hw4_data_analytics

Question 7.1

Exponential Smoothing is useful to temper my mood. Each day brings up and downs if I look at the long term of my life it is pretty good understanding this helps me put problems in context and understand that most are not long in duration. For this I choose a lower value of alpha (closer to zero) indicating there is more randomness and I want to trust the more long term trend.

Question 7.2

Has the end of summer gotten later over the last twenty years.

Approach is to get the averages per day over the full complement of years. We will then look for the critical deviation from trend for the mean. Next we will take the fitted values from es against the daily temperature means and look for the same critical deviation. The results are mostly within a day or so; therefore, the es application does not make much difference.

```
library('smooth')
```

```
## Loading required package: greybox
```

```
## Registered S3 method overwritten by 'xts':
##   method     from
##   as.zoo.xts zoo
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
## Registered S3 methods overwritten by 'forecast':
##   method             from
##   fitted.fracdiff    fracdiff
##   residuals.fracdiff fracdiff
```

```
## Package "greybox", v0.5.4 loaded.
```

```
## This is package "smooth", v2.5.3
```

```
a=read.csv("temps.csv")
e=as.character(a$DAY)
e=cbind(e,numeric(123))
e=cbind(e,numeric(123))
#determine the average by day
for (i in 1:123)
  e[i,2]=mean(as.numeric(a[i,-1]))


#day when trend is broken with row means
y=as.numeric(e[,2])
mu=mean(y)
k=sd(y)
for(q in (1:10)) {
T=q*k;
S=0;count=0; for (i in y) {
```

```
  S=max(0,S+(mu-i-k));
  count=count+1 ;
  if (S > T) {
    print(paste0(q, " ", count));
    break;
  }
 }
}
```

```
## [1] "1 102"
## [1] "2 107"
## [1] "3 109"
## [1] "4 110"
## [1] "5 111"
## [1] "6 113"
## [1] "7 115"
## [1] "8 116"
## [1] "9 116"
## [1] "10 117"
```

```
#es the row means
zz=as.numeric(es(as.numeric(e[,2]))$fitted)
mu=mean(zz)
k=sd(zz)
for(q in (1:10)) {
T=q*k;
S=0;count=0;
for (i in zz) {
  S=max(0,S+(mu-i-k));
  count=count+1 ;
  if (S > T) {
    print(paste0(q, " ", count));
    break;
  }
 }
}
```

```
## [1] "1 103"
## [1] "2 107"
## [1] "3 109"
## [1] "4 111"
## [1] "5 112"
## [1] "6 113"
## [1] "7 115"
## [1] "8 116"
## [1] "9 117"
## [1] "10 118"
```

Side Question (not part of solution)

Using the ts function with decomposition you see a slight up trend to temperature around the 15th year but then it returning to the beginning trend in the 20th year.

```r
a=read.csv("temps.csv")
b=cbind(a$X1996)
for (i in 1997:2015)
b=cbind(b,a[[paste0('X',i)]])
b=as.vector(b)
my_ts=ts(b,frequency=122)
plot(decompose(my_ts))
```