



dplyr Overview

Nick Ulle



dplyr

dplyr is a “grammar of data manipulation”.

- A popular Tidyverse package
- Functions for taking subsets
 - We already learned how to use base R to take subsets
- Functions for aggregate statistics
 - We'll learn how to use base R for this soon



Learning dplyr

Like most Tidyverse packages, dplyr has:

- [Online documentation](#)
- [A cheat sheet](#)



Subsets with dplyr

Nick Ulle





Base R versus dplyr

Nick Ulle





Transformations with dplyr

Nick Ulle



The Pipe



Nick Ulle

Pipe Operator Drawbacks

The pipe operator is the most controversial tool in R.

Disadvantages of the pipe:

- Yet another dialect to learn.
- Problems are hard to diagnose.
 - Without intermediate variables, it's hard to check each step.
- Error messages are obfuscated.

Use the pipe if you like it, avoid it if you don't.