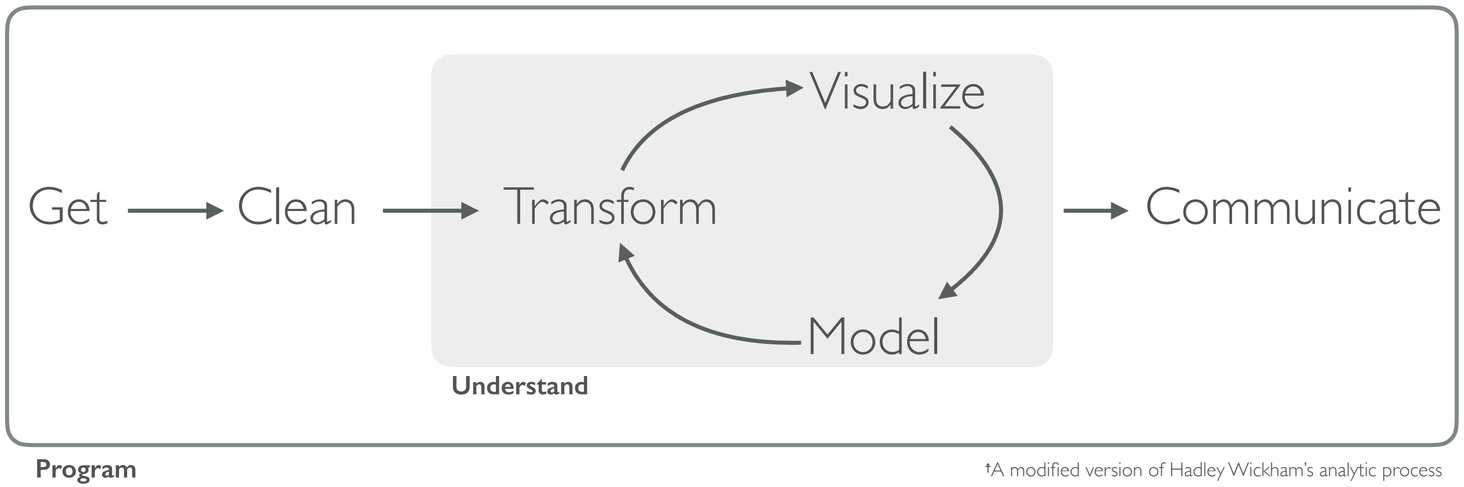
R PROGRAMMING

An Introduction

bradleyboehmke.github.io
bradleyboehmke@gmail.com
@bradleyboehmke
bradleyboehmke

SETTING THE EXPECTATIONS

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

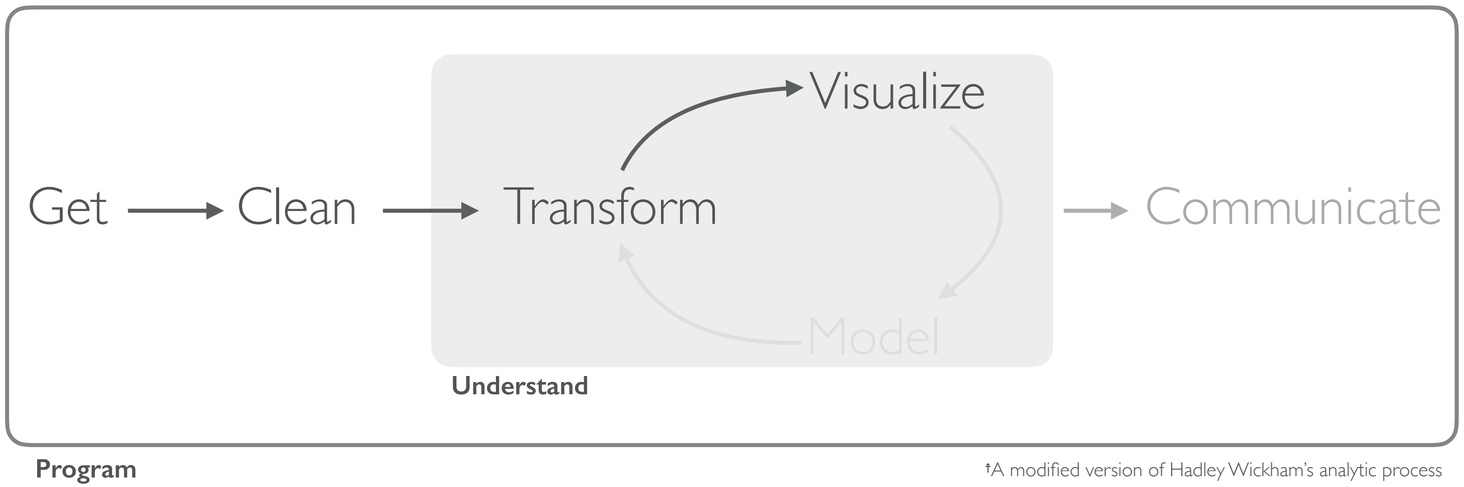


†A modified version of Hadley Wickham's analytic process

SETTING THE EXPECTATIONS

Day 1

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

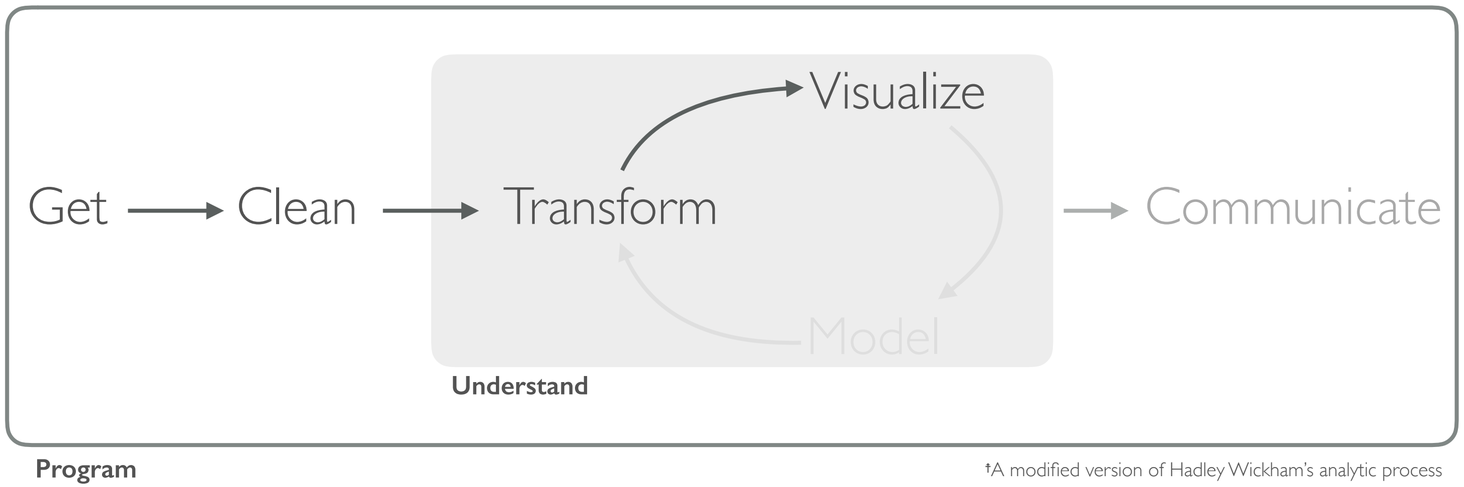


†A modified version of Hadley Wickham's analytic process

SETTING THE EXPECTATIONS

Day 2

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



†A modified version of Hadley Wickham's analytic process

You will be overwhelmed!

My Teaching Philosophy

YOUR TURN!

Lots of hands-on coding exercises

Strong proponent of collaborative work!

YOUR TURN!

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

FUNDAMENTALS

INSTALLATION

1. 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

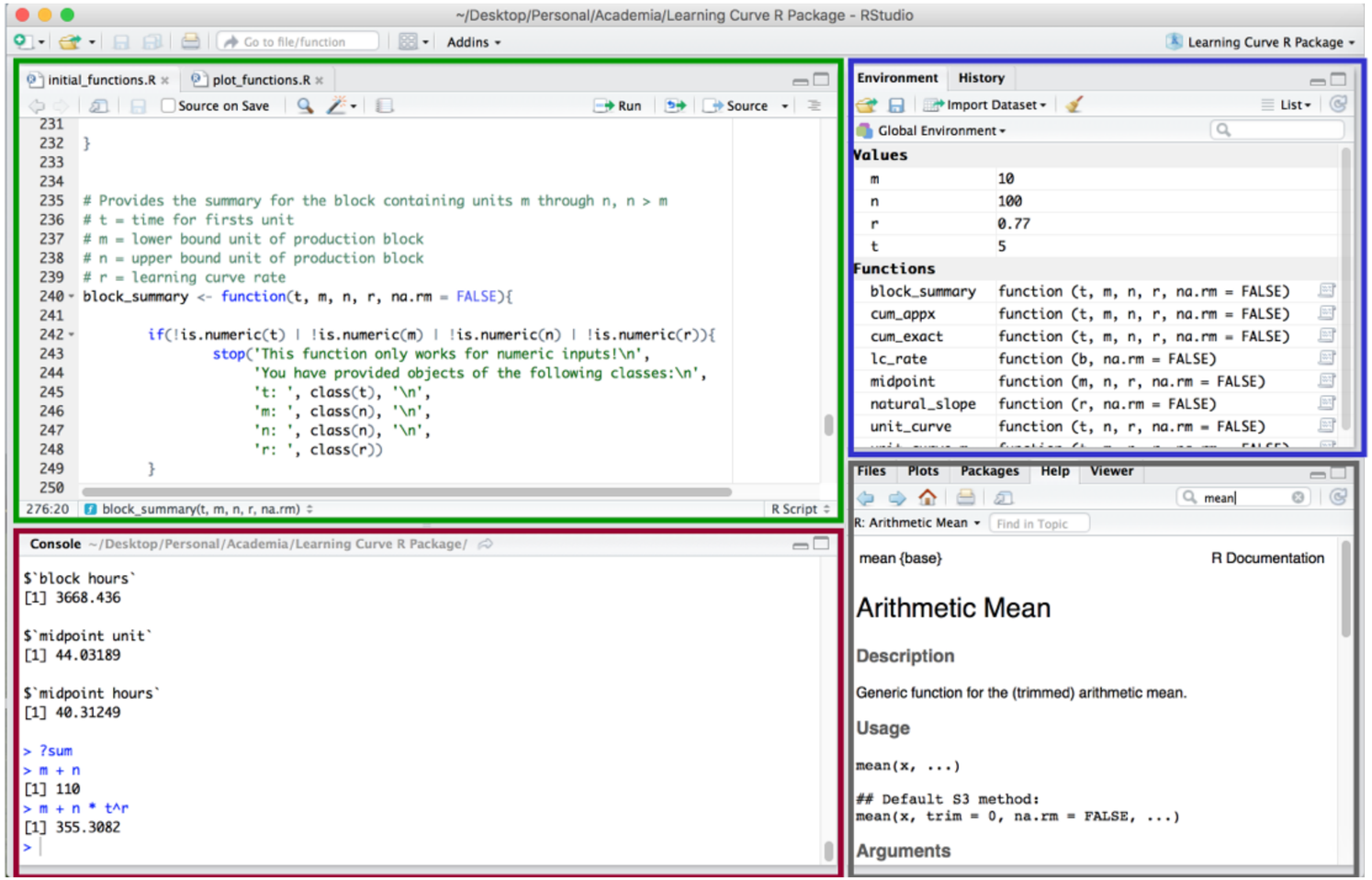
1. Go to RStudio for desktop https://www.rstudio.com/products/rstudio/download/
2. Select the install file for your OS
3. Follow the instructions of the installer.

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

UNDERSTANDING THE CONSOLE

Script files
- Saves your script
- Allows code & comments
- Can have multiple files open at a time

Console/Command line
- Can use as calculator
- Does not save code
- This is where your output is displayed



Workspace environment
- Holds your objects
- Can review history

Misc - Displays:
- files in working directory
- plots when produced
- help files/search

Thorough tutorial regarding the RStudio console: http://dss.princeton.edu/training/RStudio101.pdf

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.

SET YOUR 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

YOUR TURN!

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

R AS A CALCULATOR

```
# 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
X
Error: object 'x' not found
```

YOUR TURN!

Economic Order Quantity Model:

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

Calculate **Q** *where:*

D = 1000

K = 5

h = 0.25

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

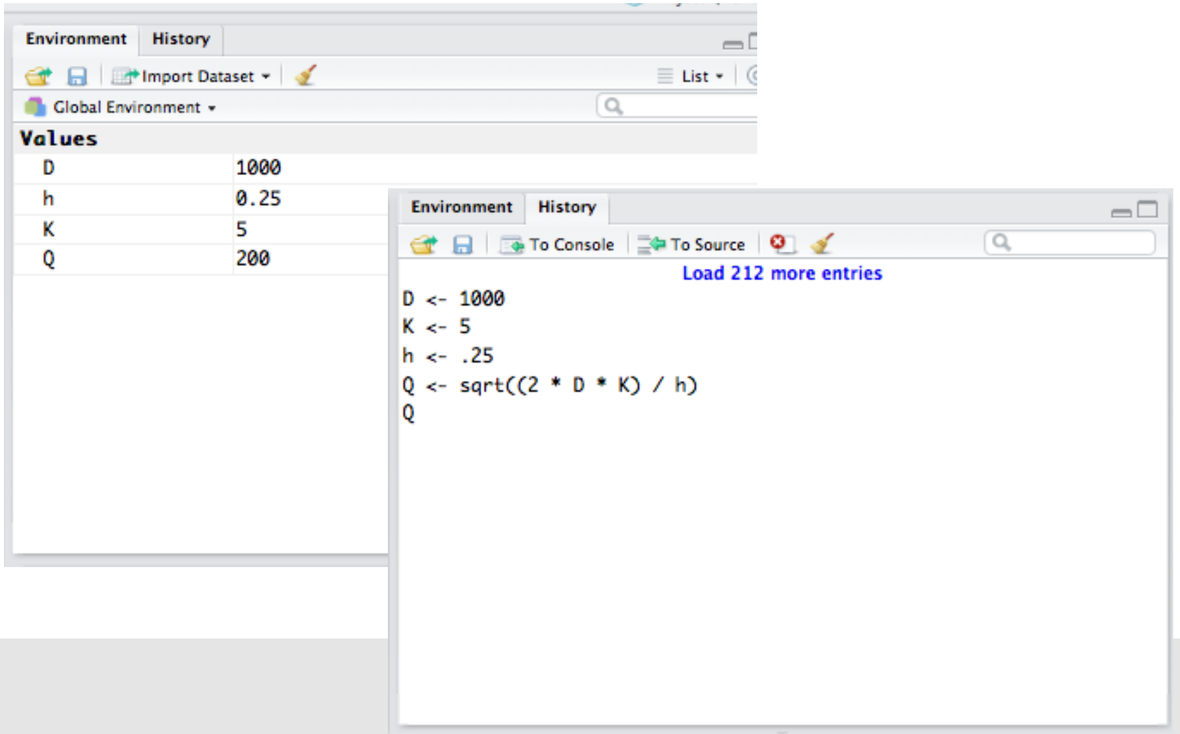
SOLUTION

```
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())
```

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")
```

YOUR TURN!

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: http://bit.ly/1x9vkzV

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")
```

WHAT TO REMEMBER

FUNCTIONS TO REMEMBER

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