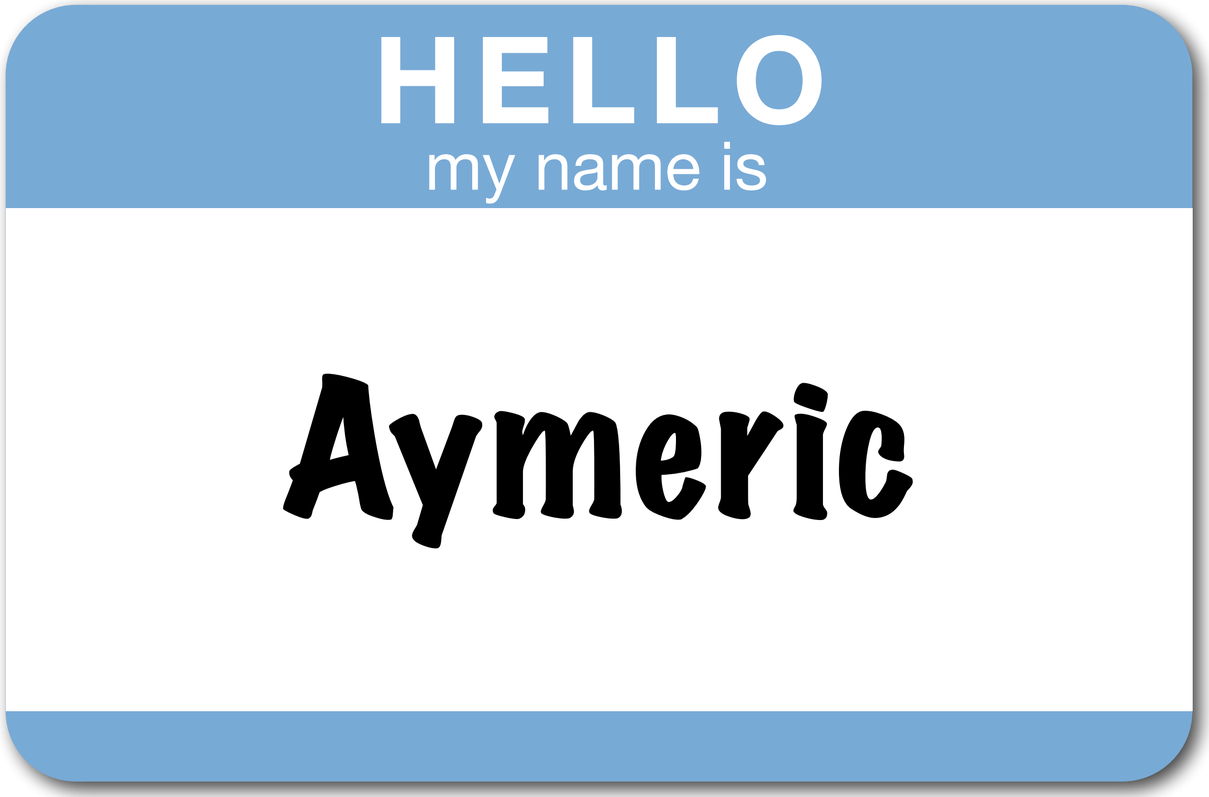
Welcome to the Tidyverse

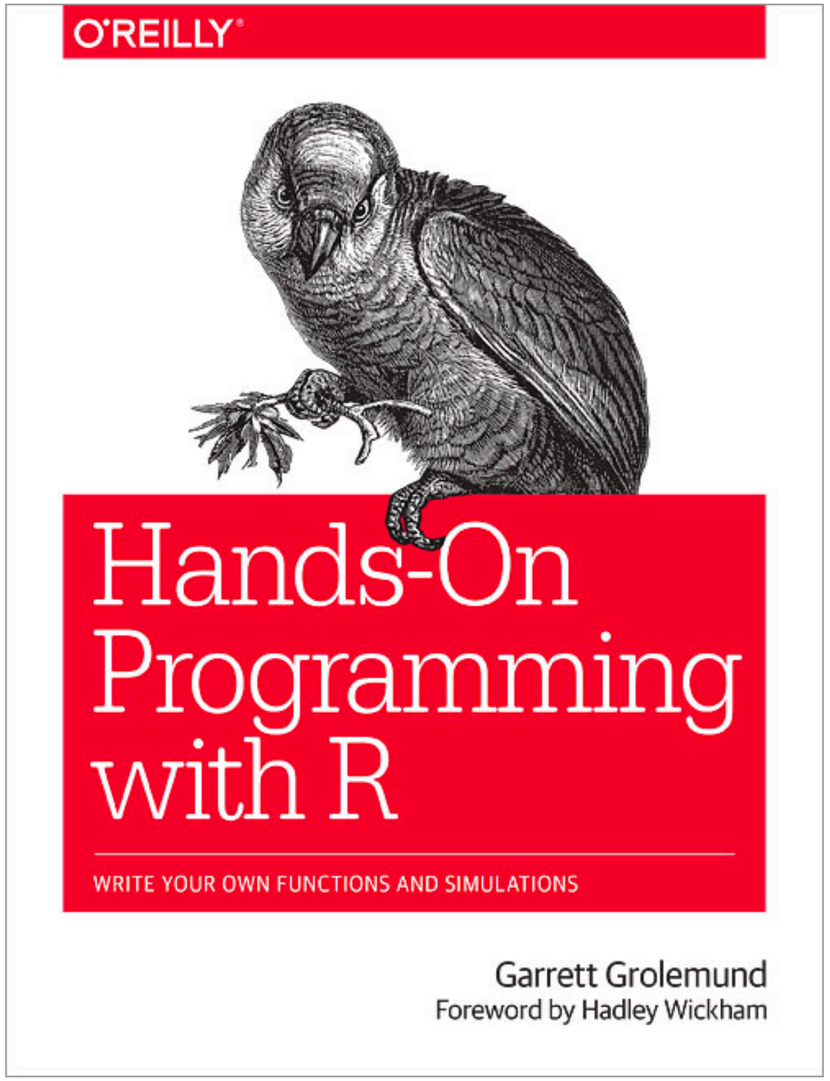


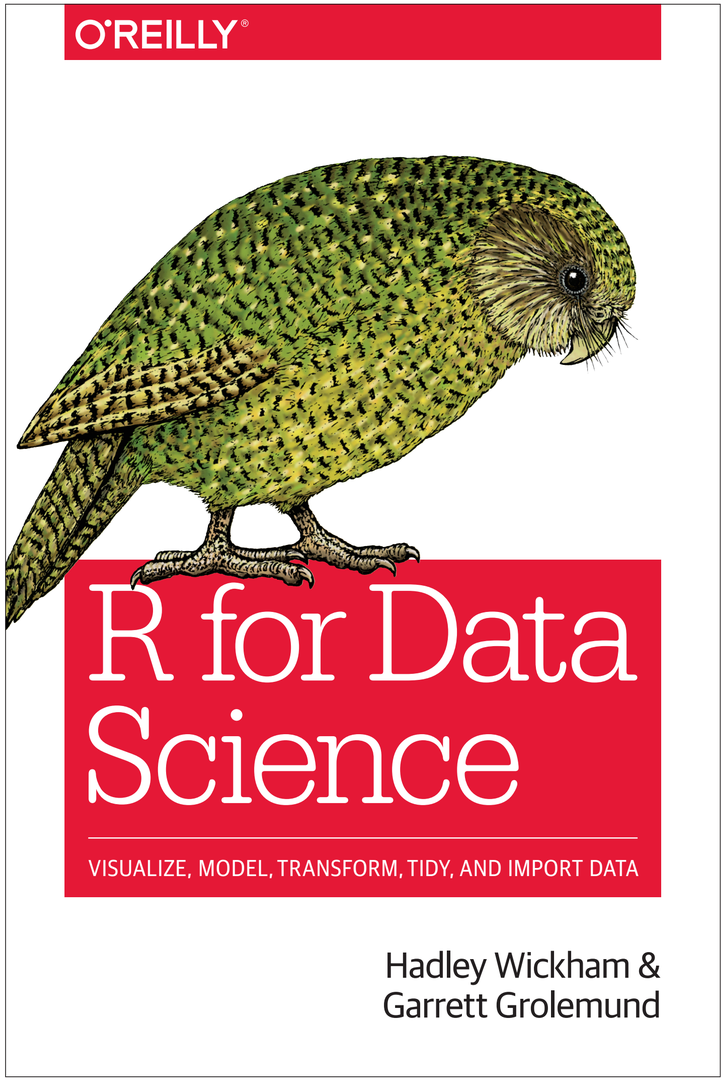
Slides CC BY-SA RStudio

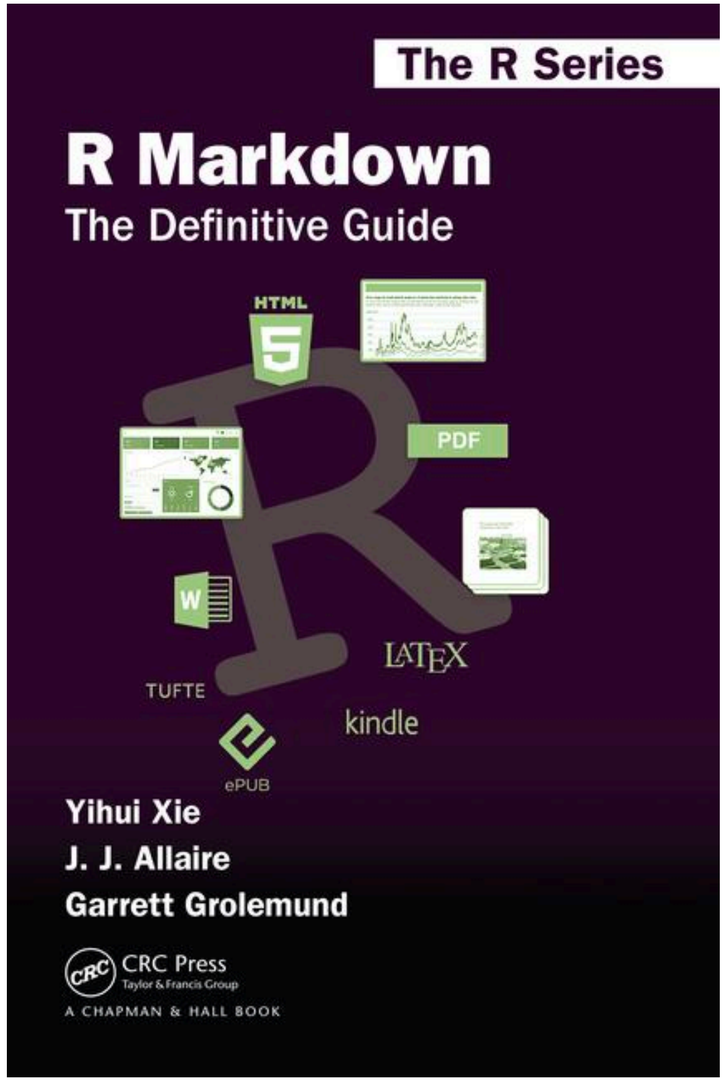


@aymeric-stamm

aymeric.stamm@cnrs.fr


O'REILLY
Hands-On Programming with R
WRITE YOUR OWN FUNCTIONS AND SIMULATIONS
Garrett Grolemund
Foreword by Hadley Wickham


O'REILLY
R for Data Science
VISUALIZE, MODEL, TRANSFORM, TIDY, AND IMPORT DATA
Hadley Wickham &
Garrett Grolemund


The R Series
R Markdown
The Definitive Guide
HTML
PDF
W
LATEX
TUFTE
kindle
ePUB
Yihui Xie
J. J. Allaire
Garrett Grolemund
CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Pop Quiz

What does **IMRAD** stand for? Poll your neighbors.

Introduction
What hypothesis was tested and why?

Methods
How was the study done?

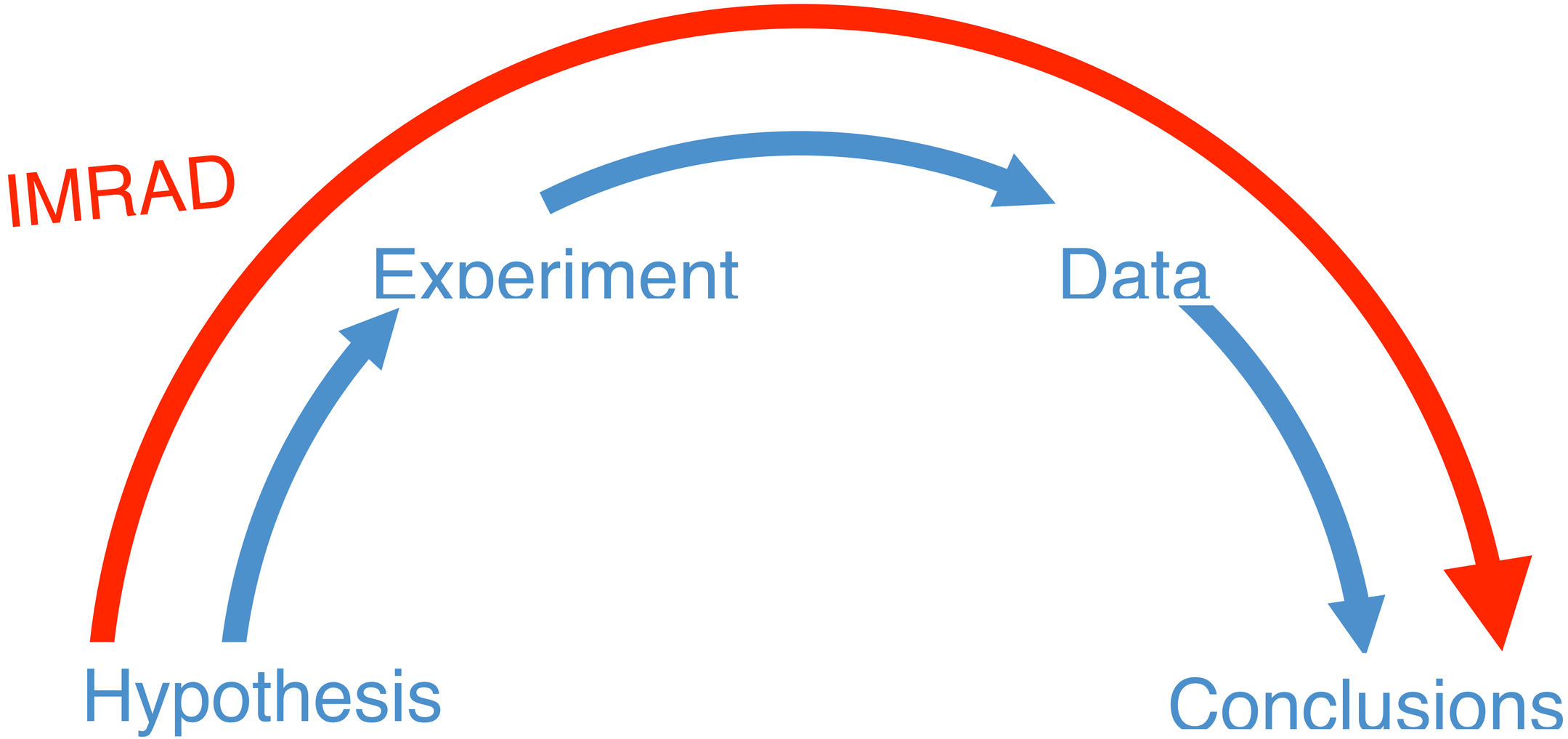
Results
What answer was discovered?

And Discussion
What does the answer imply?



CC BY-SA RStudio

(Experimental) **Science**



CC BY-SA RStudio

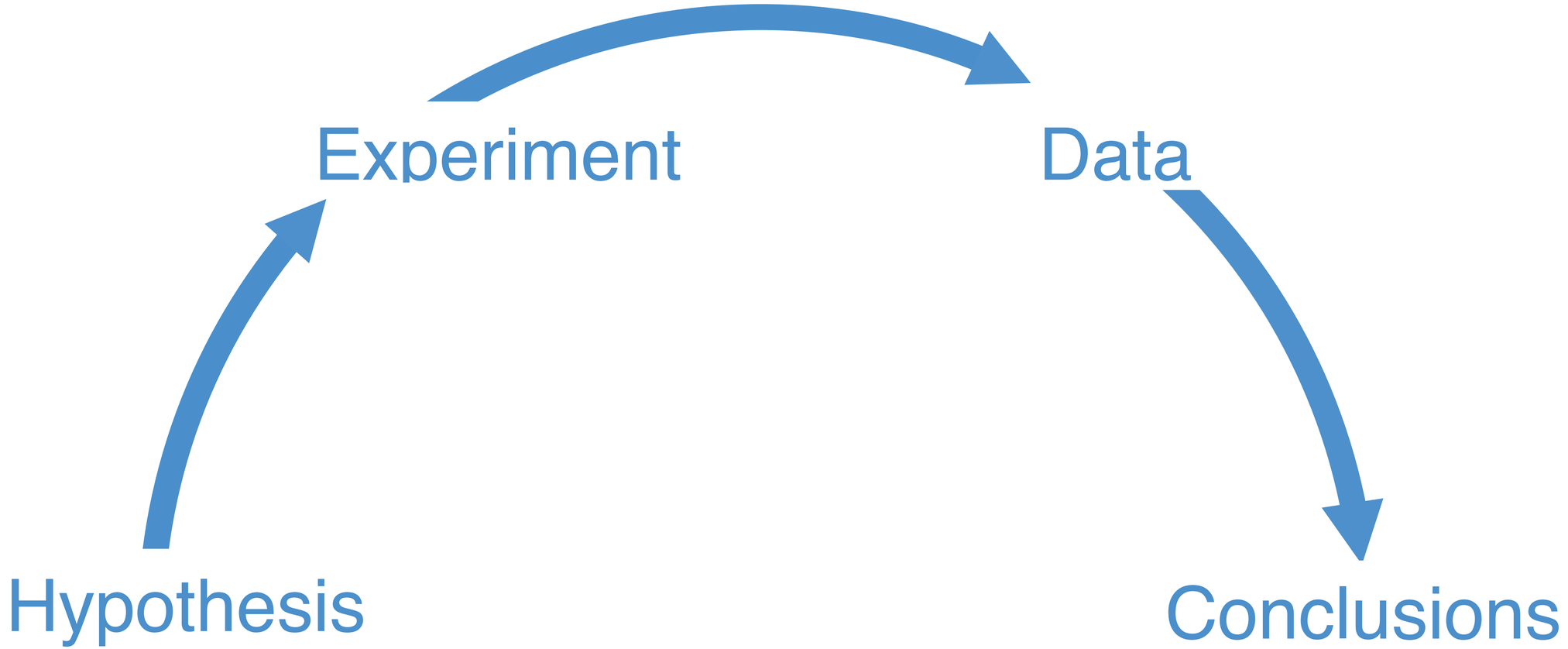
Pop Quiz

What is **data science**? Poll your neighbors.



CC BY-SA RStudio

Data Science



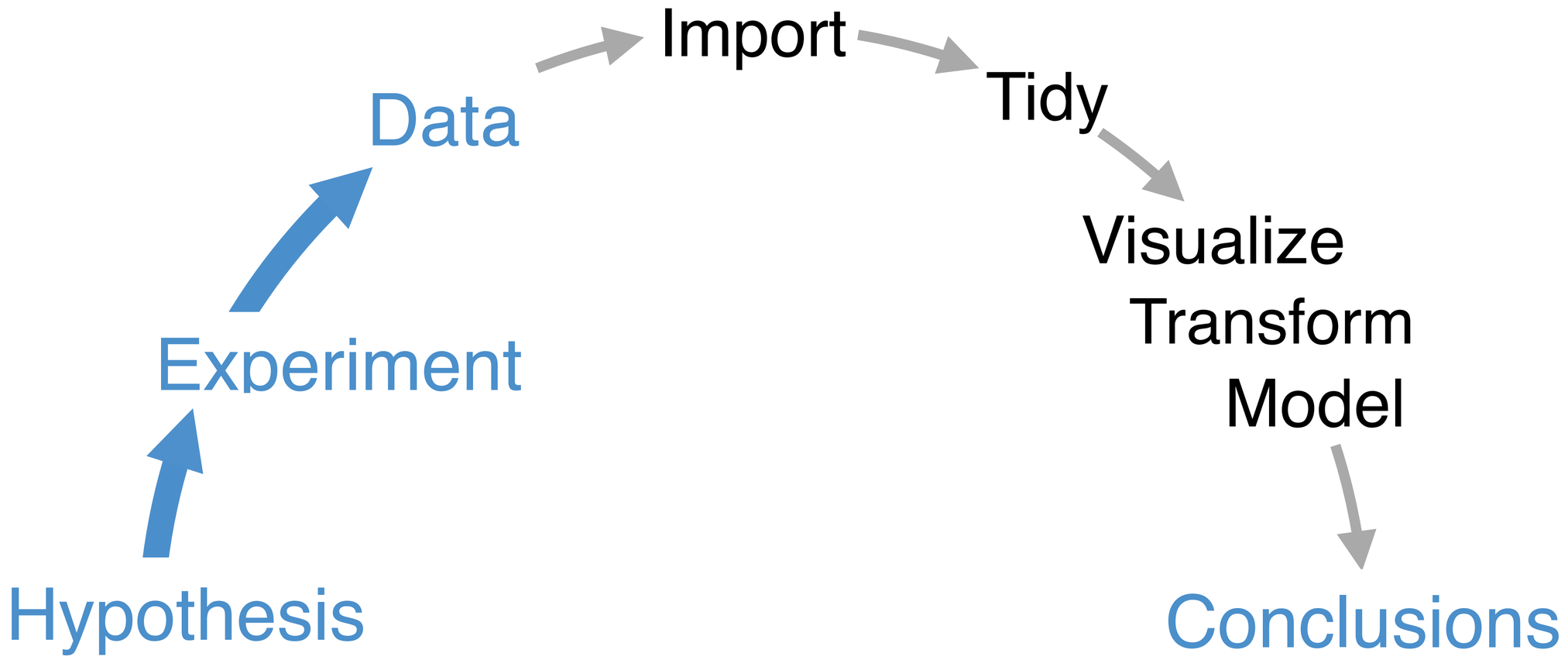
CC BY-SA RStudio

Data Science



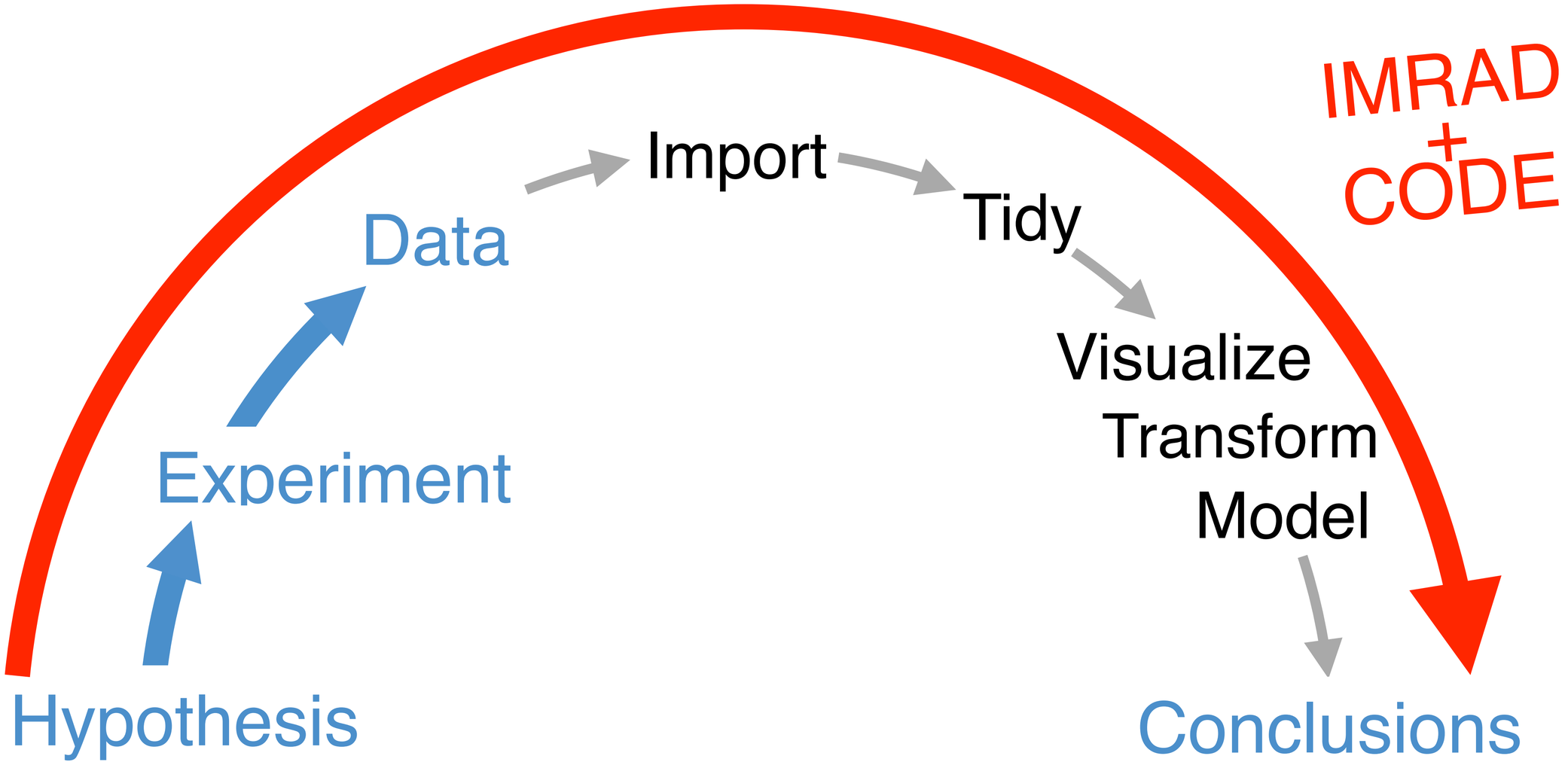
CC BY-SA RStudio

Data Science



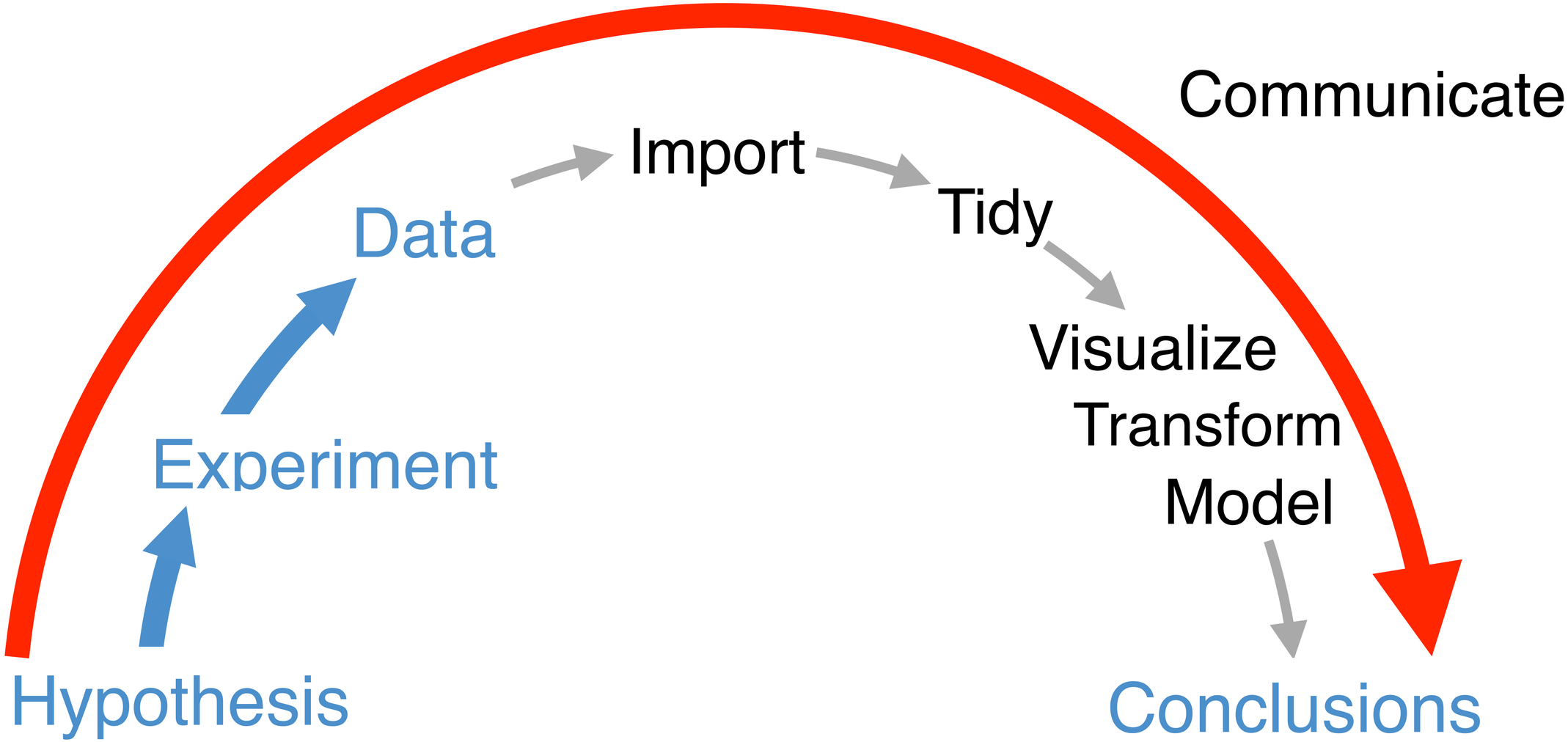
CC BY-SA RStudio

Data Science



CC BY-SA RStudio

Data Science



CC BY-SA RStudio

Data Science

Communicate

Import

Tidy

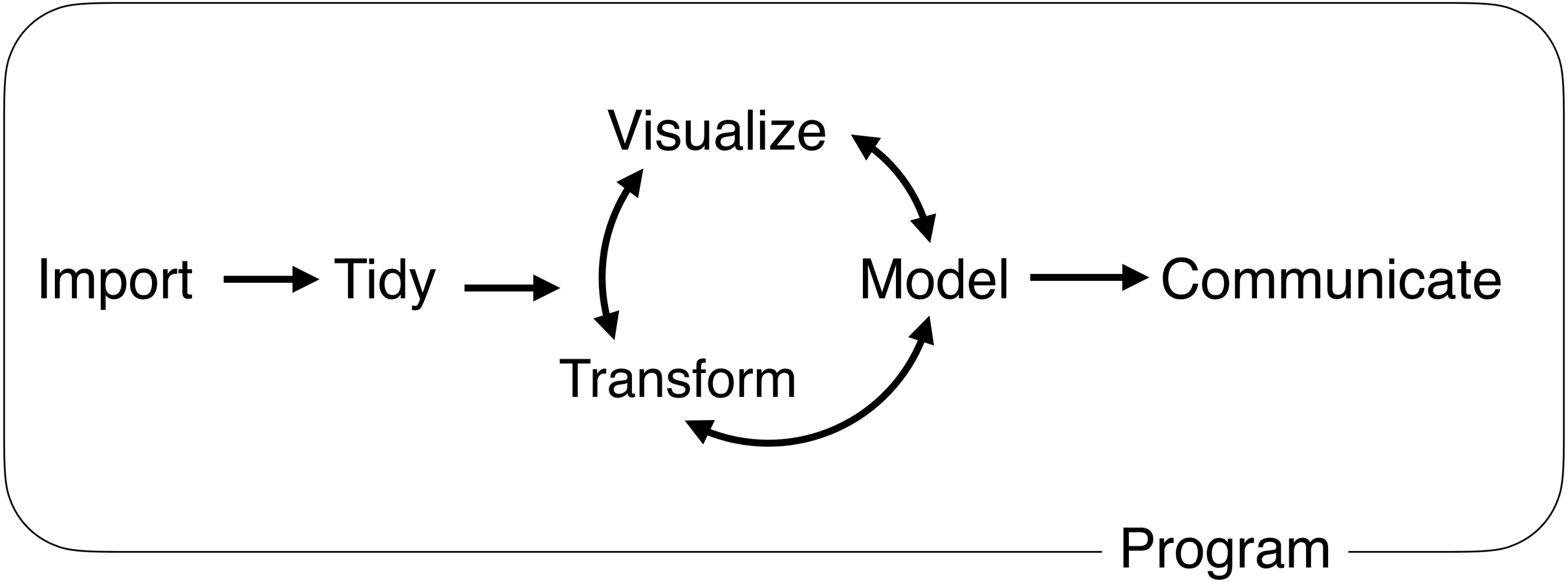
Visualize

Transform

Model

CC BY-SA RStudio

(Applied) Data Science







A language





A language



Writing software

Outline

Introduction

Visualize Data

Transform Data

Model Data

Communicate Data

Tidy Data

Join Data

Manipulate Data Types

Manipulate Lists

Your Turn

Go here for the class material

https://astamm.github.io/data-science-with-r/



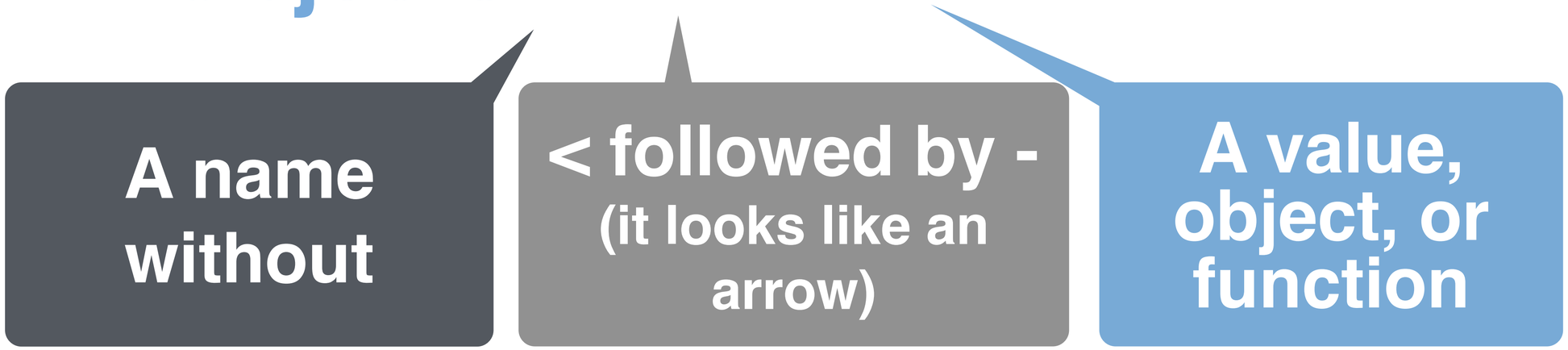
R the language

Values: 1, "Florida", "2010-01-25"

R the language

Values: 1, "Florida", "2010-01-25"

Objects: x <- 22/7



R the language

Values: 1, "Florida", "2010-01-25"

Objects: x <- c(22/7, 0.99, 3)

To put multiple values in an object, combine the values with

R the language

Values: 1, "Florida", "2010-01-25"

Objects: x <- c(22/7, 0.99, 3)

Functions: round(x, digits = 3)

A name without quotes

followed by () to run the

Arguments: values, objects, or function

Warm Up

Which of these are **numbers**?

1 **"1"** **"one"** **one**

Warm Up

Which of these are **numbers**?

1 **"1"** **"one"** **one**

number

Warm Up

Which of these are **numbers**?

1 **"1"** **"one"** **one**

number **words (strings)**

Warm Up

Which of these are **numbers**?

1	**"1"**	**one**
number	**words (strings)**	**object**

Warm Up

Which of these will work? Suppose **one <- 1**.

log(1) **log("1")** **log("one")** **log(one)**

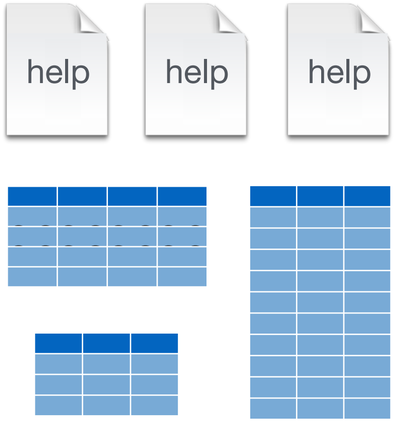
Warm Up

Which of these will work? Suppose **one <- 1**.

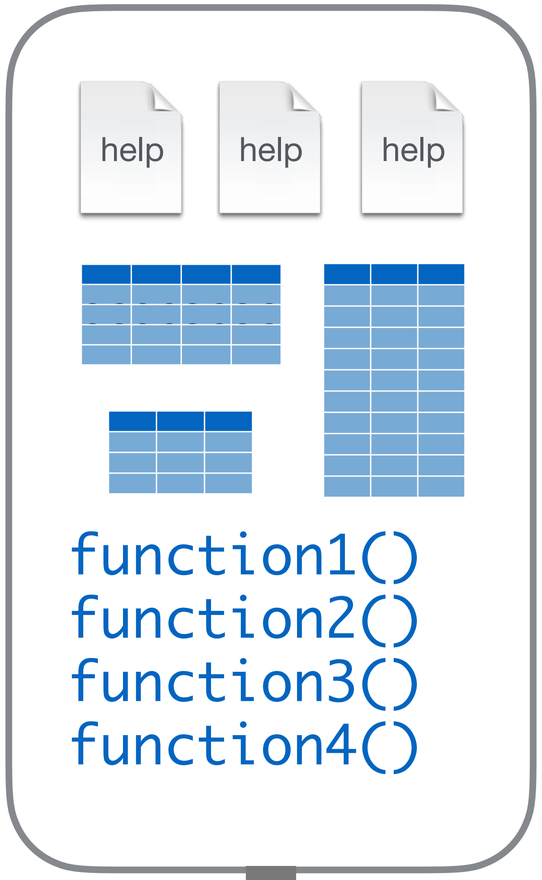
log(1) **log("1")** **log("one")** **log(one)**

R Packages

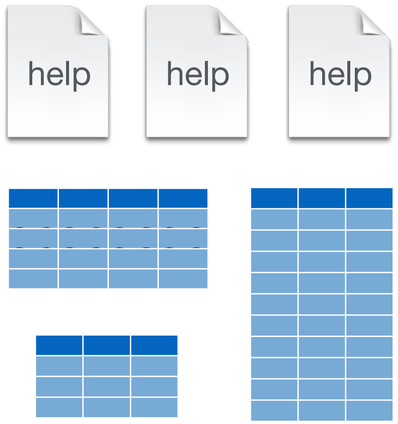




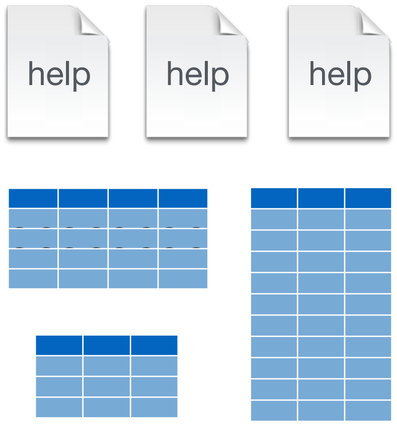

```
function1()
function2()
function3()
function4()
```



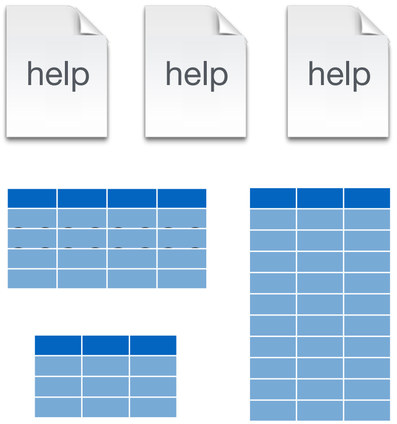

```
function1()
function2()
function3()
function4()
```



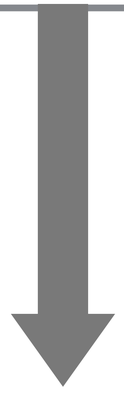

```
function5()
function6()
function7()
function8()
```



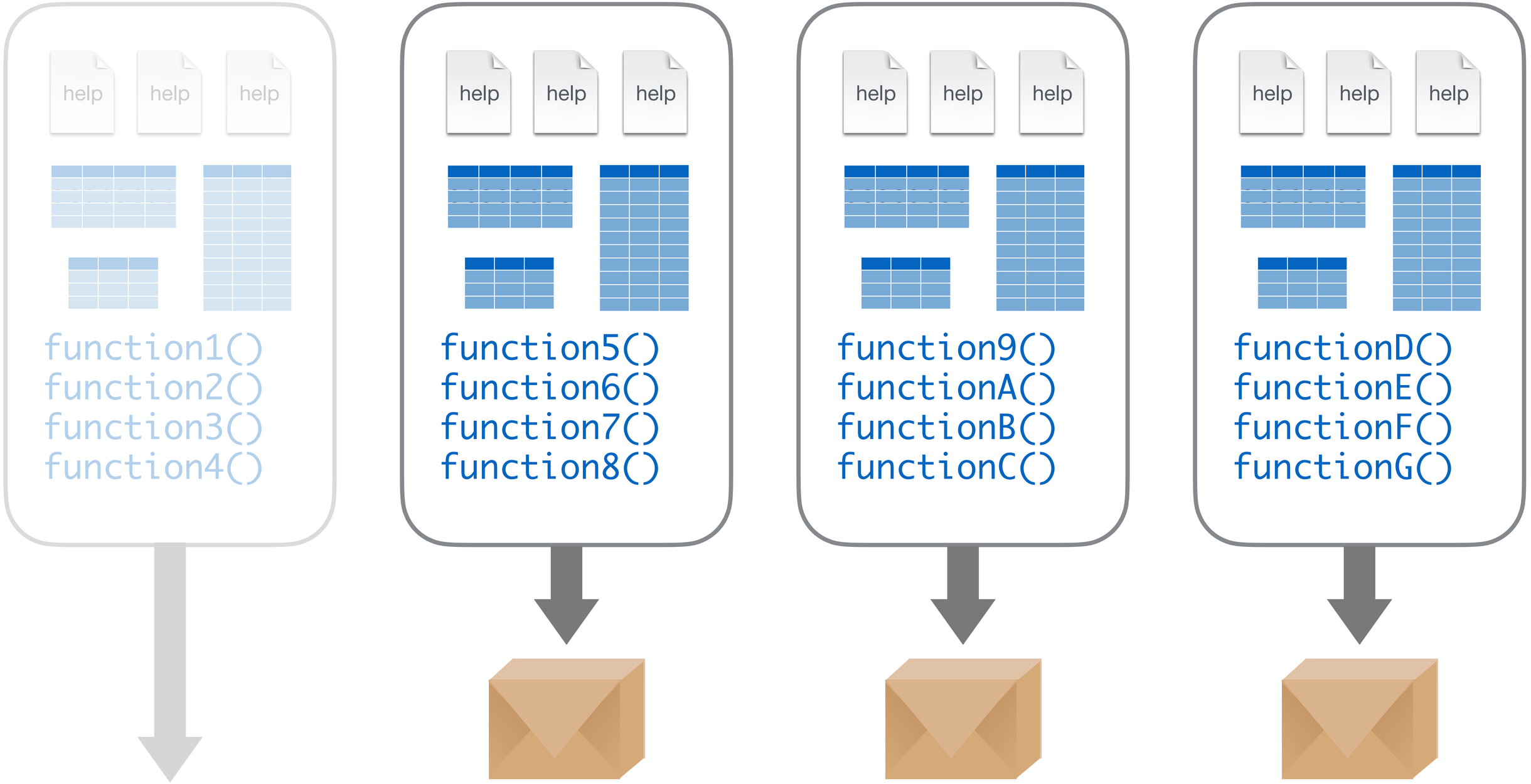

```
function9()
functionA()
functionB()
functionC()
```




```
functionD()
functionE()
functionF()
functionG()
```

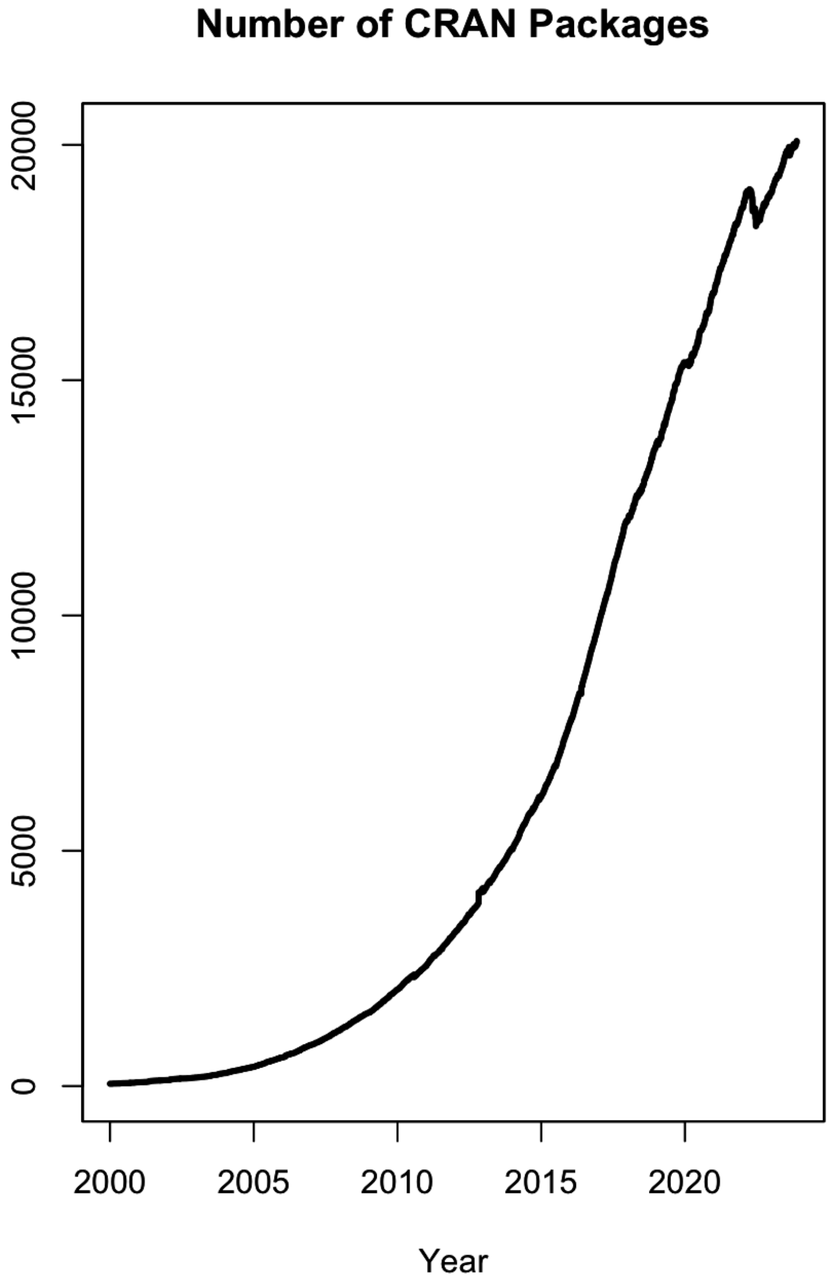


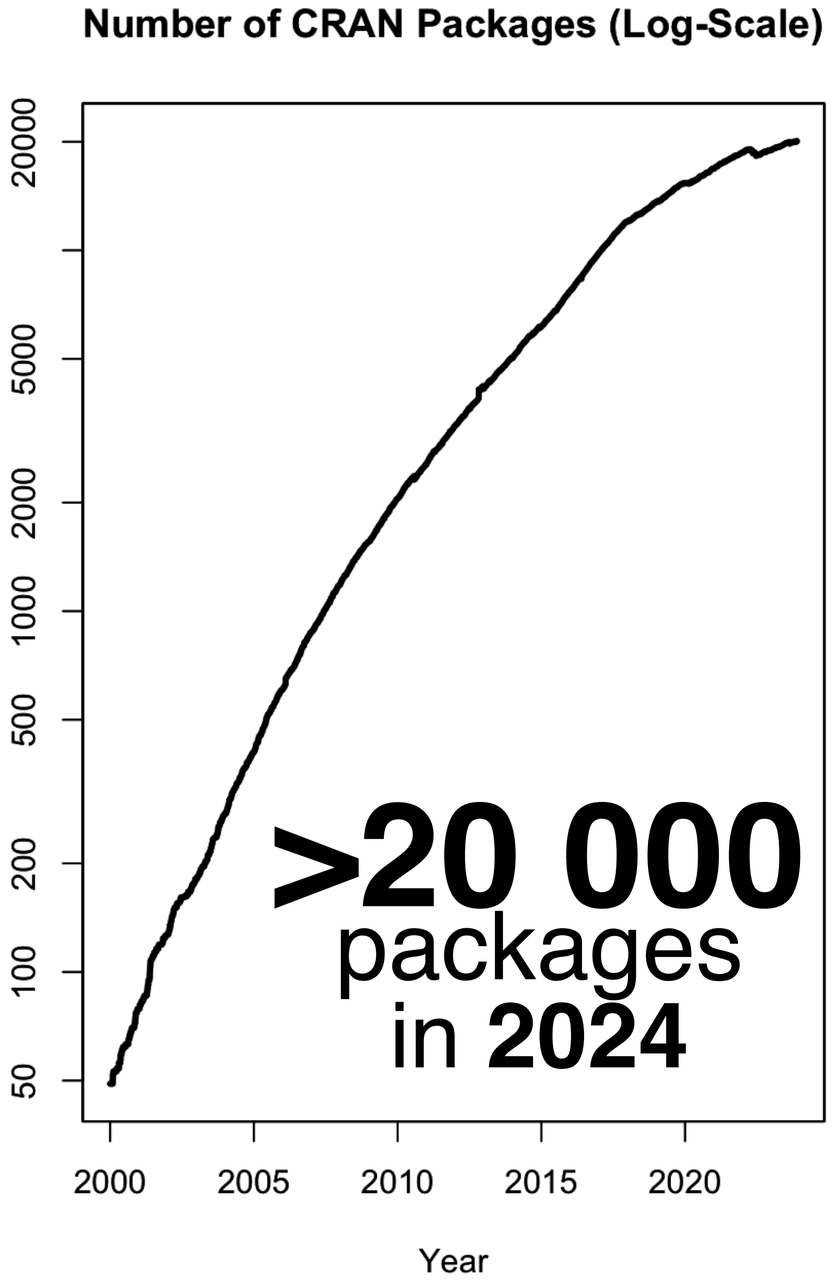
Base R



Base R

R Packages





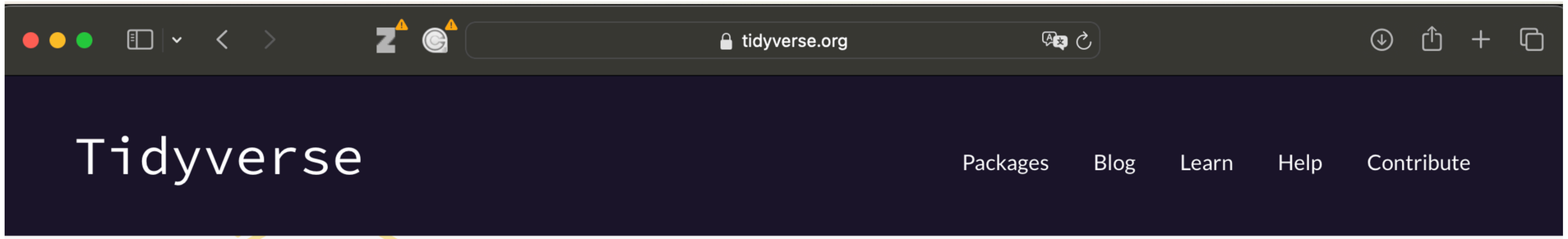
Graph by https://journal.r-project.org/news/RJ-2023-3-cran/RJ-2023-3-cran_files/figure-html5/cran_growth-1.p

cran.r-project.org

Available CRAN Packages By Name

A3	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models
AalenJohansen	Conditional Aalen-Johansen Estimation
AATtools	Reliability and Scoring Routines for the Approach-Avoidance Task
ABACUS	Apps Based Activities for Communicating and Understanding Statistics
abasequence	Coding 'ABA' Patterns for Sequence Data
abbreviate	Readable String Abbreviation
abc	Tools for Approximate Bayesian Computation (ABC)
abc.data	Data Only: Tools for Approximate Bayesian Computation (ABC)
ABC.RAP	Array Based CpG Region Analysis Pipeline
ABCanalysis	Computed ABC Analysis
abclass	Angle-Based Large-Margin Classifiers
ABCoptim	Implementation of Artificial Bee Colony (ABC) Optimization
ABCp2	Approximate Bayesian Computational Model for Estimating P2
abcrf	Approximate Bayesian Computation via Random Forests
abcrlda	Asymptotically Bias-Corrected Regularized Linear Discriminant Analysis
abctools	Tools for ABC Analyses

tidyverse.org





R packages for data science

The tidyverse is an opinionated collection of R packages designed for data science. All packages share an underlying design philosophy, grammar, and data structures.

Install the complete tidyverse with:

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

Using packages

1

```
install.packages("foo")
```

Downloads files to computer

1 x per computer

Pop Quiz!

The tidyverse contains the following packages. What function would you use to install them?

ggplot2
dplyr
tidyr
readr
purrr

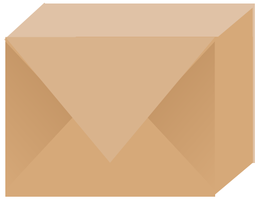
tibble
hms
stringr
lubridate
forcats

DBI
haven
httr
jsonlite
readxl

rvest
xml2
modelr
tidyverse

```
install.packages("ggplot2")
install.packages("dplyr")
install.packages("tidyr")
install.packages("readr")
install.packages("purrr")
install.packages("tibble")
install.packages("hms")
install.packages("stringr")
install.packages("lubridate")
install.packages("forcats")
install.packages("DBI")
install.packages("haven")
install.packages("httr")
install.packages("jsonlite")
install.packages("readxl")
install.packages("rvest")
install.packages("xml2")
install.packages("modelr")
install.packages("broom")
```

tidyverse



An R package that serves as a short cut for installing and loading the components of the tidyverse.

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

A shortcut!

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

does the equivalent of

```
install.packages("ggplot2")
install.packages("dplyr")
install.packages("tidyr")
install.packages("readr")
install.packages("purrr")
install.packages("tibble")
install.packages("hms")
install.packages("stringr")
install.packages("lubridate")
install.packages("forcats")
install.packages("DBI")
install.packages("haven")
install.packages("httr")
install.packages("jsonlite")
install.packages("readxl")
install.packages("rvest")
install.packages("xml2")
install.packages("modelr")
install.packages("broom")
```

Using packages

1

```
install.packages("foo")
```

Downloads files to computer

1 x per computer

2

```
library("foo")
```

Loads package

1 x per R Session

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

does the equivalent of

```
install.packages("ggplot2")
install.packages("dplyr")
install.packages("tidyr")
install.packages("readr")
install.packages("purrr")
install.packages("tibble")
install.packages("stringr")
install.packages("forcats")
install.packages("lubridate")
install.packages("hms")
install.packages("DBI")
install.packages("haven")
install.packages("httr")
install.packages("jsonlite")
install.packages("readxl")
install.packages("rvest")
install.packages("xml2")
install.packages("modelr")
install.packages("broom")
```

```
library("tidyverse")
```

does the equivalent of

```
library("ggplot2")
library("dplyr")
library("tidyr")
library("readr")
library("purrr")
library("tibble")
library("stringr")
library("forcats")
```

Quarto

Your Turn

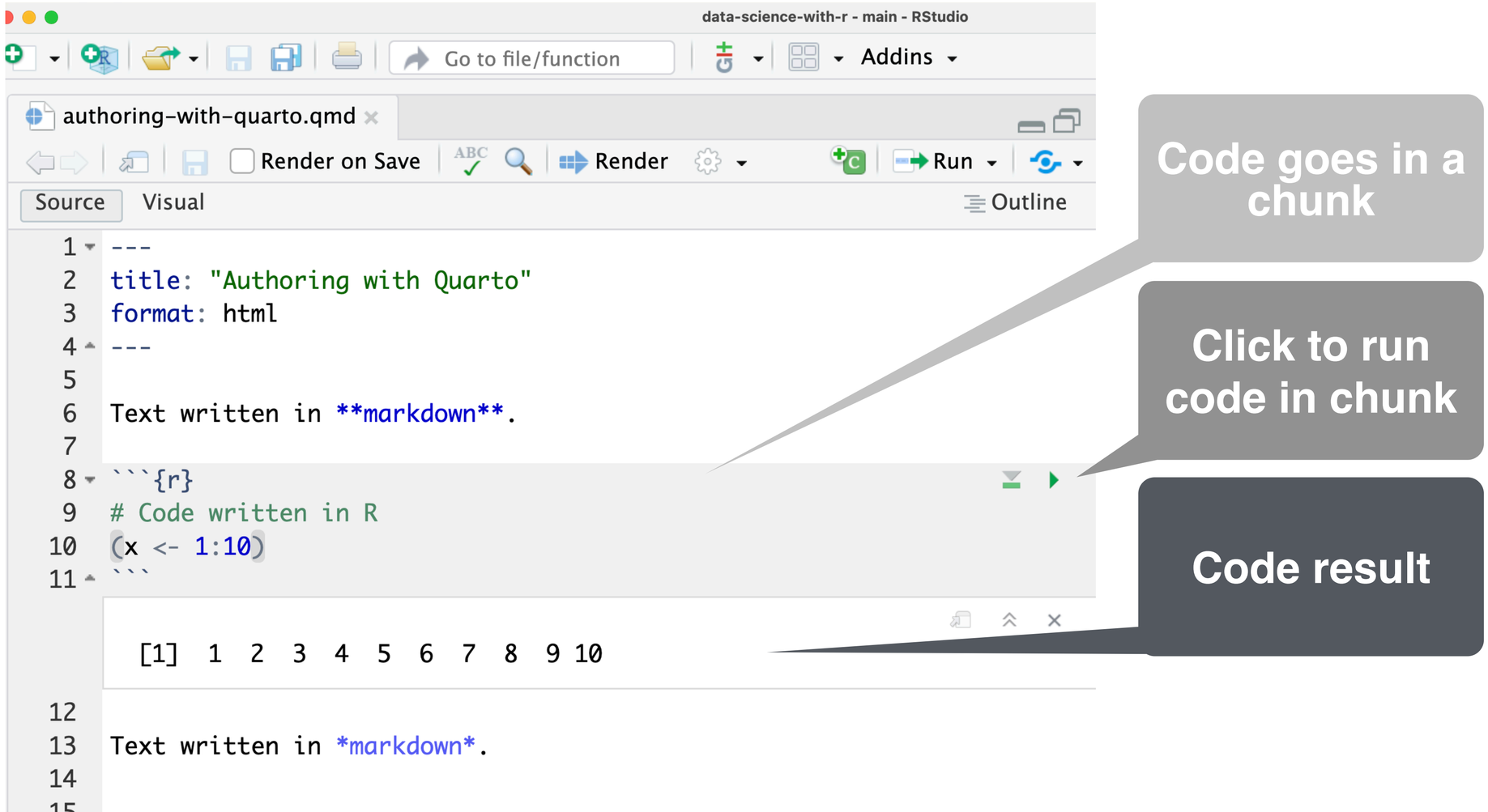
Navigate to the main page of the class: **https://astamm.github.io/data-science-with-r/**.

Download **01-Introduction-Exercises.qmd** from the outline table and open it. Read through the file and do everything it tells you to do.



Quarto

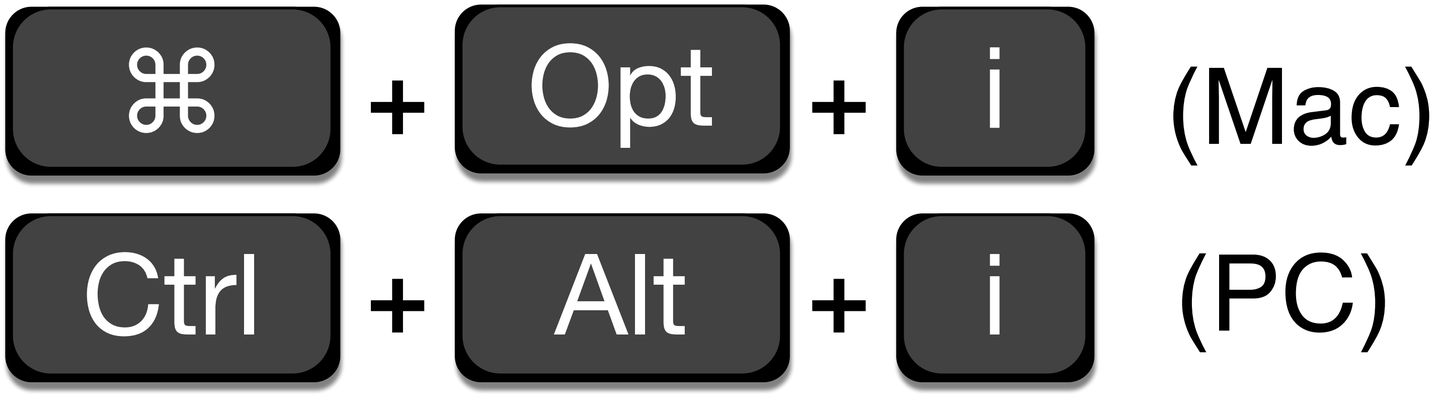
An authoring format for Data Science.



Code chunks

Insert a chunk of R code with

```
```{r}
# some code
```
```

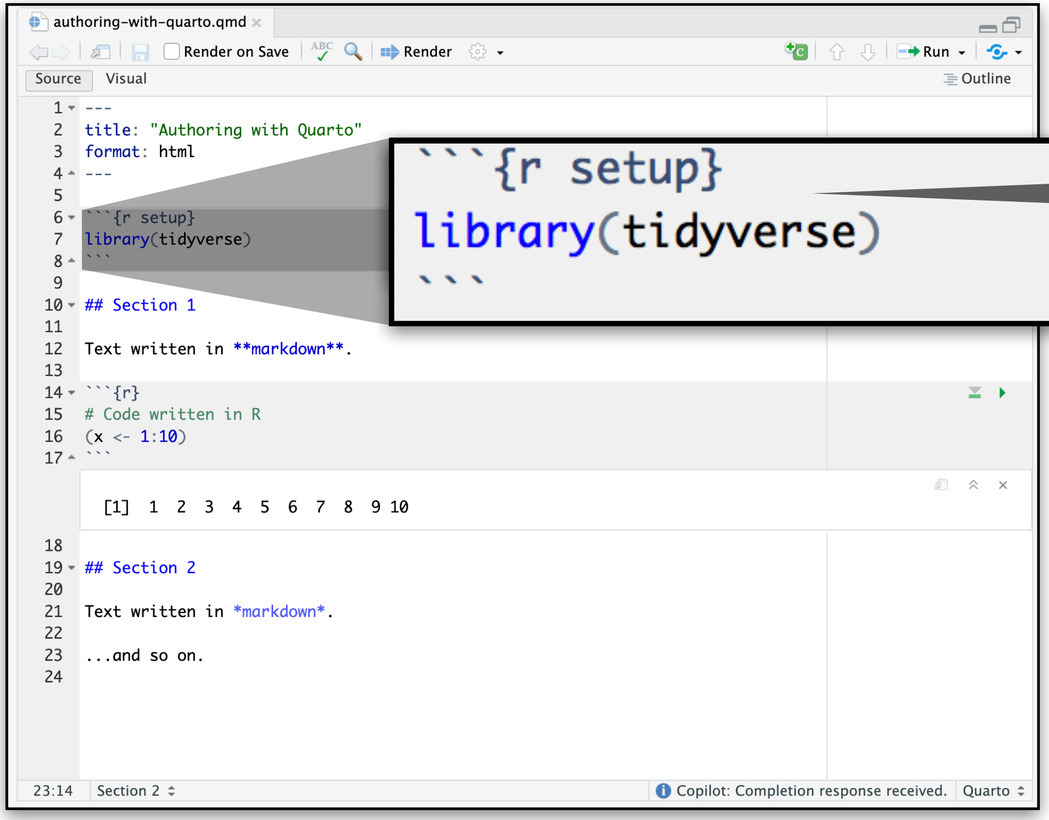


CC BY-SA RStudio

quarto®

Setup

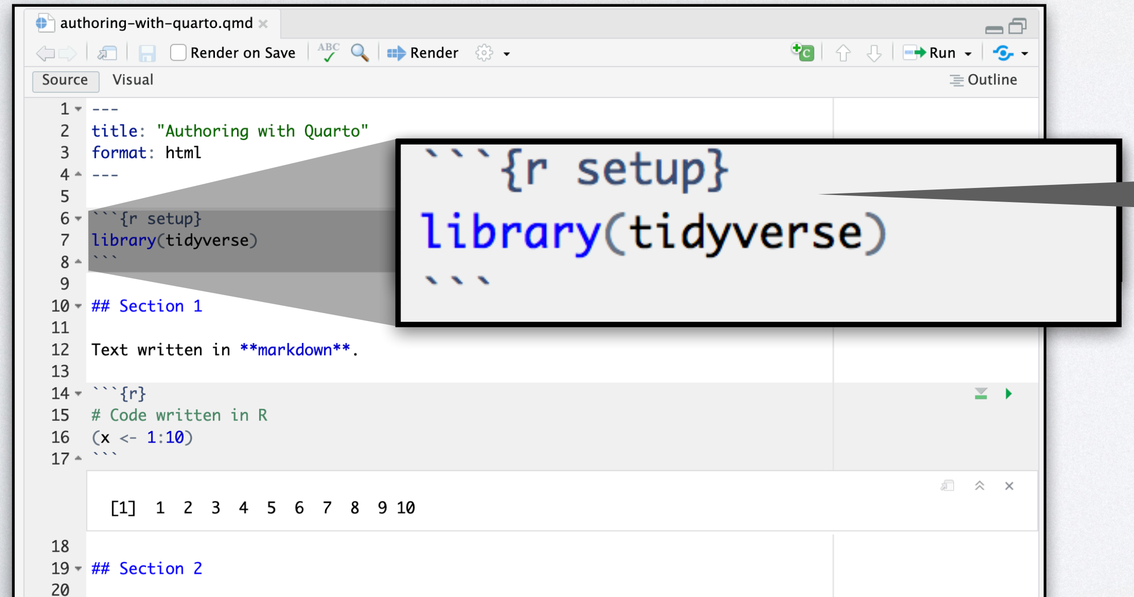
The setup chunk is always run once before anything else



chunk labels are optional, the setup label is special

Your Turn

Add a setup chunk to the top of **01-Introduction-Exercises.qmd**. Use it to load the tidyverse package, then uncomment and run the ggplot2 chunk at the bottom of your file.



chunk labels are optional, the setup label is special





Your Turn

Navigate to the main page of the class: **https://astamm.github.io/data-science-with-r/**.

Download **02-Visualize-Exercises.qmd** from the outline table and open it.

Add a setup chunk that loads the tidyverse packages.

