Project 4

Marcus McKenzie

4.1

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 (the first smoothing parameter) to be closer to 0 or 1, and why?

For work I work on robotics and autonomous vehicles. It would be informative to understand what the potential demand would be, by month, for autonomous vehicles in the future. To determine this information we could build an exponential smoothing model that looks at current transportation usage among ride sharing and personal vehicles throughout the year.

The alpha value for this model would depend on the time period that we are analyzing. For example, if we wanted to analyze the entire year, our alpha value might be closer to 0, since there may be random spikes that occur over this duration. Conversely, if we are analyzing shorter time periods, such as monthly or quarterly, we might increase the alpha value to be closer to 1, since greater changes in demand may be more reasonable for these shorter time periods.

4.2

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.

Load Data:

```r
rm(list = ls())

#data <- read.table("Documents/OMSCS/Analytics Modeling/Assignments/Assignment4/temps.txt", header= TRU

data <- read.table("temps.txt", header=TRUE, stringsAsFactors = FALSE)

head(data)
```

```
##     DAY X1996 X1997 X1998 X1999 X2000 X2001 X2002 X2003 X2004 X2005 X2006 X2007
## 1 1-Jul    98    86    91    84    89    84    90    73    82    91    93    95
## 2 2-Jul    97    90    88    82    91    87    90    81    81    89    93    85
## 3 3-Jul    97    93    91    87    93    87    87    87    86    86    93    82
## 4 4-Jul    90    91    91    88    95    84    89    86    88    86    91    86
## 5 5-Jul    89    84    91    90    96    86    93    80    90    89    90    88
## 6 6-Jul    93    84    89    91    96    87    93    84    90    82    81    87
##   X2008 X2009 X2010 X2011 X2012 X2013 X2014 X2015
## 1    85    95    87    92   105    82    90    85
## 2    87    90    84    94    93    85    93    87
## 3    91    89    83    95    99    76    87    79
```

```
## 4     90     91     85     92     98     77     84     85
## 5     88     80     88     90    100     83     86     84
## 6     82     87     89     90     98     83     87     84
```

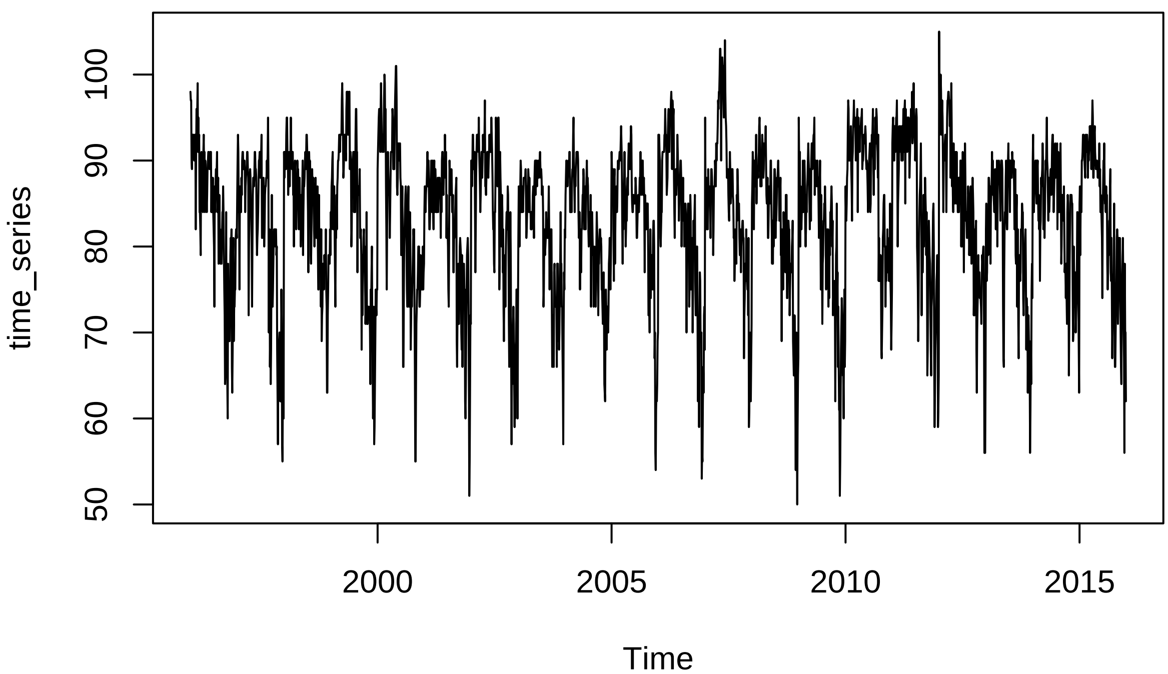
Create time series:

```r
time_series = as.vector(unlist(data[, 2:21]))

time_series = ts(time_series, start = 1996, frequency = 123)
```

Plot time series:

```r
plot(time_series)
```

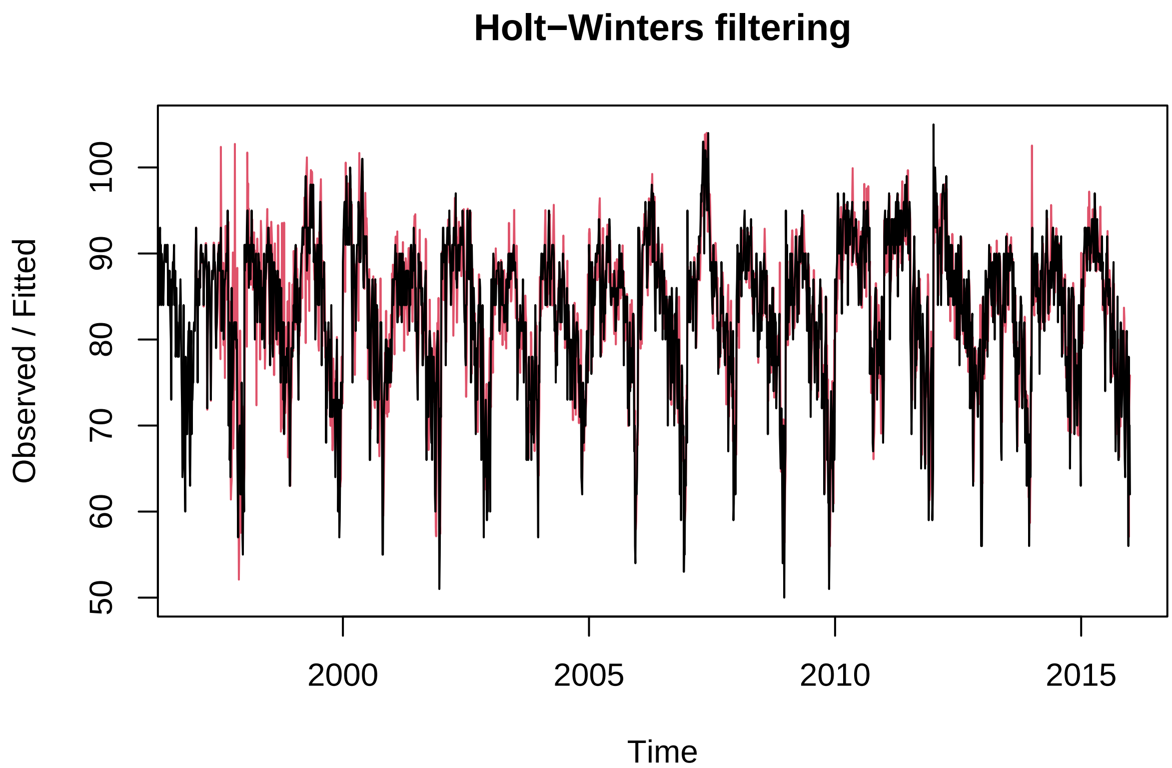


Implement HoltWinters:

```r
holtwinters <- HoltWinters(time_series, alpha=NULL, beta=NULL, gamma=NULL, seasonal = "multiplicative")
```

Plot time series:

```r
plot(holtwinters)
```



Create cusum function:

```
cusum = function(temps, avg, T, C){

  res = list()
  sum = 0
  count = 1

    while (count <= nrow(temps)){
      curr = temps[count,]

      #print("mean")
      #print(avg)
      #print(curr)
      sum = max(0, sum + (avg - curr - C))


      if (sum >= T) {
        res = count
        break
      }

      count = count + 1
      if (count >= nrow(temps)){
        res = NA
        break
```

```r
    }
  }

  return(res)
}
```

Running cusum:

```r
holtwinters <- matrix(holtwinters$fitted[,4], nrow=123)
colnames(holtwinters) <- colnames(data[,3:21])
#head(holtwinters)

c = sd(holtwinters[,1])*0.4
t = sd(holtwinters[,1])*2

res = vector()

for (col in 1:ncol(holtwinters)){
  res[col] = cusum(temps = as.matrix(holtwinters[,col]), avg = 1, T = t, C = c)
}
```

```r
out = data.frame(Year = colnames(holtwinters),Day = data[res,1])

out
```

```
##     Year    Day
## 1  X1997 30-Sep
## 2  X1998 30-Sep
## 3  X1999 30-Sep
## 4  X2000 30-Sep
## 5  X2001  1-Oct
## 6  X2002  1-Oct
## 7  X2003  1-Oct
## 8  X2004  1-Oct
## 9  X2005  2-Oct
## 10 X2006  2-Oct
## 11 X2007  2-Oct
## 12 X2008  2-Oct
## 13 X2009  2-Oct
## 14 X2010  2-Oct
## 15 X2011  1-Oct
## 16 X2012  1-Oct
## 17 X2013  1-Oct
## 18 X2014  1-Oct
## 19 X2015  2-Oct
```

Conclusion

From the results shown above we can see that the end of summer has gotten slightly later over the past 20 years. However, this change is not greatly significant, as the change only occurs over a few days. On the other hand, these results could differ based on the coefficients that are utilized in the function. Regardless, the end of summer changes should not be much more drastic than a few days.