Module 3 Notes

Justin Millar 2018-09-13

Download notes as a PDF

Upcoming Assignments

Assignments	Open Time	Due Time
Module 4 Data Quiz Module 4 Conceptual Quiz		Sunday, Sept 16 (11:55 pm EST) Sunday, Sept 16 (11:55 pm EST)

Notes from Discussion Board/Office Hours

Difference between = and <-

In R both <- and = can be used to assign values to variables/objects, however there are some slight differences in how R uses these operators behind-the-scenes. This blog post demonstrates the some these differences examples.

Typically I recommend always using \leftarrow for assigning variables, and I mostly reserve = for arguments within functions. A nice shortcut for \leftarrow is Alt + \rightarrow .

```
# My preferred uses of <- and =
q <- 0.12345
rounded_q <- round(q, digits = 0.2)</pre>
```

Subsetting dataframes

Using the subset() function:

The subset() function can be used to select specific parts of a dataframe. It includes a subset argument for defining which rows will be kept, and a select argument for choosing which columns to include. Here's some examples using the mtcarrs data set:

```
# Cars with more than 6 cylinders
subset(mtcars, subset = cyl > 6)
```

```
##
                        mpg cyl disp hp drat
                                                       qsec vs am gear carb
                                                   wt
                              8 360.0 175 3.15 3.440 17.02
                                                                           2
## Hornet Sportabout
                       18.7
                                                                      3
## Duster 360
                       14.3
                              8 360.0 245 3.21 3.570 15.84
                                                                      3
                                                                           4
## Merc 450SE
                              8 275.8 180 3.07 4.070 17.40
                                                                      3
                                                                           3
                       16.4
## Merc 450SL
                       17.3
                              8 275.8 180 3.07 3.730 17.60
                                                             0
                                                                      3
                                                                           3
## Merc 450SLC
                       15.2
                              8 275.8 180 3.07 3.780 18.00
                                                                      3
                                                                           3
## Cadillac Fleetwood 10.4
                              8 472.0 205 2.93 5.250 17.98
                                                                      3
                                                                           4
                                                                      3
## Lincoln Continental 10.4
                              8 460.0 215 3.00 5.424 17.82
                                                                           4
                              8 440.0 230 3.23 5.345 17.42
                                                                      3
                                                                           4
## Chrysler Imperial
                       14.7
                                                                           2
## Dodge Challenger
                       15.5
                              8 318.0 150 2.76 3.520 16.87
                                                                      3
## AMC Javelin
                       15.2
                              8 304.0 150 3.15 3.435 17.30
                                                             0
                                                                      3
                                                                           2
## Camaro Z28
                       13.3
                              8 350.0 245 3.73 3.840 15.41
                                                                      3
                                                                           4
                                                                      3
                                                                           2
## Pontiac Firebird
                       19.2
                              8 400.0 175 3.08 3.845 17.05
                                                             0
## Ford Pantera L
                       15.8
                              8 351.0 264 4.22 3.170 14.50
```

```
## Maserati Bora
                       15.0
                              8 301.0 335 3.54 3.570 14.60 0 1
# Cars with more than 6 cylinders and high miles per gallon
subset(mtcars, subset = cyl > 6 & hp > 150)
##
                        mpg cyl disp hp drat
                                                   wt qsec vs am gear carb
## Hornet Sportabout
                       18.7
                              8 360.0 175 3.15 3.440 17.02
                                                             0
                                                                0
## Duster 360
                       14.3
                              8 360.0 245 3.21 3.570 15.84
                                                                          4
## Merc 450SE
                       16.4
                              8 275.8 180 3.07 4.070 17.40
                                                                Λ
                                                                     3
                                                                          3
## Merc 450SL
                       17.3
                              8 275.8 180 3.07 3.730 17.60
                                                                     3
                                                                          3
## Merc 450SLC
                       15.2
                              8 275.8 180 3.07 3.780 18.00
                                                             0
                                                                0
                                                                     3
                                                                          3
## Cadillac Fleetwood 10.4
                              8 472.0 205 2.93 5.250 17.98
                                                             0
                                                                     3
                                                                          4
## Lincoln Continental 10.4
                              8 460.0 215 3.00 5.424 17.82
                                                            0
                                                                0
                                                                     3
## Chrysler Imperial
                       14.7
                              8 440.0 230 3.23 5.345 17.42
                              8 350.0 245 3.73 3.840 15.41
## Camaro Z28
                       13.3
                                                             0
                                                                0
                                                                     3
                       19.2
                              8 400.0 175 3.08 3.845 17.05
                                                                     3
                                                                          2
## Pontiac Firebird
                                                             0
                                                                     5
## Ford Pantera L
                       15.8
                              8 351.0 264 4.22 3.170 14.50
                                                               1
## Maserati Bora
                       15.0
                              8 301.0 335 3.54 3.570 14.60
                                                             0
# Cars with more than 6 cylinders for high horsepower
subset(mtcars, subset = cyl > 6 | hp > 150)
##
                        mpg cyl disp hp drat
                                                   wt qsec vs am gear carb
                              8 360.0 175 3.15 3.440 17.02
## Hornet Sportabout
                       18.7
                                                                0
## Duster 360
                       14.3
                              8 360.0 245 3.21 3.570 15.84
                                                                          4
## Merc 450SE
                              8 275.8 180 3.07 4.070 17.40
                                                                          3
                       16.4
                                                             0
                                                                0
## Merc 450SL
                       17.3
                              8 275.8 180 3.07 3.730 17.60
                                                             0
                                                                0
                                                                     3
                                                                          3
## Merc 450SLC
                              8 275.8 180 3.07 3.780 18.00
                                                                     3
                                                                          3
                       15.2
                                                                0
## Cadillac Fleetwood 10.4
                              8 472.0 205 2.93 5.250 17.98
                                                             0
                                                                0
                                                                     3
                              8 460.0 215 3.00 5.424 17.82
## Lincoln Continental 10.4
                                                             0
                                                                0
                                                                     3
                              8 440.0 230 3.23 5.345 17.42
                                                             0
                                                                     3
                                                                          4
## Chrysler Imperial
                       14.7
                                                                0
## Dodge Challenger
                       15.5
                              8 318.0 150 2.76 3.520 16.87
                                                                     3
## AMC Javelin
                       15.2
                              8 304.0 150 3.15 3.435 17.30
                                                             0
                                                                0
                                                                     3
                                                                          2
## Camaro Z28
                       13.3
                              8 350.0 245 3.73 3.840 15.41
                                                                     3
                                                                          4
## Pontiac Firebird
                       19.2
                              8 400.0 175 3.08 3.845 17.05
                                                                     3
                                                                          2
                                                             Ω
                                                                Ω
## Ford Pantera L
                       15.8
                              8 351.0 264 4.22 3.170 14.50
                                                                     5
                                                                          4
## Ferrari Dino
                              6 145.0 175 3.62 2.770 15.50 0 1
                       19.7
                                                                     5
                                                                          6
## Maserati Bora
                       15.0
                              8 301.0 335 3.54 3.570 14.60 0 1
# Just keep variables of interest
subset(mtcars, subset = cyl > 6 | hp > 100, select = c("cyl", "hp"))
##
                       cvl hp
## Mazda RX4
                         6 110
## Mazda RX4 Wag
                         6 110
## Hornet 4 Drive
                         6 110
## Hornet Sportabout
                         8 175
## Valiant
                         6 105
## Duster 360
                         8 245
## Merc 280
                         6 123
## Merc 280C
                         6 123
## Merc 450SE
                         8 180
## Merc 450SL
                         8 180
## Merc 450SLC
                         8 180
## Cadillac Fleetwood
                         8 205
## Lincoln Continental
                         8 215
```

```
## Chrysler Imperial
                          8 230
## Dodge Challenger
                          8 150
## AMC Javelin
                          8 150
## Camaro Z28
                          8 245
## Pontiac Firebird
                          8 175
## Lotus Europa
                          4 113
## Ford Pantera L
                          8 264
## Ferrari Dino
                          6 175
## Maserati Bora
                          8 335
## Volvo 142E
                          4 109
```

Using operators:

We can also select part of a dataframe using brackets []. Remember that this notation uses [rows,columns] to define a subset. It may be useful to also use the \$ operator, which most commonly used is to select a single column in a dataframe using the name of the column (as opposed to the index). This will return a vector, which may be useful for subsetting in an index (recall the Battleship example for the lecture). Here are the same examples as above just using the operators:

```
# Cars with more than 6 cylinders
mtcars[mtcars$cyl > 6,]
```

```
##
                                   disp hp drat
                                                     wt
                         mpg cyl
                                                         qsec vs am gear carb
## Hornet Sportabout
                        18.7
                                8 360.0 175 3.15 3.440 17.02
                                                                   0
                                                                        3
                                                                              2
## Duster 360
                        14.3
                                8 360.0 245 3.21 3.570 15.84
                                                                0
                                                                   0
                                                                        3
                                                                              4
## Merc 450SE
                        16.4
                                8 275.8 180 3.07 4.070 17.40
                                                                        3
                                                                              3
## Merc 450SL
                        17.3
                                8 275.8 180 3.07 3.730 17.60
                                                                0
                                                                   0
                                                                        3
                                                                              3
## Merc 450SLC
                                8 275.8 180 3.07 3.780 18.00
                                                                              3
                        15.2
## Cadillac Fleetwood
                        10.4
                                8 472.0 205 2.93 5.250 17.98
                                                                0
                                                                   0
                                                                        3
                                                                              4
## Lincoln Continental 10.4
                                8 460.0 215 3.00 5.424 17.82
                                                                        3
                                                                              4
## Chrysler Imperial
                                                                0
                                                                        3
                                                                              4
                        14.7
                                8 440.0 230 3.23 5.345 17.42
                                                                   \cap
                                                                        3
                                                                              2
## Dodge Challenger
                        15.5
                                8 318.0 150 2.76 3.520 16.87
## AMC Javelin
                                8 304.0 150 3.15 3.435 17.30
                                                                0
                                                                   0
                                                                        3
                                                                              2
                        15.2
                                                                        3
## Camaro Z28
                        13.3
                                8 350.0 245 3.73 3.840 15.41
                                                                0
                                                                              4
## Pontiac Firebird
                        19.2
                                8 400.0 175 3.08 3.845 17.05
                                                                        3
                                                                              2
## Ford Pantera L
                        15.8
                                8 351.0 264 4.22 3.170 14.50
                                                                        5
                                                                              4
                                                                0
                                8 301.0 335 3.54 3.570 14.60
## Maserati Bora
                        15.0
                                                                        5
                                                                              8
```

```
# Cars with more than 6 cylinders and high miles per gallon
mtcars[mtcars$cyl > 6 & mtcars$hp > 150,]
```

```
##
                         mpg cyl
                                  disp hp drat
                                                     wt
                                                         qsec vs am gear carb
## Hornet Sportabout
                        18.7
                               8 360.0 175 3.15 3.440 17.02
                                                                        3
                                                                             2
                                                                        3
## Duster 360
                        14.3
                               8 360.0 245 3.21 3.570 15.84
                                                                   0
                                                                             4
## Merc 450SE
                        16.4
                               8 275.8 180 3.07 4.070 17.40
                                                                0
                                                                   0
                                                                        3
                                                                             3
                                                                        3
## Merc 450SL
                        17.3
                               8 275.8 180 3.07 3.730 17.60
                                                                             3
## Merc 450SLC
                               8 275.8 180 3.07 3.780 18.00
                                                                0
                                                                   0
                                                                        3
                                                                             3
                        15.2
## Cadillac Fleetwood
                        10.4
                               8 472.0 205 2.93 5.250 17.98
                                                                0
                                                                   0
                                                                        3
                                                                             4
## Lincoln Continental 10.4
                               8 460.0 215 3.00 5.424 17.82
                                                                0
                                                                   0
                                                                        3
                                                                             4
## Chrysler Imperial
                        14.7
                               8 440.0 230 3.23 5.345 17.42
                                                                        3
                                                                             4
## Camaro Z28
                               8 350.0 245 3.73 3.840 15.41
                                                                0
                                                                        3
                                                                             4
                        13.3
                                                                   Ω
## Pontiac Firebird
                        19.2
                               8 400.0 175 3.08 3.845 17.05
                                                                        3
                                                                             2
                                                                        5
## Ford Pantera L
                        15.8
                               8 351.0 264 4.22 3.170 14.50
                                                                0
                                                                   1
                                                                             4
## Maserati Bora
                        15.0
                               8 301.0 335 3.54 3.570 14.60
```

```
# Cars with more than 6 cylinders or high horsepower
mtcars[mtcars$cyl > 6 | mtcars$hp > 150,]
                                                   wt qsec vs am gear carb
##
                        mpg cyl disp hp drat
## Hornet Sportabout
                       18.7
                              8 360.0 175 3.15 3.440 17.02
## Duster 360
                              8 360.0 245 3.21 3.570 15.84
                                                                      3
                                                                           4
                       14.3
                                                                 0
## Merc 450SE
                       16.4
                              8 275.8 180 3.07 4.070 17.40
                                                                           3
## Merc 450SL
                       17.3
                              8 275.8 180 3.07 3.730 17.60
                                                                     3
                                                                           3
                                                                      3
                                                                           3
## Merc 450SLC
                       15.2
                              8 275.8 180 3.07 3.780 18.00
## Cadillac Fleetwood 10.4
                              8 472.0 205 2.93 5.250 17.98
                                                                     3
                                                                           4
## Lincoln Continental 10.4
                              8 460.0 215 3.00 5.424 17.82
                                                                      3
                                                                           4
                              8 440.0 230 3.23 5.345 17.42
                                                                     3
## Chrysler Imperial
                       14.7
                                                                           4
## Dodge Challenger
                       15.5
                              8 318.0 150 2.76 3.520 16.87
## AMC Javelin
                                                                      3
                                                                           2
                       15.2
                              8 304.0 150 3.15 3.435 17.30
                                                             0
## Camaro Z28
                       13.3
                              8 350.0 245 3.73 3.840 15.41
                                                                     3
                                                                           4
                                                                     3
                                                                           2
## Pontiac Firebird
                              8 400.0 175 3.08 3.845 17.05
                       19.2
## Ford Pantera L
                       15.8
                              8 351.0 264 4.22 3.170 14.50
                                                                1
                                                                     5
                                                                           4
## Ferrari Dino
                       19.7
                              6 145.0 175 3.62 2.770 15.50
                                                                     5
                                                                           6
## Maserati Bora
                       15.0
                              8 301.0 335 3.54 3.570 14.60
                                                                           8
# Just keep variables of interest
mtcars[mtcars$cyl > 6 | mtcars$hp > 150, c("cyl", "hp")]
##
                       cyl hp
```

```
## Hornet Sportabout
                          8 175
## Duster 360
                          8 245
## Merc 450SE
                          8 180
## Merc 450SL
                          8 180
## Merc 450SLC
                          8 180
## Cadillac Fleetwood
                          8 205
## Lincoln Continental
                          8 215
## Chrysler Imperial
                          8 230
## Dodge Challenger
                          8 150
## AMC Javelin
                          8 150
## Camaro Z28
                          8 245
## Pontiac Firebird
                          8 175
## Ford Pantera L
                          8 264
## Ferrari Dino
                          6 175
## Maserati Bora
                          8 335
```

Both the these approaches will work the same, so use whichever you prefer.

Useful keyboard shortcuts

 $\begin{array}{c|cccc} Window Mac \\ \underline{Shortcut Shortcut Action:} \\ \hline Ctrl & Cmd & Run \\ + & + & cur- \\ \hline Enter & Enter & rent \\ & & line/selection \\ \end{array}$

Window Mac ShortcutShortcutAction: Alt Alt Run ++cur-Enter Enter rent line/selection (retain cursor position) Ctrl Cmd Run +Entire Shift Shift Enter Enter Alt Alt Add + -+ -<-Show Alt Alt Shortcuts +Shift Shift + K+ K

Full list of shortcuts available here

Extracting files from compressed (zipped) folders

In this course (and perhaps in your research) we will sometimes work with large files that must be compressed. Typically we will use .zip compression, which can be extracted by most newer PC and Mac without additional software. Other compression may require specific software to unzip (check the Properties of the folder, and try Google if you are unsure how to open). It is important that you either Extract the contents of the compressed folder, or manually move the files to a new location once to open the folder. This is because R will not allow you to specify a path to a file in a compressed folder.

Other notes

- Cheat sheet for base R
- List of other cheat sheets.
- Dataframes can be exported as CSV files using the write.csv() function.
- RStudio is available on UFApps.
- You can run RStudio on the cloud at https://rstudio.cloud/.
- This video demonstrates how to find and remove outliers using the identify() function.
- Here's a site with practice problems.