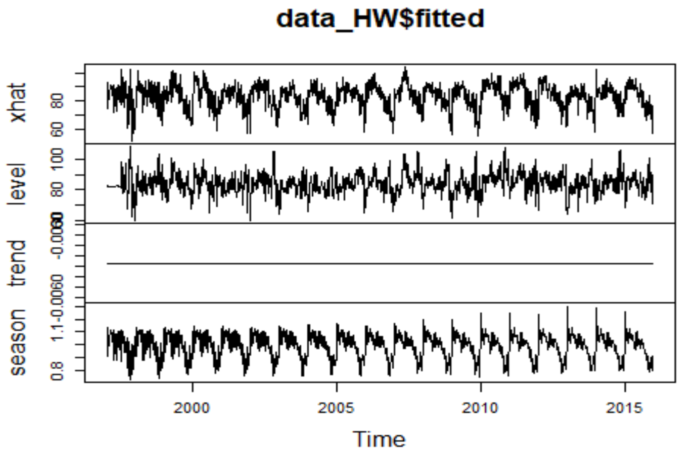




I also then proceeded to create a Holt-Winters model for this data using seasonality set to multiplicative:



Alongside this, I plotted the fitted values obtained as well:



Here, what particularly caught my eye is that the trend estimate is very close to zero, which to me suggests that the data is not really showing any significant increases or decreases over the time period.

Then, to address whether the unofficial end of summer has gotten later over the years, I developed a CUSUM model to investigate. For this, I decided that I would use the seasonal factors and see if there is a change (decrease) detected for each year. When I ran the code, I got the following:

```
## The year is: 1 and the date is: 41
## The year is: 2 and the date is: 41
## The year is: 3 and the date is: 40
## The year is: 4 and the date is: 41
## The year is: 5 and the date is: 42
## The year is: 6 and the date is: 42
## The year is: 7 and the date is: 43
## The year is: 8 and the date is: 43
## The year is: 9 and the date is: 42
## The year is: 10 and the date is: 42
## The year is: 11 and the date is: 43
## The year is: 12 and the date is: 43
## The year is: 13 and the date is: 44
## The year is: 14 and the date is: 44
## The year is: 15 and the date is: 45
## The year is: 16 and the date is: 46
## The year is: 17 and the date is: 68
## The year is: 18 and the date is: 46
## The year is: 19 and the date is: 46
```

This is essentially indicating that for every single year a change was detected. And the date of the change usually is around the 9th of August to the 15th. There is only one case where the date differs significantly and that is in 2013 (year 17). Here, the date change detected is September 6th. Ultimately though, looking at the majority of the data, I would say there is not really a change in the end of unofficial summer.

I also decided that it would be worthwhile to create a Holt-Winters model with seasonality set to additive and see what that would result in. The code for this process is found in appendix 7.2.2. What I was first surprised about was that in comparing the SSE for the multiplicative approach (68904.57) to the SSE for the additive approach (66244.25), the additive approach's SSE was actually lower! Continuing, when I looked at the fitted values, I also noticed that the values for trend are very close to zero in this situation as well, which indicates that the data is not really showing a lot of significant increases or decreases across the time period. Finally, I created a CUSUM analysis for this situation. My output for change (decrease) detection was as follows:

```
## The year is: 1 and the date is: 41
## The year is: 2 and the date is: 41
## The year is: 3 and the date is: 40
## The year is: 4 and the date is: 41
## The year is: 5 and the date is: 42
## The year is: 6 and the date is: 42
## The year is: 7 and the date is: 43
## The year is: 8 and the date is: 42
## The year is: 9 and the date is: 42
## The year is: 10 and the date is: 42
## The year is: 11 and the date is: 43
## The year is: 12 and the date is: 43
## The year is: 13 and the date is: 43
## The year is: 14 and the date is: 44
## The year is: 15 and the date is: 45
## The year is: 16 and the date is: 45
## The year is: 17 and the date is: 46
## The year is: 18 and the date is: 46
## The year is: 19 and the date is: 46
```

As can be seen, a decrease is detected for every single year and the dates of the change detection are pretty close to one another (usually occurring from August 9th to August 15th). Therefore, I would say that, based on the additive approach, unofficial summer is not ending later either.

Appendix

Question 7.2.1

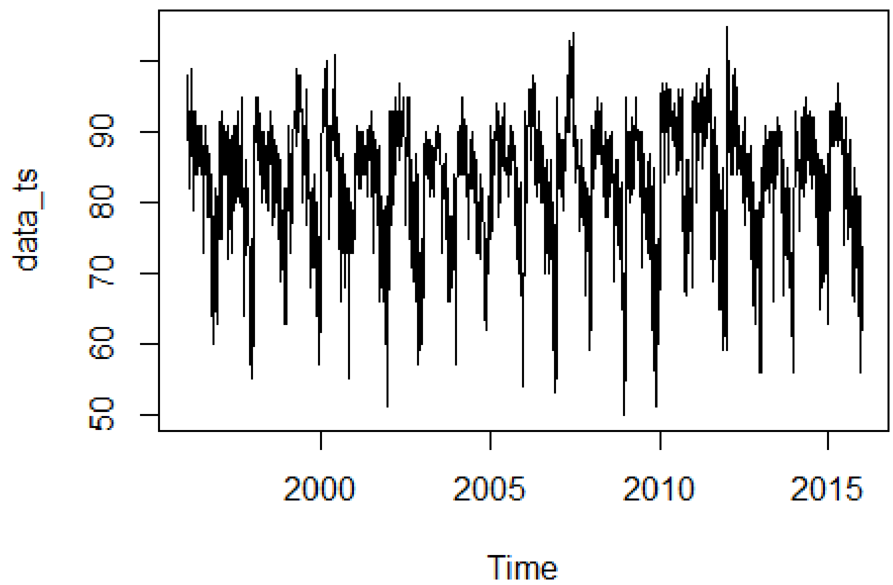
```
#Load library
library(stats);

#Load data
data <- read.table("C:\\Users\\User\\OneDrive\\Desktop\\Data 7.2\\temps.txt",
stringsAsFactors = F, header = T)

#Create time series
data_vector <- as.vector(unlist(data[,2:21]))
data_ts <- ts(data_vector, start=1996, frequency=123)
plot(data_ts)

#Additive
plot(decompose(data_ts))

#Multiplicative
plot(decompose(data_ts, type=c("multiplicative")))

#Holt-Winters Multiplicative
data_HW <- HoltWinters(data_ts, seasonal=c("multiplicative"))
plot(data_HW)

#Plot fitted values
plot(data_HW$fitted)

#CUSUM approach to detect change
#Seasonal approach
seasonal <- data_HW$fitted[,4]
seasonal <- matrix(seasonal, nrow=123)

colnames(seasonal) <- c(1997:2015)
rownames(seasonal) <- as.vector(data[,1])

dfcusum <- data.frame(matrix(nrow=nrow(seasonal), ncol=ncol(seasonal)))
colnames(dfcusum) <- colnames(seasonal)
rownames(dfcusum) <- rownames(seasonal)

for (year in 1:ncol(dfcusum)){
  dfcusum[1, year] <- 0
  mean <- mean(seasonal[1:62, year]) #July and August for summer mean
  stdev <- sd(seasonal[1:62, year])
  c_value <-  0.5 * stdev
  threshold <- 5 * stdev
  change_detect <- NULL

  for (row in 2:nrow(dfcusum)){
```

```
    dfcusum[row,year] <- max(0, dfcusum[row-1, year] +
                              (mean - seasonal[row, year] - c_value))
    if (dfcusum[row, year] >= threshold){
      change_detect <- append(change_detect, dfcusum[row, year])
      cat("The year is:", year, "and the date is:", row, "\n")
      break
    }
  }
}
```

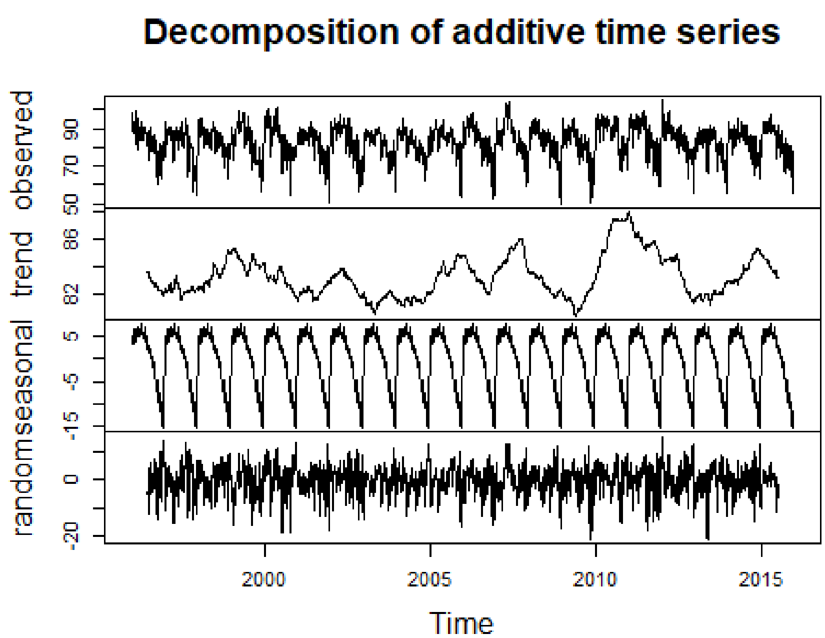
```
## The year is: 1 and the date is: 41
## The year is: 2 and the date is: 41
## The year is: 3 and the date is: 40
## The year is: 4 and the date is: 41
## The year is: 5 and the date is: 42
## The year is: 6 and the date is: 42
## The year is: 7 and the date is: 43
## The year is: 8 and the date is: 43
## The year is: 9 and the date is: 42
## The year is: 10 and the date is: 42
## The year is: 11 and the date is: 43
## The year is: 12 and the date is: 43
## The year is: 13 and the date is: 44
## The year is: 14 and the date is: 44
## The year is: 15 and the date is: 45
## The year is: 16 and the date is: 46
## The year is: 17 and the date is: 68
## The year is: 18 and the date is: 46
## The year is: 19 and the date is: 46
```

Question 7.2.2

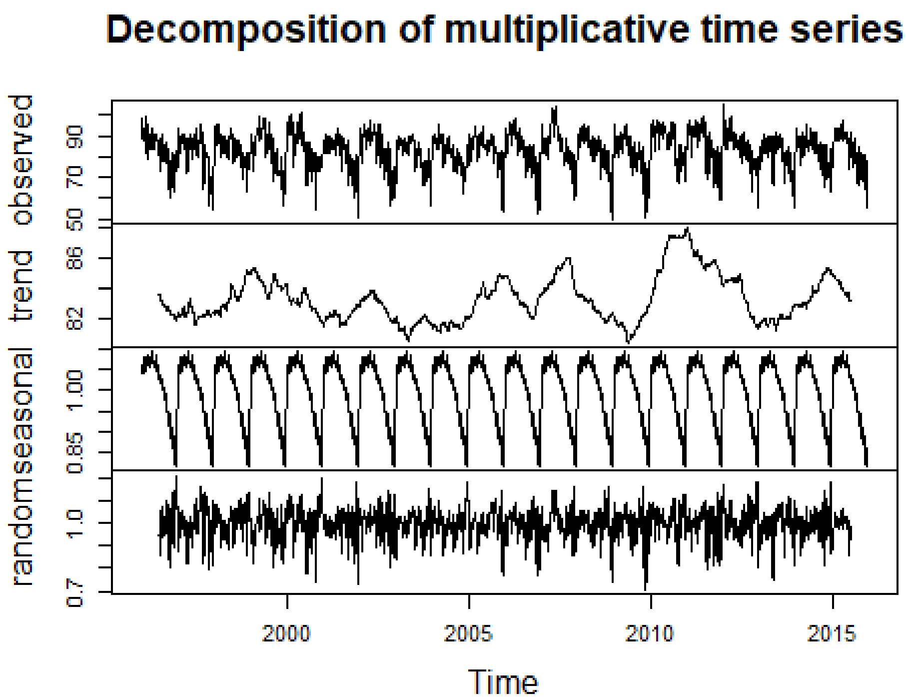
```
#Load library
library(stats);

#Load data
data <- read.table("C:\\Users\\User\\OneDrive\\Desktop\\Data 7.2\\temps.txt",
stringsAsFactors = F, header = T)

#Create time series
data_vector <- as.vector(unlist(data[,2:21]))
data_ts <- ts(data_vector, start=1996, frequency=123)
plot(data_ts)
```



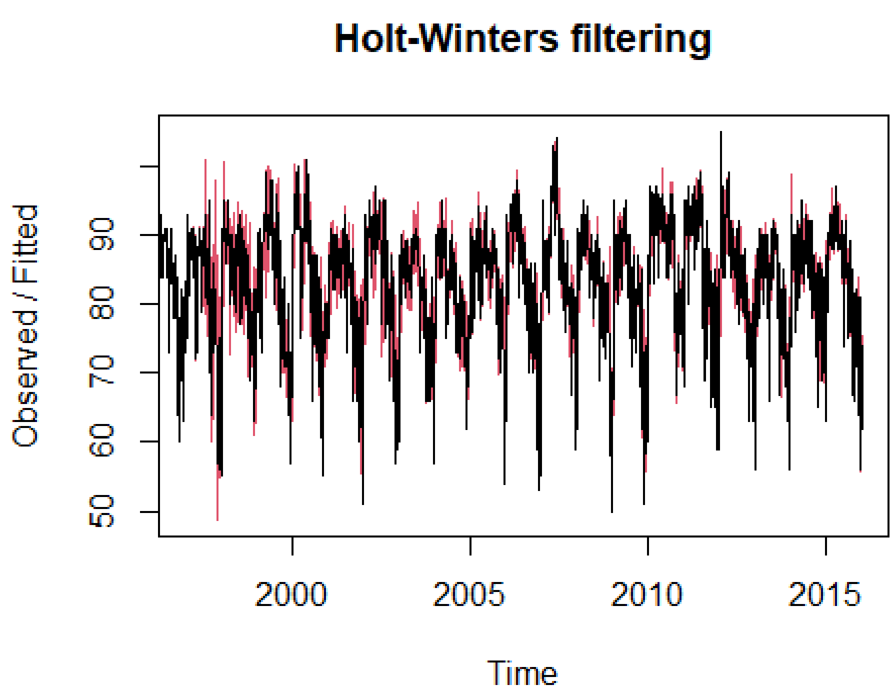

```r
#Additive
plot(decompose(data_ts))
```



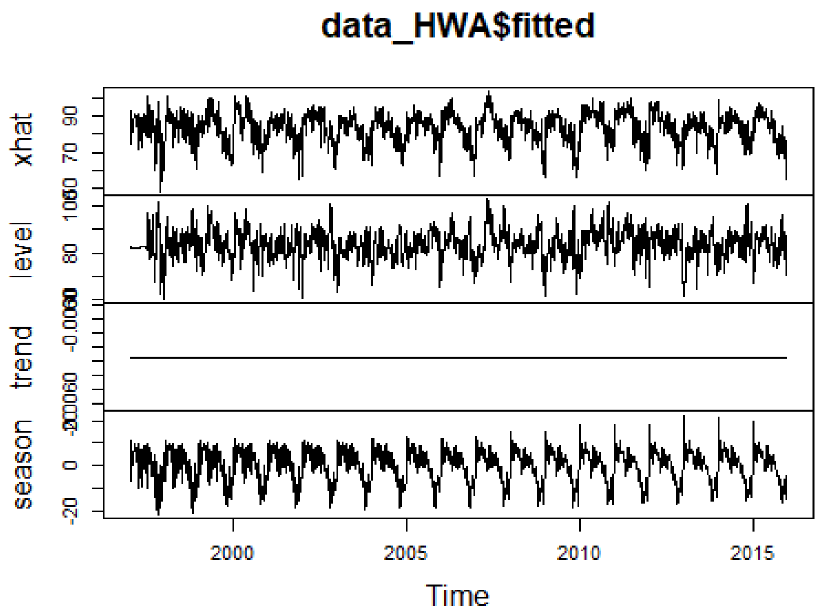

```r
#Multiplicative
plot(decompose(data_ts, type=c("multiplicative")))
```




```r
#Holt-Winters Additive
data_HWA <- HoltWinters(data_ts, seasonal=c("additive"))
plot(data_HWA)
```



```
#Plot fitted values
plot(data_HWA$fitted)
```




```
#CUSUM approach to detect change
#Seasonal approach
seasonal <- data_HWA$fitted[,4]
seasonal <- matrix(seasonal, nrow=123)

colnames(seasonal) <- c(1997:2015)
rownames(seasonal) <- as.vector(data[,1])

dfcusum <- data.frame(matrix(nrow=nrow(seasonal), ncol=ncol(seasonal)))
colnames(dfcusum) <- colnames(seasonal)
rownames(dfcusum) <- rownames(seasonal)

for (year in 1:ncol(dfcusum)){
  dfcusum[1, year] <- 0
  mean <- mean(seasonal[1:62, year]) #July and August for summer mean
  stdev <- sd(seasonal[1:62, year])
  c_value <-  0.5 * stdev
  threshold <- 5 * stdev
  change_detect <- NULL

  for (row in 2:nrow(dfcusum)){
    dfcusum[row,year] <- max(0, dfcusum[row-1, year] +
                               (mean - seasonal[row, year] - c_value))
    if (dfcusum[row, year] >= threshold){
      change_detect <- append(change_detect, dfcusum[row, year])
      cat("The year is:", year, "and the date is:", row, "\n")
      break
```

```
    }
  }
}
```

```
## The year is: 1 and the date is: 41
## The year is: 2 and the date is: 41
## The year is: 3 and the date is: 40
## The year is: 4 and the date is: 41
## The year is: 5 and the date is: 42
## The year is: 6 and the date is: 42
## The year is: 7 and the date is: 43
## The year is: 8 and the date is: 42
## The year is: 9 and the date is: 42
## The year is: 10 and the date is: 42
## The year is: 11 and the date is: 43
## The year is: 12 and the date is: 43
## The year is: 13 and the date is: 43
## The year is: 14 and the date is: 44
## The year is: 15 and the date is: 45
## The year is: 16 and the date is: 45
## The year is: 17 and the date is: 46
## The year is: 18 and the date is: 46
## The year is: 19 and the date is: 46
```