Week 1: Welcome to R Statistical Programming

08/29/2019

Jake Campbell

Why R?

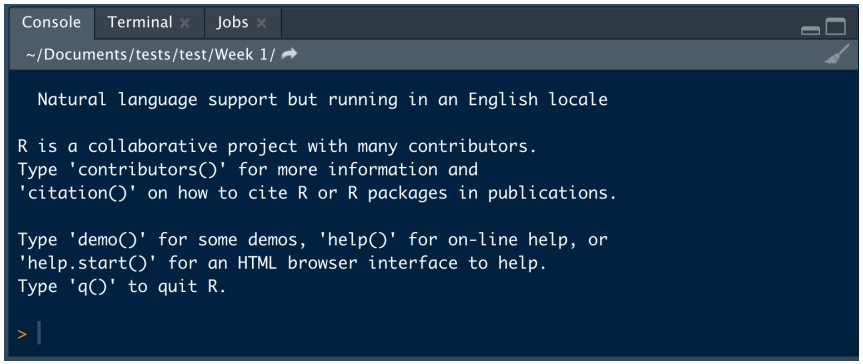
- It's open source
  - Why pay for licenses for other tools?
- There is lots of good documentation
- Extremely large and helpful community
  - Stack overflow is your friend

Rstudio

- R is the language, but it can be pretty ugly to use
- Rstudio is a GUI for R that will make your life easier
    - You can live without it, but why would you?

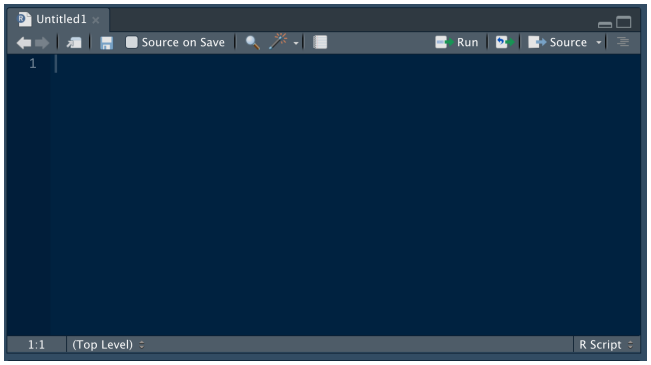
Rstudio: The Console

- The console is where we'll push code and see output
- We can type code directly into it
- The code we type in isn't saved!
    - We can go back in history, but it's not unlimited



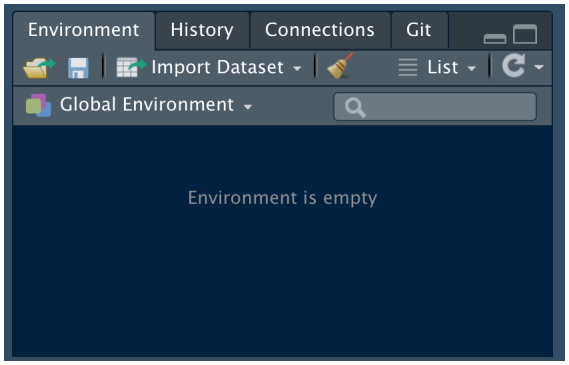
Rstudio: The Text Editor

- Here we can load scripts directly into R
  - Doesn't necessarily need to be R scripts
- We can push code to the console using that run button
  - Always easier to have that as a keyboard shortcut



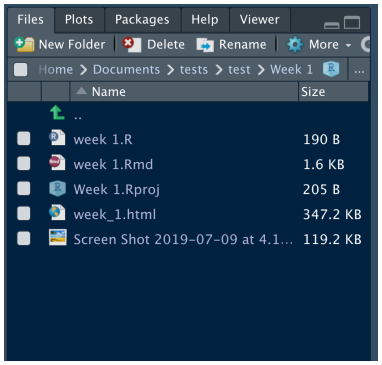
Rstudio: The Environment

- This is where we can see what is stored in our environment
  - Could be data, functions, models, etc.
- Essentially where what we save is stored
- Also includes a history tab for past code as well as a Git staging area



Rstudio: The Viewer Pane

- The viewer pane has a lot of different functions
  - We can use it as a file viewer
  - This is where plots we create will show
  - This is where the help screen is



Bare Bones R

- R can be used as a calculator

```
1 + 1
```

```
## [1] 2
```

```
25 / 5
```

```
## [1] 5
```

```
4 * 3
```

```
## [1] 12
```

Assigning to Objects

- We can store different aspects of our work in objects
    - Data, output, functions, etc.
- Store what you want to keep, or else it's gone!

Assigning to Objects

- We can use the assignment operator `<-` to store objects

```
test_output <- 1 + 1
```

- Might be more intuitive to use `=`, but that will cause confusion later on

Assigning to Objects

- `test_output` is now stored in our environment
- When we call `test_output`, it will show what we stored in it

```
test_output
```

```
## [1] 2
```

-We can also perform functions on this object

```
test_output + 3
```

```
## [1] 5
```

Object Names

- We should be descriptive, but not overly-complicated with our object names
  - `foo` isn't descriptive and wouldn't mean anything to us
  - `test_output_from_model_1_set_b_where_i_added_a_variable` is way too complicated for a name
- Just be straightforward and succint
  - `model_gbm` is probably holding a `gbm` model in it
  - Quick and to the point

Functions

- R has several built in functions
    - These allow us to code common problems much more quickly
- For example, to get the absolute value of a number, we could use `abs()`
- We can look up the documentation for a function by typing `?` before it

```
abs(x = -1)
```

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

```
##
## # Look up the help page
## ?abs
```