

# R Notebook

Write a function that:

- 1 . Takes data frame, row selection and column selection as arguments
2. Subsets data according to the given selections. Possible selections are:

By index/indexes

By name/names

By logical vector

3. For each column of the subsetted data performs some computations:

For numeric data – any of `sum()`, `mean()`, `sd()`, etc. As you wish

For non-numeric data (characters and logical vectors, factors) – frequency table

4. Returns an output as a data frame and calculation results as a list

```
subset_df <- function(df, row_selection, column_selection){  
  
  column_func <- function(x){  
    if (is.numeric(x)){  
      return(mean(x))  
    } else {  
      return(table(x))  
    }  
  }  
  
  df_subs <- df[row_selection, column_selection, drop=F]  
  d <- sapply(df_subs, column_func)  
  
  return (list(df_subs, d))  
}
```

Apply it in a few examples:

```
library("datasets")  
df <- mtcars  
l <- subset_df(df, c(1,2,3), c(T, F))  
l[[1]]
```

```
##           mpg disp drat  qsec am carb  
## Mazda RX4    21.0  160 3.90 16.46  1   4  
## Mazda RX4 Wag 21.0  160 3.90 17.02  1   4  
## Datsun 710    22.8  108 3.85 18.61  1   1
```

```
l[[2]]
```

```
##      mpg      disp      drat      qsec      am      carb
## 21.600000 142.666667  3.883333 17.363333  1.000000  3.000000
```

```
library("ggplot2")
df <- presidential
l <- subset_df(df, 1:10, c("name", "start", "end"))
l[[1]]
```

```
## # A tibble: 10 x 3
##   name      start      end
##   <chr>    <date>    <date>
## 1 Eisenhower 1953-01-20 1961-01-20
## 2 Kennedy    1961-01-20 1963-11-22
## 3 Johnson    1963-11-22 1969-01-20
## 4 Nixon      1969-01-20 1974-08-09
## 5 Ford       1974-08-09 1977-01-20
## 6 Carter     1977-01-20 1981-01-20
## 7 Reagan     1981-01-20 1989-01-20
## 8 Bush       1989-01-20 1993-01-20
## 9 Clinton    1993-01-20 2001-01-20
## 10 Bush      2001-01-20 2009-01-20
```

```
l[[2]]
```

```
## $name
## x
##      Bush      Carter      Clinton Eisenhower      Ford      Johnson
##      2          1          1          1          1          1
## Kennedy      Nixon      Reagan
##      1          1          1
##
## $start
## x
## 1953-01-20 1961-01-20 1963-11-22 1969-01-20 1974-08-09 1977-01-20
##      1          1          1          1          1          1
## 1981-01-20 1989-01-20 1993-01-20 2001-01-20
##      1          1          1          1
##
## $end
## x
## 1961-01-20 1963-11-22 1969-01-20 1974-08-09 1977-01-20 1981-01-20
##      1          1          1          1          1          1
## 1989-01-20 1993-01-20 2001-01-20 2009-01-20
##      1          1          1          1
```