

Summarizing your Data in RSudio

Reed College, Instructional Technology Services

Load the mtcars data

```
data(mtcars)
```

List summary statistics of each variable

```
summary(mtcars)
```

```
##      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
```

List summary statistics for one variable

```
summary(mtcars$mpg)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    10.40   15.42   19.20   20.09   22.80   33.90
```

List summary statistics for several variables

```
summary(mtcars[,c(1,2,4:6)])
```

```
##      mpg          cyl          hp          drat
##  Min.   :10.40   Min.    :4.000   Min.    : 52.0   Min.    :2.760
## 1st Qu.:15.43   1st Qu.:4.000   1st Qu.: 96.5   1st Qu.:3.080
## Median :19.20   Median :6.000   Median :123.0   Median :3.695
## Mean   :20.09   Mean   :6.188   Mean   :146.7   Mean   :3.597
```

```
## 3rd Qu.:22.80 3rd Qu.:8.000 3rd Qu.:180.0 3rd Qu.:3.920
## Max. :33.90 Max. :8.000 Max. :335.0 Max. :4.930
## wt
## Min. :1.513
## 1st Qu.:2.581
## Median :3.325
## Mean :3.217
## 3rd Qu.:3.610
## Max. :5.424
```

Using the describe command

```
# install.packages("psych")
library(psych)
```

```
## Warning: package 'psych' was built under R version 3.1.1
```

Describe each variable

```
describe(mtcars)
```

```
## vars n mean sd median trimmed mad min max range skew
## mpg 1 32 20.09 6.03 19.20 19.70 5.41 10.40 33.90 23.50 0.61
## cyl 2 32 6.19 1.79 6.00 6.23 2.97 4.00 8.00 4.00 -0.17
## disp 3 32 230.72 123.94 196.30 222.52 140.48 71.10 472.00 400.90 0.38
## hp 4 32 146.69 68.56 123.00 141.19 77.10 52.00 335.00 283.00 0.73
## drat 5 32 3.60 0.53 3.70 3.58 0.70 2.76 4.93 2.17 0.27
## wt 6 32 3.22 0.98 3.33 3.15 0.77 1.51 5.42 3.91 0.42
## qsec 7 32 17.85 1.79 17.71 17.83 1.42 14.50 22.90 8.40 0.37
## vs 8 32 0.44 0.50 0.00 0.42 0.00 0.00 1.00 1.00 0.24
## am 9 32 0.41 0.50 0.00 0.38 0.00 0.00 1.00 1.00 0.36
## gear 10 32 3.69 0.74 4.00 3.62 1.48 3.00 5.00 2.00 0.53
## carb 11 32 2.81 1.62 2.00 2.65 1.48 1.00 8.00 7.00 1.05
## kurtosis se
## mpg -0.37 1.07
## cyl -1.76 0.32
## disp -1.21 21.91
## hp -0.14 12.12
## drat -0.71 0.09
## wt -0.02 0.17
## qsec 0.34 0.32
## vs -2.00 0.09
## am -1.92 0.09
## gear -1.07 0.13
## carb 1.26 0.29
```

Describe one variable

```
describe(mtcars$mpg)
```

```
## vars n mean sd median trimmed mad min max range skew kurtosis
## 1 1 32 20.09 6.03 19.2 19.7 5.41 10.4 33.9 23.5 0.61 -0.37
## se
## 1 1.07
```

Describe several variables

```
describe(mtcars[,c(1, 4:6)])
```

```
##      vars  n   mean    sd median trimmed  mad   min    max  range skew
## mpg     1 32  20.09  6.03  19.20   19.70  5.41 10.40  33.90 23.50 0.61
## hp      2 32 146.69 68.56 123.00  141.19 77.10 52.00 335.00 283.00 0.73
## drat    3 32   3.60  0.53   3.70   3.58  0.70  2.76   4.93   2.17 0.27
## wt      4 32   3.22  0.98   3.33   3.15  0.77  1.51   5.42   3.91 0.42
##      kurtosis    se
## mpg     -0.37  1.07
## hp      -0.14 12.12
## drat    -0.71  0.09
## wt      -0.02  0.17
```

Describe variables by a grouping variable

```
describeBy(mtcars$mpg, mtcars$cyl)
```

```
## group: 4
##   vars  n  mean    sd median trimmed  mad   min  max range  skew kurtosis
## 1     1 11 26.66 4.51     26   26.44 6.52 21.4 33.9  12.5 0.26    -1.65
##      se
## 1 1.36
## -----
## group: 6
##   vars n  mean    sd median trimmed  mad   min  max range  skew kurtosis
## 1     1 7 19.74 1.45   19.7   19.74 1.93 17.8 21.4   3.6 -0.16    -1.91
##      se
## 1 0.55
## -----
## group: 8
##   vars n  mean    sd median trimmed  mad   min  max range  skew kurtosis
## 1     1 14 15.1 2.56   15.2   15.15 1.56 10.4 19.2   8.8 -0.36    -0.57
##      se
## 1 0.68
```