## RPROGRAMING

An Introduction

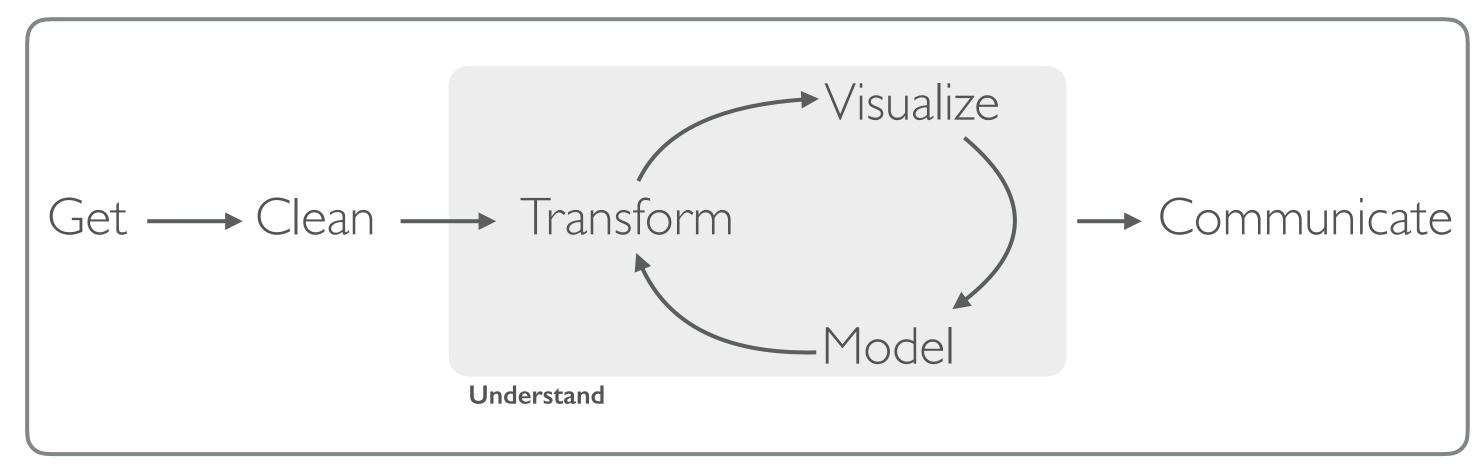
## HELLO my name is

# Bradd

- bradleyboehmke.github.io
- □ bradleyboehmke@gmail.com
- bradleyboehmke

#### SETTING THE EXPECTATIONS

- Introduction to R
- Intermediate R
- Applied Analytics with R



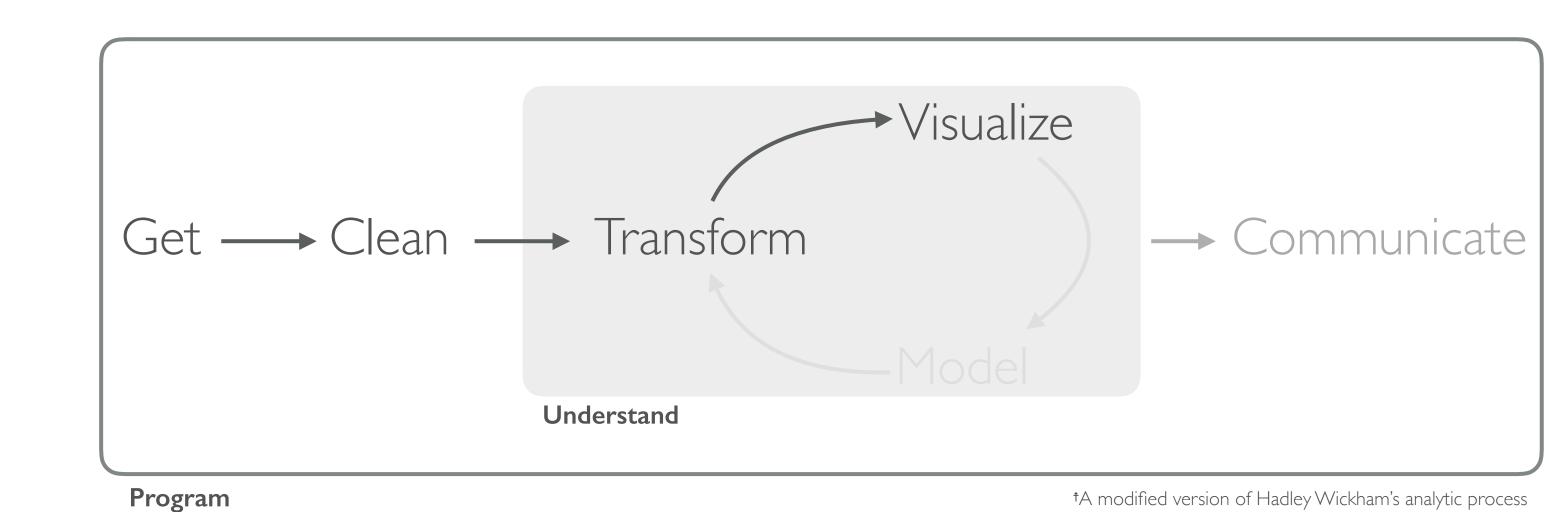
Program

†A modified version of Hadley Wickham's analytic process

#### SETTING THE EXPECTATIONS

#### Day I

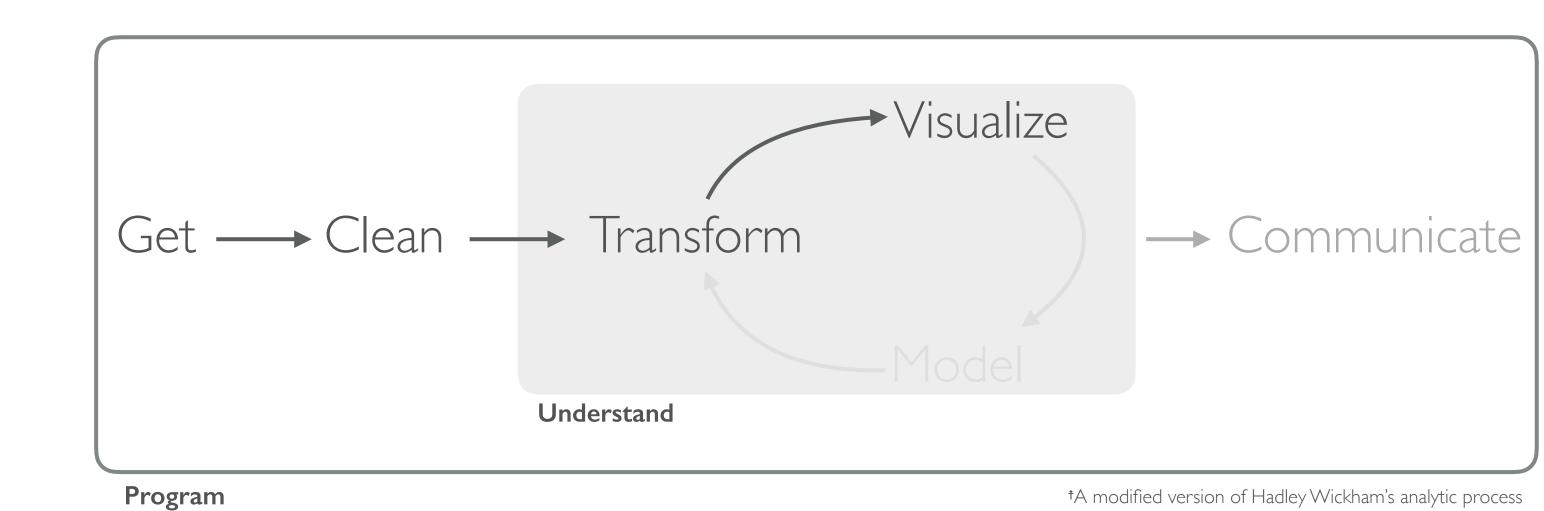
- Fundamentals
- Visualization with ggplot2
- Data transformation
- Data exploration
- Data structures



#### SETTING THE EXPECTATIONS

#### Day 2

- Importing/exporting data
- Tidy data
- Relational data
- Factors, dates, & strings
- Little bit of R markdown





You will be overwhelmed!

My Teaching Philosophy

#### YOURTURN!

Lots of hands-on coding exercises

# Strong proponent of collaborative work!



#### YOURTURN!

Introduce yourself to your neighbors. Who are you and what is your experience with R?

### FUNDAMENTALS

#### INSTALLATION



- I. Go to https://cran.r-project.org/
- 2. Click "Download R for Mac/Windows"
- 3. Download the appropriate file:
  - Windows users click Base, and download the installer for the latest R version
  - Mac users select the file R-3.X.X.pkg that aligns with your OS version
- 4. Follow the instructions of the installer



- I. Go to RStudio for desktop https://www.rstudio.com/products/rstudio/download/
- 2. Select the install file for your OS
- 3. Follow the instructions of the installer.

Note: There are other R IDE's available: Emacs, Microsoft R Open, Notepad++, etc.

#### UNDERSTANDING THE CONSOLE

#### ~/Desktop/Personal/Academia/Learning Curve R Package - RStudio ♀ ☐ ☐ ☐ ☐ ☐ Go to file/function □ → Addins → Learning Curve R Package + Script files Workspace environment initial\_functions.R × Plot\_functions.R × Environment History - Holds your objects - Saves your script 😭 🖳 🔝 Import Dataset 🕶 🦂 Run Source - = - Allows code & comments - Can review history Global Environment \* Q 232 - Can have multiple files Values 233 10 open at a time 235 # Provides the summary for the block containing units m through n, n > m 236 # t = time for firsts unit 0.77 237 # m = lower bound unit of production block 5 238 # n = upper bound unit of production block Functions 239 # r = learning curve rate function (t, m, n, r, na.rm = FALSE) block\_summary 240 - block\_summary <- function(t, m, n, r, na.rm = FALSE){ function (t, m, n, r, na.rm = FALSE) cum\_appx 242 ~ if(!is.numeric(t) | !is.numeric(m) | !is.numeric(n) | !is.numeric(r)){ cum\_exact function (t, m, n, r, na.rm = FALSE) 243 stop('This function only works for numeric inputs!\n', function (b, na.rm = FALSE)lc\_rate 'You have provided objects of the following classes:\n', function (m, n, r, na.rm = FALSE)midpoint 245 't: ', class(t), '\n', natural\_slope function (r, na.rm = FALSE) 'm: ', class(n), '\n', function (t, n, r, na.rm = FALSE) 247 'n: ', class(n), '\n', 248 'r: ', class(r)) 249 Files Plots Packages Help Viewer Misc - Displays: 250 (a) 🖒 🏠 🗐 Q mean - files in working 276:20 Dlock\_summary(t, m, n, r, na.rm) = R Script \$ R: Arithmetic Mean - Find in Topic directory Console/Command line Console ~/Desktop/Personal/Academia/Learning Curve R Package/ Academia/Learning Curve R Package/ - plots when produced R Documentation mean {base} - Can use as calculator \$`block hours` - help files/search - Does not save code [1] 3668.436 Arithmetic Mean - This is where your output \$`midpoint unit` is displayed [1] 44.03189 Description \$`midpoint hours` Generic function for the (trimmed) arithmetic mean. [1] 40.31249 Usage mean(x, ...) ## Default S3 method: > m + n \* t^r mean(x, trim = 0, na.rm = FALSE, ...)[1] 355.3082

Thorough tutorial regarding the RStudio console: <a href="http://dss.princeton.edu/training/RStudio101.pdf">http://dss.princeton.edu/training/RStudio101.pdf</a>

Arguments

#### GETTING HELP

```
# provides details for specific function
help(sqrt)

# provides same information as help(functionname)
?sqrt

# provides examples for said function
example(sqrt)
```

#### External to R:

Google: just add "with R" at the end of any search.

Stack Overflow: a searchable Q&A site oriented toward programming issues. 75% of my answers come from SO

Cross Validated: a searchable Q&A site oriented toward statistical analysis.

**R-bloggers:** a central hub of content from over 500 bloggers who provide news and tutorials about R.

#### SETYOUR WORKING DIRECTORY

```
# get your current working directory
getwd()
[1] "/Users/bradboehmke/Dropbox/Academia/University of Cincinnati/Intro to R Bootcamp"

# set your working directory
setwd("/Users/bradboehmke/Dropbox/Academia/University of Cincinnati")

getwd()
[1] "/Users/bradboehmke/Dropbox/Academia/University of Cincinnati"
```

Keeping your files organized is critical

#### YOURTURN!

Set your working directory to the "Intro to R" folder you downloaded for this course.

#### RASACALCULATOR

```
# Uses PEMDAS convention for order of operations
4 + 3 / 10 ^ 2
## [1] 4.03
4 + (3 / 10 ^ 2)
## [1] 4.03
(4 + 3) / 10 ^ 2
## [1] 0.07
# large/small numbers will be displayed in scientific notation
1 / 17 ^ 7
## [1] 2.437011e-09
# Undefined calculations result in Inf or NaN
1 / 0
## [1] Inf
Inf - Inf
## [1] NaN
```

## THE ASSIGNMENT (<-) OPERATOR

```
x <- 3 # GOOD
x = 3 # BAD
# we can increment (build onto) existing objects
x < -x + 1
X
## [1] 4
# must be specific
Error: object 'x' not found
```

#### YOURTURN!

#### Economic Order Quantity Model:

$$Q = \sqrt{\frac{2DK}{h}}$$

Calculate Q where:

$$D = 1000$$

$$K = 5$$

$$h = 0.25$$

hint: 
$$sqrt(x) = \sqrt{x}$$

### SOLUTION

```
D <- 1000

K <- 5

h <- .25

Q <- sqrt((2 * D * K) / h)

Q

## [1] 200
```

#### WORKSPACE ENVIRONMENT

- You should now have 4 objects in your global environment
- History tab will show your recent code

To list and remove objects in your global environment:

```
# list all objects
ls()
## [1] "D" "h" "K" "Q"

# remove defined object from the environment
rm(D)

# removes everything in the working environment -- use with caution!
rm(list = ls())
```

```
Environment History
🚰 🔚 🔛 Import Dataset 🕶 🥑
                                                        List *
Values
                  1000
                  0.25
                                  Environment History
                                                                                               🕣 🕞 💽 To Console 🚅 To Source 🥝 🎻
                                                          Load 212 more entries
                                 D <- 1000
                                 Q <- sqrt((2 * D * K) / h)
```

#### PACKAGES

The fundamental unit of shareable code is the package.

**CRAN:** 10,000+ **Bioconductor:** 1,000+

GitHub: Many more plus beta versions for updated packages not yet published

So how do we install these packages?

```
# install packages from CRAN
install.packages("packagename")

# install packages from Bioconductor
source("http://bioconductor.org/biocLite.R")  # only required the first time
biocLite()  # only required the first time
biocLite("packagename")

# install packages from GitHub
install.packages("devtools")  # only required the first time
devtools::install_github("username/packagename")
```

#### YOURTURN!

Download these packages from CRAN:

tidyverse nycflights13

#### SOLUTION

```
install.packages("tidyverse")
install.packages("nycflights13")

# alternative
install.packages(c("tidyverse", "nycflights13"))
```

For a full list of useful packages see this guide: <a href="http://bit.ly/1x9vkzV">http://bit.ly/1x9vkzV</a>

#### PACKAGES

#### Loading packages:

```
# load the package to use in the current R session
library(tidyverse)

# use a particular function within a package without loading the package
stringr::str_replace()
```

#### Getting help on packages:

```
# provides details regarding contents of a package
help(package = "tidyr")

# list vignettes available for a specific package
vignette(package = "tidyr")

# view specific vignette
vignette("tidy-data")
```

## WHATTO REMEMBER

#### FUNCTIONS TO REMEMBER

Operator/Function	Description
help(), ?, example()	Get help on functions and provide examples
<pre>getwd(), setwd()</pre>	Get and set your working directory
+, -, *, /, ^	Arithmetic
<-	Assignment operator
ls(), rm()	list and remove objects in your global environment
<pre>install.packages(), library()</pre>	Install and load packages
vignette()	View/list package vignette