**Question 6.2**

1. Using July through October daily-high-temperature data for Atlanta for 1996 through 2015, use a CUSUM approach to identify when unofficial summer ends (i.e., when the weather starts cooling off) each year. You can get the data that you need from the file temps.txt or online, for example at <http://www.iweathernet.com/atlanta-weather-records> or <https://www.wunderground.com/history/airport/KFTY/2015/7/1/CustomHistory.html> . You can use R if you’d like, but it’s straightforward enough that an Excel spreadsheet can easily do the job too.
2. Use a CUSUM approach to make a judgment of whether Atlanta’s summer climate has gotten warmer in that time (and if so, when).

**Answer1.**

The day a change in trend is detected is: 19-Oct

(Have also attached Excel Method)

**R – Method**

#Load dara from temperature table

#temp\_data = read\_tsv("temps.txt")

#checking for NA/missing values in data frame

sum(is.na(temp\_data))

#Exploring the data

tail(temp\_data)[1:5,]

#adding a column to data frame, with average value for each day

temps<-cbind(temp\_data,rowMeans(temp\_data[,-1]))

#renaming the column to Average

colnames(temps)[ncol(temp\_data)] <- "Average"

#checking the data frame

head(temp\_data)[1:3,]

#adding a col to temps data frame to store St value in each step

temp\_data[,"St"]<-NA

#Standard Deviation and Mean of the data points

print(std\_data<-sd(temp\_data[,"Average"]))

print(mean\_data<-mean(temp\_data[,"Average"]))

#allowable slack C : typically set to 1 sigma

C<-std\_data

#threshold set to 5\*sigma

print(T<-5\*std\_data)

#using Zero start method

temp\_data[1,"St"]<-0

for(i in 2:nrow(temp\_data)){

temp\_data[i,"St"]<-max(0,(temp\_data[i-1,"St"]+mean\_data-temp\_data[i,"Average"]-C))

}

temp\_data$St

cat("The day a change in trend is detected is:",temp\_data[which(temp\_data$St>T),"DAY"][1])

#The day a change in trend is detected is: 19-Oct

temp\_data[,"Date"]<-as.Date(temp\_data[,"DAY"],"%d-%B")

temp\_data[,"Date"]<-format(temp\_data[,"Date"],format="%m/%d")

options(repr.plot.width=10, repr.plot.height=5)

ggplot(data = temp\_data, aes(x = Date, y =`St`,group=2)) +

geom\_line()+

geom\_hline(yintercept=T,color="red")+

scale\_x\_discrete(breaks = unique(temp\_data$Date)[seq(1,123,5)])+

xlab("Days") +

ylab("CUSUM") +

ggtitle("CUSUM Chart for July-October Daily-high Temperature for Atlanta 1996-2015")+

theme(plot.title = element\_text(hjust = 0.5))