

Introduction to R for Data Management and Analysis

Marcel Ramos

Tuesday, June 25, 2019

Notes from previous class

- t-test with a formula vs vector
- object names with a period (.)
- apply over a data.frame

General R Learning Tips

- Troubleshooting is important
 - Know your errors
- Know your inputs and your outputs
- Practice, practice, practice
- Make use of online resources, help pages
- Be curious
 - test your questions with examples

Troubleshooting

- Getting help within R
 - Documentation pages (examples)
- Getting help outside of R
 - stackoverflow.com
 - Quick-R website
- Running code line by line
 - Identify where in the script the error is occurring
 - Isolate the issue by breaking the process into small pieces

Exploring objects

- head & tail
- class
- length or dim
- dimnames
- attributes
- str
- args on a function name

- Create a catalog of functions that are useful for a particular task
- Use your reference websites / textbooks
 - R Cookbook by Winston Chang - ggplot2
 - Quick-R
 - R for Data Science - Hadley Wickham
 - Handbook for Statistical Analyses using R - 3rd Ed.

Data Analysis Workflow

- RStudio setup
- import
- clean
 - subset / divide
- reshape
- summarize
- explore
- graph
- report