

BST 219

Core Principles of Data Science

Lecture 8: R Basics Continued
September 26, 2024

Recipe of the Day!

Matcha Tres Leches Cake (Credit to Jennifer)



When you find the bug in your code
and fix it on the first try

Agenda

- Announcements
 - Lab this week! FXB G03 and Zoom
 - Homework #1 due 9/27 by 11:59pm
 - Plots: for this HW only axis labels do not need to be “pretty” or intuitive/informative. We will be learning how to do this in the next module.
- Wrap up of coding basics
 - Importing data
 - Writing functions
- Start the visualization module!



`dplyr::filter()` KEEP ROWS THAT satisfy your CONDITIONS

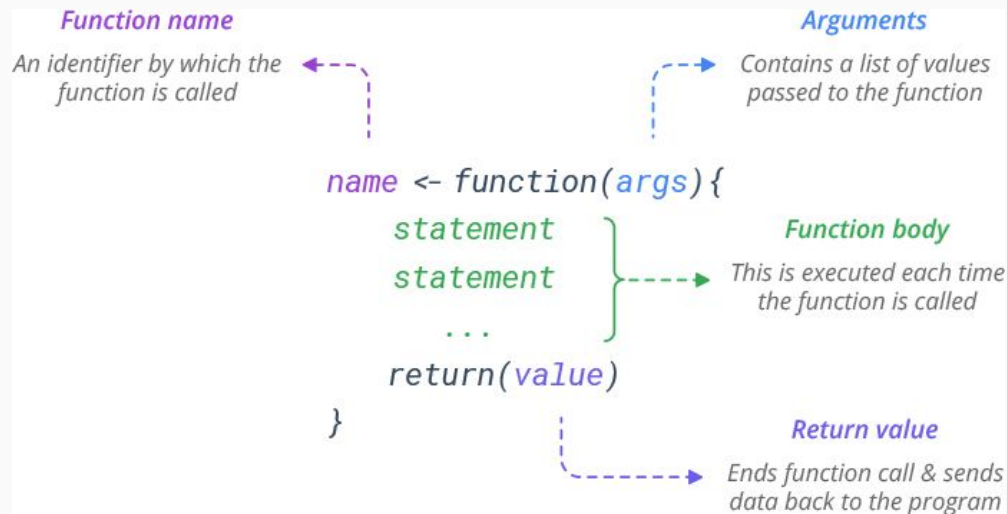
keep rows from... this data... ONLY IF... type is "otter" AND site is "bay"

```
filter(df, type == "otter" & site == "bay")
```

A cartoon illustration for the `dplyr::filter()` function. A green character is pointing to a map of the United States. A purple character is standing next to a table. The table has columns 'type', 'food', and 'site'. The rows are: 'otter', 'urchin', 'bay'; 'shark', 'seal', 'channel'; 'otter', 'abalone', 'bay'; 'otter', 'crab', 'wharf'. The first row is highlighted in green. The second row has an 'X' next to 'channel'. The third row has a checkmark next to 'bay'. The fourth row has an 'X' next to 'wharf'.

type	food	site
otter	urchin	bay
shark	seal	channel
otter	abalone	bay
otter	crab	wharf

The Anatomy of Functions



Good Practices for Functions

1. Use meaningful names for your function and the arguments
2. Make your functions short and simple - try to have each function do one thing
3. Use an explicit `return()` statement at the end of your function

Where we are in the data science pipeline

(But we'll come back to this with more advanced techniques!)

