Intro to R/RStudio

How can we interact with R?

RStudio

Scripts

Type multiple lines

before running

Plots

browser, package

browser, and for

Displays plots

Also used as file

getting help

· Savable code

Command Line

RStudio Interface

Console

Runs R commands

- Displays output
- Runs scripts or single commands

Graphical Output

<u>Environment</u>

Environment: Shows data

- History: Shows commands ran
- Both are savable

How can we load our data into R?

Analyzing Patient Data

<u>Set Directory</u> → <u>Load Data</u>

setwd()

read.csv()

RStudio

Getting help: read.csv()
Arguments & defaults

How do we work with loaded data?

Assigning Variables

- What is a variable?
- Defining a simple variable (<-)
- Printing variables
- Arithmetic w/ variables operators
- Loading datasets into variables
- SIDENOTES: how-to commenting

Working with Data/Variables

- Type of data (class ())
- Checking dimensions (dim())
- Accessing data: single values, slicing, combinations, by column name
- Applying calculations (max, min, mean, median, sd)
- Applying with MARGIN (column vs row)
- Column & row specific calculations

Plotting Data

- Creating a basic plot
- EXAMPLE: Finding errors in data w/ plotting

Data Types & Structures

How is data stored in R?

Data Types

- What & Why
- 6 Types: Charactar, numeric integer, logical, complex (factors later)
- Examining datatypes: class(),

typeof(),
length(),
attributes()

Data StructuresWhat & Why

- Types: atomic vectors, list, matrix, dataframw
- Vectors: about, modes, examining, creating, adding, missing data, special types, mixing types, attributes
- Matrix: rows/columns, constructing, dimensions
- List: why(not limited to single types), creating lists, accessing, assessing
- Dataframes: what/why, creating, special functions

How do I deal with categorical data in R?

Factors

Factors

What is a factor?

- Categorical representation
- Stored as integers with associated labels
- Levels
- Ordered vs unordered

1

Creating & Modifying Factors

- factor(), level(), nlevel()
- Ordering factors
- Converting factors

Using Factors

- Loading datasets w/ strings as factors
- Observing with str()
- Using summary()
- Removing factor levels: reassigning, dropping levels, shortcut

Accessing Data

How do I access my stored data?

<u>3 ways!</u>

<u>by Index</u>

- [row(s), column(s)]
- Single values & sequences
- Combinations

by Name

- Getting columns by names: names ()
- Addressing by name: \$, []
- Why this is best practice

by Logical Vector

- What? (vector of TRUE/FALSE)
- Creating w/ relational operators
- Selecing using logical operators
- logical vectors & assignments (<-)

Looping

How do I do repeated operations?

<u>3 ways!</u>

Vectorized Operations

- What/why
- Examples (i.e. sum ())
- Importance
- Vector Recycling

Apply Functions

- What/Why/How
- Different apply functions
- When to use
- Example

For Loops

- When to use (vs apply)
- 3 golden rules of for loops in R
- Syntax
- Example: For loop of same example as apply – wrong way then right way
- Timing our various repeated operations