Test Result 20170313

Jiayu (Alice) Wu

2017-03-09

Contents

Vector Computation

- Test: write a function findNum(y,x,l) to compute the position in vetor y where number x repeats l times.

Method 1

Clever solution (20min)

```
findNum<-function(y,x,l) {
       a<-which(y==x)
       b<-which(diff(y,1,l-1)==0)
       return(a[a%in%b])
       }
findNum(y,1,2)
findNum(y,1,2)
```

Method 2

Violent loop solution (7min)

```
findNum<-function(y,x,l) {
       result<-NULL
       for (i in 1:(length(y)-l+1)) {
              if (all(y[i:(i+l-1)]==x)) {
                     result<-c(result,i)
              }
       }
       print(result)
}
findNum(y,1,2)
```

Data Structure

- Test: compute the gap between daily maximum and minimum temperature (6min)

```
raw <- read.delim("data/weather.txt",check.names = F, na.strings = ".")
library(reshape2)
data<-melt(raw,id=c("year", "month", "element"),
           variable.name = "day", na.rm = TRUE)
df<-dcast(data,year+month+day~element, value.var = "value")
day<-as.Date(paste0(df$year,"-",df$month,"-",df$day))
result<-data.frame(tdiff = df$tmax-df$tmin, row.names = day)
```

Group Processing

- Test: Compute the mean of 10th quantile of ArrDelay for each Unique Carrier each month (10min)

```
#   hfflights
library(hflights)
str(hflights)
library(hflights)
str(hflights)
table(hflights$Year)
data<-hflights[,c("UniqueCarrier","Month", "ArrDelay")]
hfflights<-tapply(data$ArrDelay,list(data$UniqueCarrier, data$Month),
                  quantile,probs = 0.1,na.rm = T)
```