

Loops_and_Control_Flow.R

015004

Thu Feb 21 13:28:18 2019

```
#####
# Load packages #
#####

#install.packages("JM")
library(JM)

## Warning: package 'JM' was built under R version 3.5.2
## Loading required package: MASS
## Loading required package: nlme
## Loading required package: splines
## Loading required package: survival
## Warning: package 'survival' was built under R version 3.5.2
#####
# Loops and Control Flow #
#####

## Calculate the sum per row
A <- matrix(rnorm(1e06), 1000, 1000)
res1 <- replicate(100, z <- apply(A, 1, sum))
res2 <- replicate(100, z <- rowSums(A)) # Use specialized functions

## compute the cumulative sum
x <- rnorm(1000000, 10, 10)

cSum <- 0
for (i in 1:length(x)){
  cSum <- cSum + x[i]
  cSum
}

tail(cumsum(x))

## [1] 10015816 10015819 10015834 10015842 10015854 10015864

system.time({
  cSum <- 0
  for (i in 1:length(x)){
    cSum <- cSum + x[i]
    cSum
  }
})

##      user  system elapsed
##    0.07    0.00    0.07
```

```

system.time({
  cumsum(x)
}) # better

##      user      system elapsed
##         0         0         0

## Create a dichotomous variable for age
for (i in 1:dim(pbc2.id)[1]) {
  pbc2.id$ageCat[i] <- as.numeric(pbc2.id$age[i] > 42)
}

pbc2.id$ageCat <- as.numeric(pbc2.id$age > 42)

## calculate the mean weight of males and females in 100 datasets
datlist <- list()

i <- 1
set.seed(2015+i)
patient <- c(1:20)
weight <- rnorm(20, 70, 10)
sex <- sample(1:2, 20, replace = T)
sex <- factor(sex, levels = 1:2, labels = c("male", "female"))

datlist[[i]] <- data.frame(patient, weight, sex)

i <- 2
set.seed(2015+i)
patient <- c(1:20)
weight <- rnorm(20, 70, 10)
sex <- sample(1:2, 20, replace = T)
sex <- factor(sex, levels = 1:2, labels = c("male", "female"))

datlist[[i]] <- data.frame(patient, weight, sex)

for (i in 1:100) {
  set.seed(2015+i)
  patient <- c(1:20)
  weight <- rnorm(20, 70, 10)
  sex <- sample(1:2, 20, replace = T)
  sex <- factor(sex, levels = 1:2, labels = c("male", "female"))

  datlist[[i]] <- data.frame(patient, weight, sex)
}

means <- matrix(NA, length(datlist), 2)

i <- 1
dat <- datlist[[i]]
means[i, ] <- tapply(dat$weight, dat$sex, mean)

i <- 2
dat <- datlist[[i]]

```

```
means[i, ] <- tapply(dat$weight, dat$sex, mean)

for (i in 1:length(datlist)) {
  dat <- datlist[[i]]
  means[i, ] <- tapply(dat$weight, dat$sex, mean)
}

means
```

```
##           [,1]      [,2]
## [1,] 67.43233 71.23607
## [2,] 64.97965 72.20599
## [3,] 65.24479 71.31810
## [4,] 66.99264 67.66554
## [5,] 73.48859 65.96814
## [6,] 73.75327 74.12918
## [7,] 67.86328 66.57248
## [8,] 67.76939 72.09949
## [9,] 67.13370 66.89012
## [10,] 69.87276 73.52814
## [11,] 66.45083 65.74966
## [12,] 70.91945 68.87740
## [13,] 59.08741 70.32179
## [14,] 73.79177 68.27906
## [15,] 70.61414 69.37221
## [16,] 72.95609 69.47359
## [17,] 78.27397 70.28429
## [18,] 65.46974 69.62863
## [19,] 72.05220 72.07861
## [20,] 70.92732 67.30444
## [21,] 63.85819 72.52005
## [22,] 76.03918 63.37934
## [23,] 70.96035 77.16855
## [24,] 65.92766 63.08480
## [25,] 73.75418 73.49044
## [26,] 70.46921 67.87070
## [27,] 72.38152 68.36856
## [28,] 74.23517 72.78862
## [29,] 71.59158 71.38895
## [30,] 71.95000 72.94791
## [31,] 72.63845 72.96826
## [32,] 70.11789 73.66195
## [33,] 66.56750 70.89077
## [34,] 66.04598 66.45523
## [35,] 69.35357 65.89027
## [36,] 66.52157 66.06234
## [37,] 58.45010 66.77747
## [38,] 76.01192 73.10698
## [39,] 71.12626 72.75966
## [40,] 72.53766 69.66504
## [41,] 68.27829 69.47538
## [42,] 68.28131 70.94434
## [43,] 65.14775 68.59821
```

```

## [44,] 73.35305 72.78260
## [45,] 72.28328 68.16351
## [46,] 71.14535 68.93143
## [47,] 64.91764 61.73179
## [48,] 64.62555 68.84194
## [49,] 72.26236 69.87325
## [50,] 65.75143 68.21834
## [51,] 69.52712 73.08446
## [52,] 65.23999 68.01814
## [53,] 69.90426 70.58894
## [54,] 67.83676 68.22100
## [55,] 68.56953 70.36872
## [56,] 72.27409 73.10898
## [57,] 73.36143 65.15850
## [58,] 71.39609 72.74777
## [59,] 72.79471 72.72307
## [60,] 69.96473 78.02253
## [61,] 67.63354 76.91379
## [62,] 67.26969 68.98290
## [63,] 71.51521 74.90335
## [64,] 69.83345 69.57401
## [65,] 70.66560 69.79079
## [66,] 68.08066 65.94654
## [67,] 67.47622 74.84203
## [68,] 72.75592 73.99124
## [69,] 71.91537 63.25603
## [70,] 69.35163 68.81534
## [71,] 75.28391 73.14252
## [72,] 75.33244 75.67009
## [73,] 70.28109 74.13563
## [74,] 69.25352 70.99000
## [75,] 72.99973 71.23398
## [76,] 75.90840 61.75362
## [77,] 66.19016 70.13656
## [78,] 75.51152 73.18479
## [79,] 69.06143 73.74246
## [80,] 69.87969 69.90532
## [81,] 69.42643 69.60038
## [82,] 72.04434 68.57077
## [83,] 66.57876 70.70148
## [84,] 69.82176 64.19132
## [85,] 65.69448 72.20433
## [86,] 75.20922 74.89963
## [87,] 68.44756 69.50870
## [88,] 72.33716 67.18127
## [89,] 73.05528 71.03692
## [90,] 67.16111 74.34021
## [91,] 69.33992 73.66316
## [92,] 69.63129 71.32907
## [93,] 68.25694 69.65056
## [94,] 71.67755 63.85267
## [95,] 71.19492 74.48421
## [96,] 70.85746 65.13255
## [97,] 74.51213 71.15306

```

```
## [98,] 71.61401 74.26523
## [99,] 67.79296 69.78505
## [100,] 72.49128 70.07939

## select datasets where more than 39% of the patients are females
newList <- list()
k <- 1
for (i in 1:length(datlist)) {
  dat <- datlist[[i]]
  if (sum(dat$sex == "female")/20 >= 0.4) {
    newList[[k]] <- dat
    k <- k + 1
  }
}

length(newList)
```

```
## [1] 85

## select datasets where more than 49% of the patients are males
newList <- list()
k <- 1
for (i in 1:length(datlist)) {
  dat <- datlist[[i]]
  if (sum(dat$sex == "male")/20 >= 0.5) {
    newList[[k]] <- dat
    k <- k + 1
  }
}
```

```
## other examples
for (i in 1:10){
  if (i < 5) {
    print(i)
  }
}
```

```
## [1] 1
## [1] 2
## [1] 3
## [1] 4

for (i in 1:10){
  if (i < 5) {
    print(2*i)
  } else {
    print(i)
  }
}
```

```
## [1] 2
## [1] 4
## [1] 6
## [1] 8
## [1] 5
```

```
## [1] 6
## [1] 7
## [1] 8
## [1] 9
## [1] 10
```