HM03_final

```r
library(tidyr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(data.table)
```

```
##
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:dplyr':
##
##     between, first, last
```

```r
data("iris")
head(iris)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1          5.1         3.5          1.4         0.2  setosa
## 2          4.9         3.0          1.4         0.2  setosa
## 3          4.7         3.2          1.3         0.2  setosa
## 4          4.6         3.1          1.5         0.2  setosa
## 5          5.0         3.6          1.4         0.2  setosa
## 6          5.4         3.9          1.7         0.4  setosa
```

```r
#The which() function will return the position of the
elements(i.e., row number/column number/array index) in a logical
vector which are TRUE.
split_data <-  function(data) {
  numeric_index <- sapply(data, is.numeric) | sapply(data,
is.integer) # gives index of numeric values
  numeric_data <- data[,which(numeric_index), drop=F] # selects
numeric data
  non_numeric_data <- data[,which(!numeric_index), drop = F] #
selects non numeric data
  list(numeric_data=numeric_data,
non_numeric_data=non_numeric_data) # returns both data as list
}
```

```r
calc <- function(data, row_s, col_s) {
  data <- data[row_s, col_s]
  s_data <- split_data(data)
  means <- colMeans(s_data$numeric_data)  #the same apply(x, 2,
mean)
  sums <- colSums(s_data$numeric_data) #the same as apply(x, 2,
sum)
  f_table <- lapply(s_data$non_numeric_data, table)  #use lapply
because i want list
  list(mean=means, sum=sums, freq_table=f_table)
}
calc(iris, row_s = 1:nrow(iris), col_s = 1:ncol(iris))

## $mean
## Sepal.Length  Sepal.Width Petal.Length  Petal.Width
##     5.843333     3.057333     3.758000     1.199333
##
## $sum
## Sepal.Length  Sepal.Width Petal.Length  Petal.Width
##        876.5        458.6        563.7        179.9
##
## $freq_table
## $freq_table$Species
##
##     setosa versicolor  virginica
##         50         50         50

data <-
read.csv('/home/sedreh/ITMO/semester2/Statistic-R/2/bosson.csv',
sep = ";")
#head(data)
calc(data, row_s = 1:nrow(data), col_s = 1:ncol(data))

## $mean
## aneurysm      bmi     risk
## 47.56938 22.76071  1.38756
##
## $sum
## aneurysm      bmi     risk
## 9942.000 4756.988  290.000
##
## $freq_table
## $freq_table$country
##
##  France Vietnam
##      99     110
##
## $freq_table$gender
##
##   F   M
##  51 158
```