Write Clean Code in R

This lesson reviews best practices associated with clean coding.

Learning Objectives

At the end of this activity, you will be able to:

- write code using Hadley Wickham's style guide

What you need

You need `R` and `RStudio` to complete this tutorial. Also we recommend have you have an `earth-analytics` directory setup on your computer with a `/data` directory with it.

- How to Setup R / R Studio
- Setup your working directory

Resources

- Hadley Wickham's style guide

Clean code means that your code is organized in a way that is easy for you and for someone else to follow / read. Certain conventions are suggested to make code easier to read. For example, many guides suggest the use of a space after a comment. Like so:

```
#poorly formatted  comments are missing the space after the pound sign.
# good comments have a space after the pound sign
```

While these types of guidelines may seem unimportant when you first begin to code, after a while you're realize that consistently formatted code is much easier for your eye to scan and quickly understand.

Consistent, Clean Code

Take some time to review Hadley Wickham's style guide. From here on in, we will follow this guide for all of the assignments in this class.

Object Naming Best Practices

1. **Keep object names short:** this makes them easier to read when scanning through code.
2. **Use meaningful names:** For example: `precip` is a more useful name that tells us something about the object compared to `x`
3. **Don't start names with numbers!** Objects that start with a number are NOT VALID in R.
4. **Avoid names that are existing functions in R:** e.g., `if`, `else`, `for`, see here

A few other notes about object names in `R`:

- `R` is case sensitive (e.g., `weight_kg` is different from `Weight_kg`).
- Avoid other function names (e.g., `c`, `T`, `mean`, `data`, `df`, `weights`).
- Use nouns for variable names, and verbs for function names.

- Avoid using dots in object names - e.g. `my.dataset` - dots have a special meaning in R (for methods) and other programming languages. Instead use underscores `my_dataset`.

Challenge

Take a look at the code below.

- Create a list of all of the things that could be improved to make the code easier to read / work with.
- Add to that list things that don't fit the Hadley Wickham style guide standards.
- Try to run the code in R. Any issues?

```r
#my code

#load stuff
library(ggplot2)

#turn off factors
options(stringsAsFactors = FALSE)

1variable <- 3 * 6
meanVariable <- 1variable

#calculate something important
mean-variable <- meanvariable * 5

thefinalthingthatineedtocalculate <- mean-variable + 5

#get things that are important
download.file(url = "https://ndownloader.figshare.com/files/7010681",
              destfile = "data/boulder-precip.csv")

my.data <- read.csv(file="data/boulder-precip.csv")
head(my_data)

str(my.data)

qplot(x=my.data$DATE,
      y=my.data$PRECIP)
```