HW_02

Oksana

05 04 2019

```
library(boot)
```

```
## Warning: package 'boot' was built under R version 3.5.2
```

```
data(melanoma)
head(melanoma)
```

```
##   time status sex age year thickness ulcer
## 1   10      3   1  76 1972      6.76     1
## 2   30      3   1  56 1968      0.65     0
## 3   35      2   1  41 1977      1.34     0
## 4   99      3   0  71 1968      2.90     0
## 5  185      1   1  52 1965     12.08     1
## 6  204      1   1  28 1971      4.84     1
```

```
titanic <- read.csv("titanic.csv", sep = ';')
head(titanic)
```

```
##   pclass survived gender     age
## 1      1      yes      F 29.0000
## 2      1      yes      M  0.9167
## 3      1       no      F  2.0000
## 4      1       no      M 30.0000
## 5      1       no      F 25.0000
## 6      1      yes      M 48.0000
```

```
my_data_function <- function(your_data, rw = 0, clmn = 0) {
  if (rw == 0 && clmn == 0) {
    res <- list(subs = as.data.frame(your_data))
  } else if (rw == 0) {
    res <- list(subs = as.data.frame(your_data[, clmn]))
  } else if (clmn == 0) {
    res <- list(subs = as.data.frame(your_data[rw,]))
  } else {
    res <- list(subs = as.data.frame(your_data[rw,clmn]))
  }

  calc <- calculation(res)
  res <- c(res, calculus = calc)
  res
}

calculation <- function(sub_data){
  calc = list()

  for (col in colnames(sub_data$subs)) {
    if (is.numeric(sub_data$subs[[col]])) {
      tmp <- mean(sub_data$subs[[col]])
      names(tmp) <- col
```

```
      calc <- c(calc, tmp)
    } else {
      tmp <- summary(sub_data$subs[[col]])
      names_tmp <- paste0(col, names(tmp))
      names(tmp) <- names_tmp
      calc <- c(calc, tmp)
    }
  }
  return(calc)
}
```

```
#Examples of subsetting and computing

###1
a <- my_data_function(melanoma, rw = 1:10, clmn = c(1,3,6))
```

```
## Warning in if (rw == 0) {: the condition has length > 1 and only the first
## element will be used
```

```
## Warning in if (clmn == 0) {: the condition has length > 1 and only the
## first element will be used
```

```
head(a$subs)
```

```
##   time sex thickness
## 1   10   1      6.76
## 2   30   1      0.65
## 3   35   1      1.34
## 4   99   0      2.90
## 5  185   1     12.08
## 6  204   1      4.84
```

```
a[-1]
```

```
## $calculus.time
## [1] 151.6
##
## $calculus.sex
## [1] 0.7
##
## $calculus.thickness
## [1] 5.724
```

```
###2
a <- my_data_function(melanoma, clmn = c(1,3,6))
head(a$subs)
```

```
##   time sex thickness
## 1   10   1      6.76
## 2   30   1      0.65
## 3   35   1      1.34
## 4   99   0      2.90
## 5  185   1     12.08
## 6  204   1      4.84
```

```
a[-1]
```

```
## $calculus.time
```

```
## [1] 2152.8
##
## $calculus.sex
## [1] 0.3853659
##
## $calculus.thickness
## [1] 2.919854
```

```
###3
a <- my_data_function(melanoma, c(1,2,3))
```

```
## Warning in if (rw == 0) {: the condition has length > 1 and only the first
## element will be used
```

```
head(a)
```

```
## $subs
##   time status sex age year thickness ulcer
## 1   10      3   1  76 1972      6.76     1
## 2   30      3   1  56 1968      0.65     0
## 3   35      2   1  41 1977      1.34     0
##
## $calculus.time
## [1] 25
##
## $calculus.status
## [1] 2.666667
##
## $calculus.sex
## [1] 1
##
## $calculus.age
## [1] 57.66667
##
## $calculus.year
## [1] 1972.333
```

```
###4
a <- my_data_function(melanoma, clmn  = c(T,F, F, F, T, F, F))
head(a$subs)
```

```
##   time year
## 1   10 1972
## 2   30 1968
## 3   35 1977
## 4   99 1968
## 5  185 1965
## 6  204 1971
```

```
a[-1]
```

```
## $calculus.time
## [1] 2152.8
##
## $calculus.year
## [1] 1969.907
```

```
###5
a <- my_data_function(melanoma, clmn = c("time", "age"))
head(a$subs)
```

```
##   time age
## 1   10  76
## 2   30  56
## 3   35  41
## 4   99  71
## 5  185  52
## 6  204  28
```

```
a[-1]
```

```
## $calculus.time
## [1] 2152.8
##
## $calculus.age
## [1] 52.46341
```

```
###6
a <- my_data_function(melanoma, rw = 3:6)
```

```
## Warning in if (rw == 0) {: the condition has length > 1 and only the first
## element will be used
```

```
head(a$subs)
```

```
##   time status sex age year thickness ulcer
## 3   35      2   1  41 1977      1.34     0
## 4   99      3   0  71 1968      2.90     0
## 5  185      1   1  52 1965     12.08     1
## 6  204      1   1  28 1971      4.84     1
```

```
a[-1]
```

```
## $calculus.time
## [1] 130.75
##
## $calculus.status
## [1] 1.75
##
## $calculus.sex
## [1] 0.75
##
## $calculus.age
## [1] 48
##
## $calculus.year
## [1] 1970.25
##
## $calculus.thickness
## [1] 5.29
##
## $calculus.ulcer
## [1] 0.5
```

```
###7
a <- my_data_function(melanoma, rw = c(3,4,5,6))
```

```
## Warning in if (rw == 0) {: the condition has length > 1 and only the first
## element will be used
```

```
head(a$subs)
```

```
##   time status sex age year thickness ulcer
## 3   35      2   1  41 1977      1.34     0
## 4   99      3   0  71 1968      2.90     0
## 5  185      1   1  52 1965     12.08     1
## 6  204      1   1  28 1971      4.84     1
```

```
a[-1]
```

```
## $calculus.time
## [1] 130.75
##
## $calculus.status
## [1] 1.75
##
## $calculus.sex
## [1] 0.75
##
## $calculus.age
## [1] 48
##
## $calculus.year
## [1] 1970.25
##
## $calculus.thickness
## [1] 5.29
##
## $calculus.ulcer
## [1] 0.5
```

```
###8
a <- my_data_function(melanoma)
head(a$subs)
```

```
##   time status sex age year thickness ulcer
## 1   10      3   1  76 1972      6.76     1
## 2   30      3   1  56 1968      0.65     0
## 3   35      2   1  41 1977      1.34     0
## 4   99      3   0  71 1968      2.90     0
## 5  185      1   1  52 1965     12.08     1
## 6  204      1   1  28 1971      4.84     1
```

```
a[-1]
```

```
## $calculus.time
## [1] 2152.8
##
## $calculus.status
## [1] 1.790244
##
```

```
## $calculus.sex
## [1] 0.3853659
##
## $calculus.age
## [1] 52.46341
##
## $calculus.year
## [1] 1969.907
##
## $calculus.thickness
## [1] 2.919854
##
## $calculus.ulcer
## [1] 0.4390244
```

```
###9
a <- my_data_function(titanic, rw = 1:1000)
```

```
## Warning in if (rw == 0) {: the condition has length > 1 and only the first
## element will be used
```

```
head(a$subs)
```

```
##   pclass survived gender     age
## 1      1      yes      F 29.0000
## 2      1      yes      M  0.9167
## 3      1       no      F  2.0000
## 4      1       no      M 30.0000
## 5      1       no      F 25.0000
## 6      1      yes      M 48.0000
```

```
a[-1]
```

```
## $calculus.pclass
## [1] 2.171
##
## $calculus.survivedno
## [1] 584
##
## $calculus.survivedyes
## [1] 416
##
## $calculus.genderF
## [1] 374
##
## $calculus.genderM
## [1] 626
##
## $calculus.age
## [1] 29.98167
```