Solutions to G. Grolemund & H. Wickhams's R for Data Science, Chapter 4

Krista DeStasio

7/16/2017

Contents

A Brief Introduction to This File

This R file walks through G. Grolemund & H. Wickhams's online text, "R for Data Science." Much of the code is sourced directly from the book and credit belongs to the authors. Here, some sections of code are heavily commented so that the beginning R programmer can read through and understand what each line of code does and compare it to their own as they work through the text. Throughout, the book provides the primary and most thorough explanation. **For the greatest learning benefit, I suggest you attempt each exercise on your own before looking at the code or write-ups provided here.** Of course, there is more than one way to write code and you may find a more elegant solution that you prefer.

For those new to R and RStudio, it may be of additional benefit to knit the document and examine how the code in the Rmd file is visually expressed in the resultant knitted document. For example, see how the `["R for Data Science."](http://r4ds.had.co.nz/index.html)` is expressed as a hyperlink in the preceeding paragraph where it was not surrounded by tick-marks and compare that to how the same text is expressed in this paragraph when surrounded by ticks. See also the difference in appearance when knitting to different document types (HTML, PDF, Word).

Tip: *If you are using RStudio, click the text next to the orange # box at the bottom of the editor window to easily navigate the code chunks.*

Tip: *Use the* `?` *before any command to view the documentation on that function. Do this often. For example, type* `?setwd` *to see a description, usage, arguments, and more for the function* `setwd()`.

Tip: Find RStudio Cheatsheets at https://www.rstudio.com/resources/cheatsheets/

Chapter 4, Workflow: basics

```
seq(1, 10)
```

```
##  [1]  1  2  3  4  5  6  7  8  9 10
```

```
x <- "hello world"
(y <- seq(1, 10, length.out = 5))
```

```
## [1]  1.00  3.25  5.50  7.75 10.00
```

Practice

1. Why does this code not work?

```
my_variable <- 10
my_varıable
```

```
my_variable <- 10
my_variable
```

```
## [1] 10
```

2. Tweak each of the following R commands so that they run correctly:

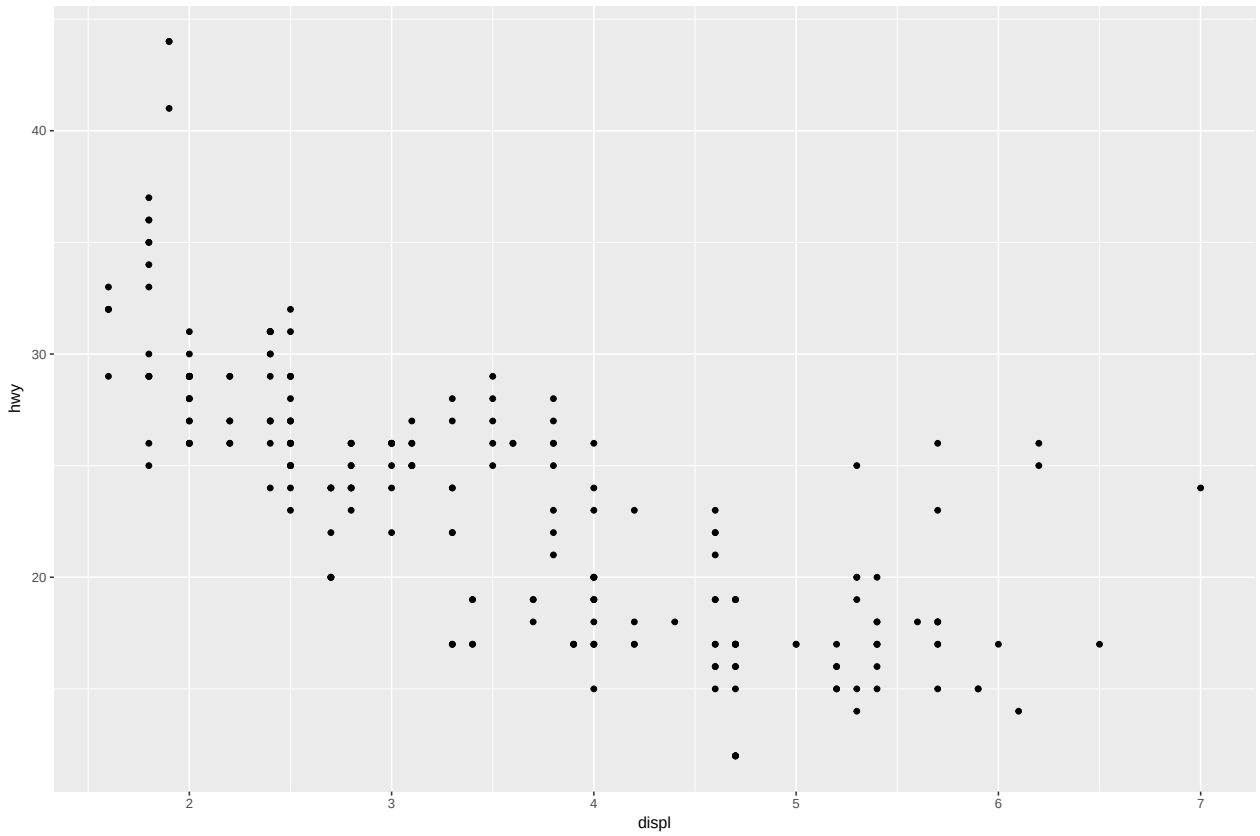
```
library(tidyverse)

ggplot(dota = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy))

fliter(mpg, cyl = 8)
filter(diamond, carat > 3)
```

```
library(tidyverse)

ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy))
```




```
filter(mpg, cyl == 8)
```

```
## # A tibble: 70 x 11
##    manufacturer             model displ  year   cyl      trans   drv
##           <chr>             <chr> <dbl> <int> <int>      <chr> <chr>
##  1         audi         a6 quattro   4.2  2008     8   auto(s6)     4
##  2    chevrolet c1500 suburban 2wd   5.3  2008     8   auto(l4)     r
##  3    chevrolet c1500 suburban 2wd   5.3  2008     8   auto(l4)     r
##  4    chevrolet c1500 suburban 2wd   5.3  2008     8   auto(l4)     r
##  5    chevrolet c1500 suburban 2wd   5.7  1999     8   auto(l4)     r
##  6    chevrolet c1500 suburban 2wd   6.0  2008     8   auto(l4)     r
##  7    chevrolet           corvette   5.7  1999     8 manual(m6)     r
##  8    chevrolet           corvette   5.7  1999     8   auto(l4)     r
##  9    chevrolet           corvette   6.2  2008     8 manual(m6)     r
## 10    chevrolet           corvette   6.2  2008     8   auto(s6)     r
## # ... with 60 more rows, and 4 more variables: cty <int>, hwy <int>,
## #   fl <chr>, class <chr>
```

```
filter(diamonds, carat > 3)
```

```
## # A tibble: 32 x 10
##    carat     cut color clarity depth table price     x     y     z
##    <dbl>   <ord> <ord>   <ord> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
##  1  3.01 Premium     I      I1  62.7    58  8040  9.10  8.97  5.67
##  2  3.11    Fair     J      I1  65.9    57  9823  9.15  9.02  5.98
##  3  3.01 Premium     F      I1  62.2    56  9925  9.24  9.13  5.73
##  4  3.05 Premium     E      I1  60.9    58 10453  9.26  9.25  5.66
##  5  3.02    Fair     I      I1  65.2    56 10577  9.11  9.02  5.91
```

```
##  6  3.01    Fair     H      I1  56.1    62 10761  9.54  9.38  5.31
##  7  3.65    Fair     H      I1  67.1    53 11668  9.53  9.48  6.38
##  8  3.24 Premium     H      I1  62.1    58 12300  9.44  9.40  5.85
##  9  3.22   Ideal     I      I1  62.6    55 12545  9.49  9.42  5.92
## 10  3.50   Ideal     H      I1  62.8    57 12587  9.65  9.59  6.03
## # ... with 22 more rows
```

3. Press Alt + Shift + K. What happens? How can you get to the same place using the menus?

`Alt + Shift + k` provides a keyboard shortcut quick reference. To access the reference sheet from the menus, go to Tools > Keyboard shortcuts help