R data frames exercises

Alex Monteil

2025-01-15

1) Recreate the dataframe by creating vectors first

```
age <- c(22, 25, 26)
weight <- c(150, 165, 120)
sex <- c("M", "M", "F")
df <- data.frame(age, weight, sex)
rownames(df) <- c("Sam", "Frank", "Amy")
print(df)
```

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

2) Check is mtcars is a data frame

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

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

3) Convert the matrix into a data frame

```
mat <- matrix(1:25, nrow = 5)
df2 <- as.data.frame(mat)
print(df2)
```

```
##   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
```

4) Set the mtcars built in data frame in a df_cars variable

```r
df_cars <- mtcars
```

5) Display the first 6 rows of df_cars

```r
head(df_cars)
```

```
##                   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
```

6) What is the average mpg value for all cars ?

```r
mean(df_cars$mpg)
```

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

7) Select cars that have 6 cylinders

```r
subset(df_cars, subset = 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
```

8) Select columns am, gear and carb

```r
df_cars[,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
```

9) Create a new column called performance which is calculated as hp divided by wt

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

```
##                   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
```

10) Your performance column will have several decimal precision. Figure out how to use round to reduce it to 2 decimal places

```
df_cars$performance <- round(df_cars$performance, digits = 2)
head(df_cars)
```

```
##                    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
```

11) What is the mpg of the Hornet Sportabout ?

```
df_cars["Hornet Sportabout", "mpg"]
```

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