

# Homework 1: `mtcars` data exploration

*Rachel Han*

*15/09/2019*

## Contents

<code>mtcars</code> as tibble	1
Dimensions of data	1
Glimpse	2
Row names	2
Summary	2

## `mtcars` as tibble

`mtcars` data set is of type `data.frame`, not a tibble. A tibble lets us see the data without printing the entire data set to the console. An example is shown below:

```
mtcars_tibble <- as_tibble(mtcars)
mtcars_tibble
```

```
## # A tibble: 32 x 11
##   mpg   cyl  disp    hp  drat    wt  qsec    vs  am  gear  carb
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1  21     6   160   110   3.9   2.62  16.5     0     1     4     4
## 2  21     6   160   110   3.9   2.88  17.0     0     1     4     4
## 3  22.8   4   108    93   3.85   2.32  18.6     1     1     4     1
## 4  21.4   6   258   110   3.08   3.22  19.4     1     0     3     1
## 5  18.7   8   360   175   3.15   3.44  17.0     0     0     3     2
## 6  18.1   6   225   105   2.76   3.46  20.2     1     0     3     1
## 7  14.3   8   360   245   3.21   3.57  15.8     0     0     3     4
## 8  24.4   4   147.    62   3.69   3.19   20      1     0     4     2
## 9  22.8   4   141.    95   3.92   3.15  22.9     1     0     4     2
## 10 19.2   6   168.   123   3.92   3.44  18.3     1     0     4     4
## # ... with 22 more rows
```

## Dimensions of data

let's see how many rows and columns are there:

```
rownum <- nrow(mtcars_tibble)
colnum <- ncol(mtcars_tibble)
cat("row:",rownum)
```

```
## row: 32
```

```
cat("column:",colnum)
```

```
## column: 11
```

## Glimpse

We can use the `glimpse` function to see the glimpse of the data. Notice that the data is transposed upon applying this function: the column names are now in rows.

```
glimpse(mtcars_tibble)
```

```
## Observations: 32
## Variables: 11
## $ mpg <dbl> 21.0, 21.0, 22.8, 21.4, 18.7, 18.1, 14.3, 24.4, 22.8, 19....
## $ cyl <dbl> 6, 6, 4, 6, 8, 6, 8, 4, 4, 6, 8, 8, 8, 8, 8, 8, 4, 4, ...
## $ disp <dbl> 160.0, 160.0, 108.0, 258.0, 360.0, 225.0, 360.0, 146.7, 1...
## $ hp <dbl> 110, 110, 93, 110, 175, 105, 245, 62, 95, 123, 123, 180, ...
## $ drat <dbl> 3.90, 3.90, 3.85, 3.08, 3.15, 2.76, 3.21, 3.69, 3.92, 3.9...
## $ wt <dbl> 2.620, 2.875, 2.320, 3.215, 3.440, 3.460, 3.570, 3.190, 3...
## $ qsec <dbl> 16.46, 17.02, 18.61, 19.44, 17.02, 20.22, 15.84, 20.00, 2...
## $ vs <dbl> 0, 0, 1, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0, 0, 0, 1, 1, ...
## $ am <dbl> 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, ...
## $ gear <dbl> 4, 4, 4, 3, 3, 3, 4, 4, 4, 4, 3, 3, 3, 3, 3, 4, 4, ...
## $ carb <dbl> 4, 4, 1, 1, 2, 1, 4, 2, 2, 4, 4, 3, 3, 3, 4, 4, 1, 2, ...
```

## Row names

Notice that by changing the data type to tibble, the row names of `mtcars` are not the strings of car names, but numbers representing the row.

Originally, the row names are:

```
rownames(mtcars)
```

```
## [1] "Mazda RX4"           "Mazda RX4 Wag"       "Datsun 710"
## [4] "Hornet 4 Drive"      "Hornet Sportabout"   "Valiant"
## [7] "Duster 360"         "Merc 240D"           "Merc 230"
## [10] "Merc 280"           "Merc 280C"           "Merc 450SE"
## [13] "Merc 450SL"         "Merc 450SLC"         "Cadillac Fleetwood"
## [16] "Lincoln Continental" "Chrysler Imperial"   "Fiat 128"
## [19] "Honda Civic"        "Toyota Corolla"      "Toyota Corona"
## [22] "Dodge Challenger"   "AMC Javelin"         "Camaro Z28"
## [25] "Pontiac Firebird"   "Fiat X1-9"           "Porsche 914-2"
## [28] "Lotus Europa"       "Ford Pantera L"      "Ferrari Dino"
## [31] "Maserati Bora"      "Volvo 142E"
```

After conversion:

```
rownames(mtcars_tibble)
```

```
## [1] "1" "2" "3" "4" "5" "6" "7" "8" "9" "10" "11" "12" "13" "14"
## [15] "15" "16" "17" "18" "19" "20" "21" "22" "23" "24" "25" "26" "27" "28"
## [29] "29" "30" "31" "32"
```

## Summary

We can use the function `summary` to get some basic statistical information about the data.

```
summary(mtcars_tibble)
```

##	mpg	cyl	disp	hp
##	Min. :10.40	Min. :4.000	Min. : 71.1	Min. : 52.0
##	1st Qu.:15.43	1st Qu.:4.000	1st Qu.:120.8	1st Qu.: 96.5
##	Median :19.20	Median :6.000	Median :196.3	Median :123.0
##	Mean :20.09	Mean :6.188	Mean :230.7	Mean :146.7
##	3rd Qu.:22.80	3rd Qu.:8.000	3rd Qu.:326.0	3rd Qu.:180.0
##	Max. :33.90	Max. :8.000	Max. :472.0	Max. :335.0
##	drat	wt	qsec	vs
##	Min. :2.760	Min. :1.513	Min. :14.50	Min. :0.0000
##	1st Qu.:3.080	1st Qu.:2.581	1st Qu.:16.89	1st Qu.:0.0000
##	Median :3.695	Median :3.325	Median :17.71	Median :0.0000
##	Mean :3.597	Mean :3.217	Mean :17.85	Mean :0.4375
##	3rd Qu.:3.920	3rd Qu.:3.610	3rd Qu.:18.90	3rd Qu.:1.0000
##	Max. :4.930	Max. :5.424	Max. :22.90	Max. :1.0000
##	am	gear	carb	
##	Min. :0.0000	Min. :3.000	Min. :1.000	
##	1st Qu.:0.0000	1st Qu.:3.000	1st Qu.:2.000	
##	Median :0.0000	Median :4.000	Median :2.000	
##	Mean :0.4062	Mean :3.688	Mean :2.812	
##	3rd Qu.:1.0000	3rd Qu.:4.000	3rd Qu.:4.000	
##	Max. :1.0000	Max. :5.000	Max. :8.000	