R Data Frames Exercises

Author: Suchanya Namwong

For this exercise we will test the knowledge of data frames!

Recreate the following dataframe by creating vectors and using the `data.frame` function:

```
name <- c('Sam', 'Frank', 'Amy')
age  <- c(22, 25, 26)
weight <- c(150, 165, 120)
sex  <- c('M', 'M', 'F')
personal <- data.frame(row.names = name, age, weight, sex)
personal
```

```
##       age weight sex
## Sam    22    150   M
## Frank  25    165   M
## Amy    26    120   F
```

Check if `mtcars` is a dataframe using `is.data.frame()`

```
is.data.frame(mtcars)
```

```
## [1] TRUE
```

Use `as.data.frame()` to convert a matrix into a dataframe:

```
mat <- matrix(1:25,nrow = 5)
mat
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    1    6   11   16   21
## [2,]    2    7   12   17   22
## [3,]    3    8   13   18   23
## [4,]    4    9   14   19   24
## [5,]    5   10   15   20   25
```

```
as.data.frame(mat)
```

```
##   V1 V2 V3 V4 V5
## 1  1  6 11 16 21
## 2  2  7 12 17 22
## 3  3  8 13 18 23
## 4  4  9 14 19 24
## 5  5 10 15 20 25
```

Set the built-in data frame `mtcars` as a variable `df`. We'll use this `df` variable for the rest of the exercises.

```
df <- mtcars
```

Display the first 6 rows of `df`

```
head(df)
```

```
##                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

What is the average `mpg` value for all the cars?

```
mean(df$mpg)
```

```
## [1] 20.09062
```

Select the rows where all cars have 6 cylinders (`cyl` column)

```
df[df$cyl == 6,]
```

```
##                 mpg cyl  disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4      21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag  21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
## Hornet 4 Drive 21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
## Valiant        18.1   6 225.0 105 2.76 3.460 20.22  1  0    3    1
## Merc 280       19.2   6 167.6 123 3.92 3.440 18.30  1  0    4    4
## Merc 280C      17.8   6 167.6 123 3.92 3.440 18.90  1  0    4    4
## Ferrari Dino   19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
```

Select the columns `am`,`gear`, and `carb`.

```
df[c('am', 'gear', 'carb')]
```

```
##                   am gear carb
## Mazda RX4          1    4    4
## Mazda RX4 Wag      1    4    4
## Datsun 710         1    4    1
## Hornet 4 Drive     0    3    1
## Hornet Sportabout  0    3    2
```

```
## Valiant             0    3    1
## Duster 360          0    3    4
## Merc 240D           0    4    2
## Merc 230            0    4    2
## Merc 280            0    4    4
## Merc 280C           0    4    4
## Merc 450SE          0    3    3
## Merc 450SL          0    3    3
## Merc 450SLC         0    3    3
## Cadillac Fleetwood  0    3    4
## Lincoln Continental 0    3    4
## Chrysler Imperial   0    3    4
## Fiat 128            1    4    1
## Honda Civic         1    4    2
## Toyota Corolla      1    4    1
## Toyota Corona       0    3    1
## Dodge Challenger    0    3    2
## AMC Javelin         0    3    2
## Camaro Z28          0    3    4
## Pontiac Firebird    0    3    2
## Fiat X1-9           1    4    1
## Porsche 914-2       1    5    2
## Lotus Europa        1    5    2
## Ford Pantera L      1    5    4
## Ferrari Dino        1    5    6
## Maserati Bora       1    5    8
## Volvo 142E          1    4    2
```

Create a new column called `performance`, which is calculated by `hp/wt`.

```
df$performance <- df$hp/df$wt
head(df)
```

```
##                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
##                   performance
## Mazda RX4            41.98473
## Mazda RX4 Wag        38.26087
## Datsun 710           40.08621
## Hornet 4 Drive       34.21462
## Hornet Sportabout    50.87209
## Valiant              30.34682
```

Your performance column will have several decimal place precision. Figure out how to use `round()` (check help(round)) to reduce this accuracy to only 2 decimal places.

```
df$performance <- round(df$performance, 2)
head(df)
```

```
##                    mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
##                   performance
## Mazda RX4               41.98
## Mazda RX4 Wag           38.26
## Datsun 710              40.09
## Hornet 4 Drive          34.21
## Hornet Sportabout       50.87
## Valiant                 30.35
```

What is the average `mpg` for cars that have more than 100 `hp` AND a `wt` value of more than 2.5.

```
df.ht.wt <- df[df$hp > 100 & df$wt > 2.5,]
mean(df.ht.wt$mpg)
```

```
## [1] 16.86364
```

What is the mpg of the `Hornet Sportabout`?

```
df[['Hornet Sportabout', 'mpg']]
```

```
## [1] 18.7
```