Welcome to the Tidyverse

Slides CC BY-SARStudio

Outline

9/3	Introduction and Visualize Data	2:00 - 3:00
9/5	Transform Data	2:00 - 3:00
9/10	Model Data	2:00 - 3:00
9/12	Import Data	2:00 - 3:00
9/17	Reproducible Reporting	2:00 – 2300
9/19	Versioning & Maps	2:00 - 3:00

Pop Quiz

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

Introduction
What hypothesis was tested and why?

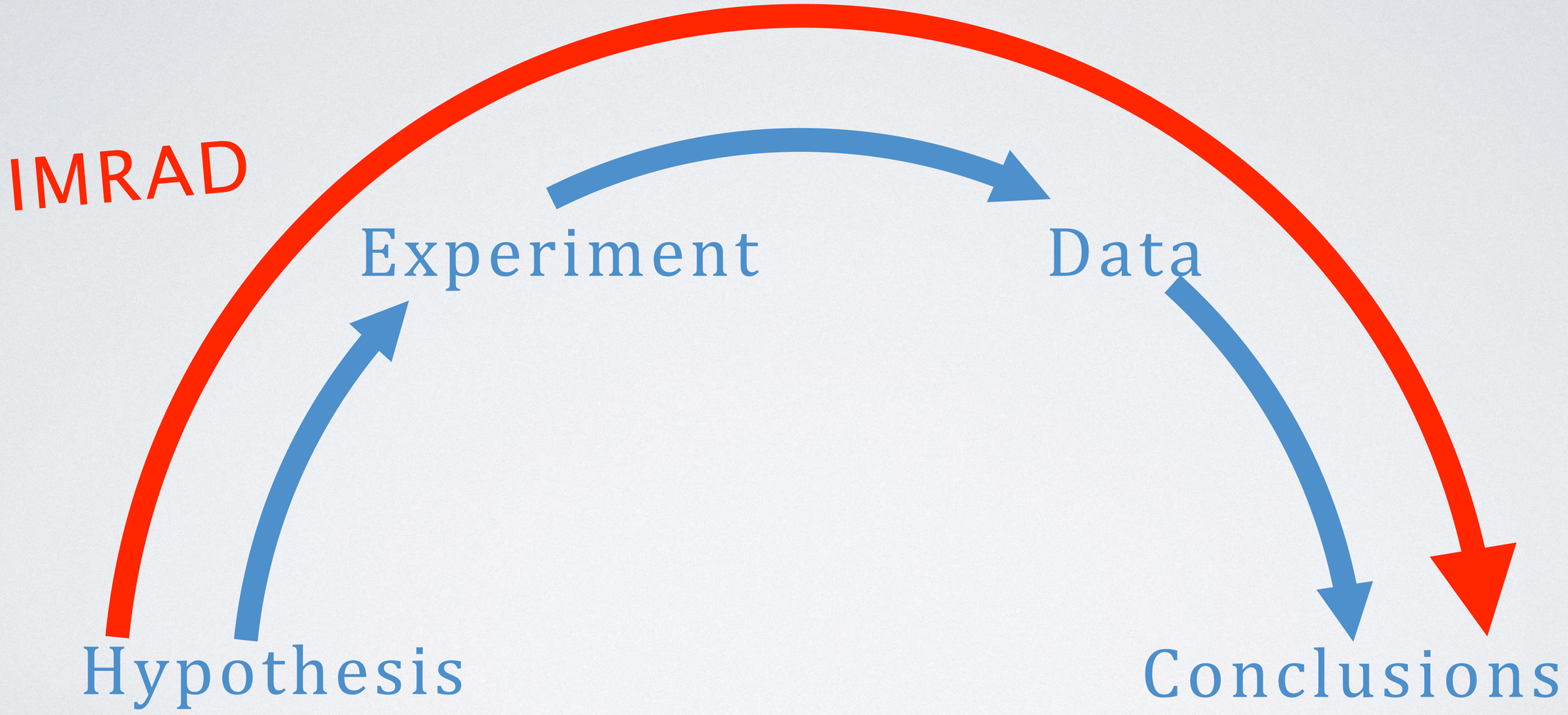
Methods
How was the study done?

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

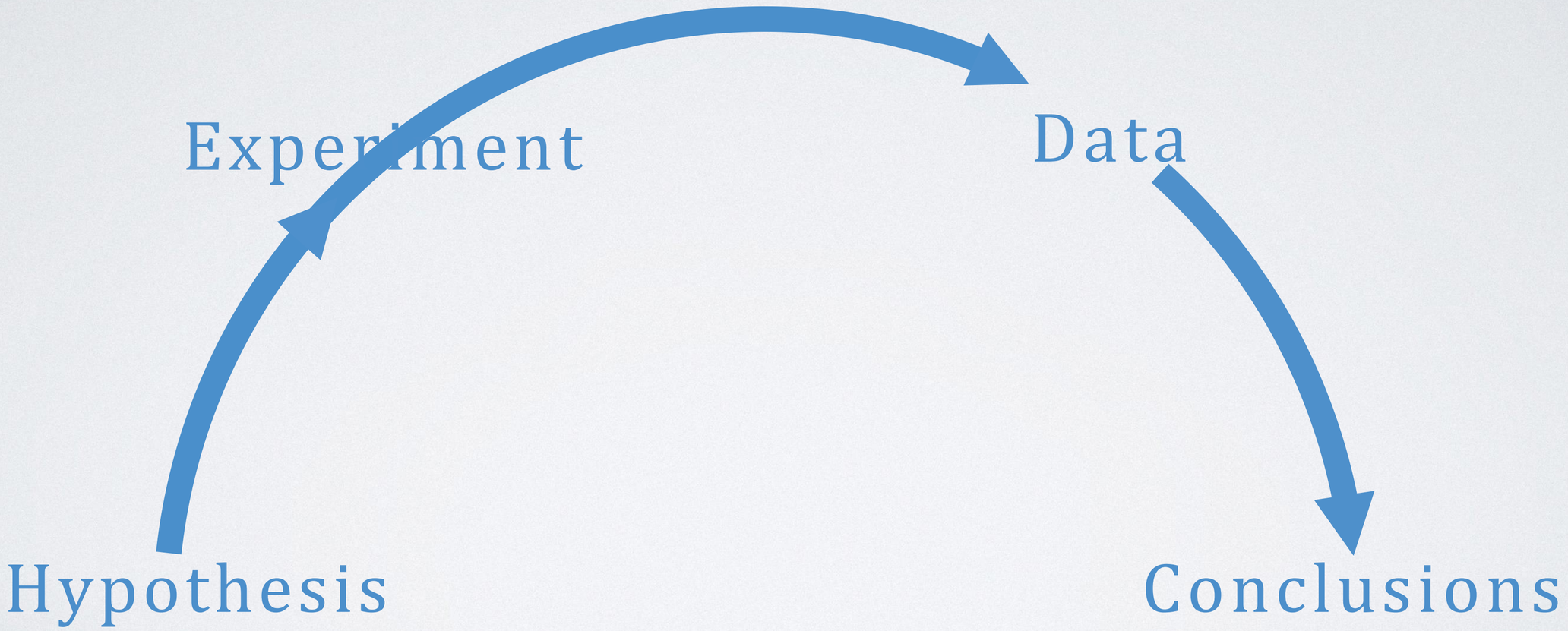
Pop Quiz

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

Science with non-trivial data?

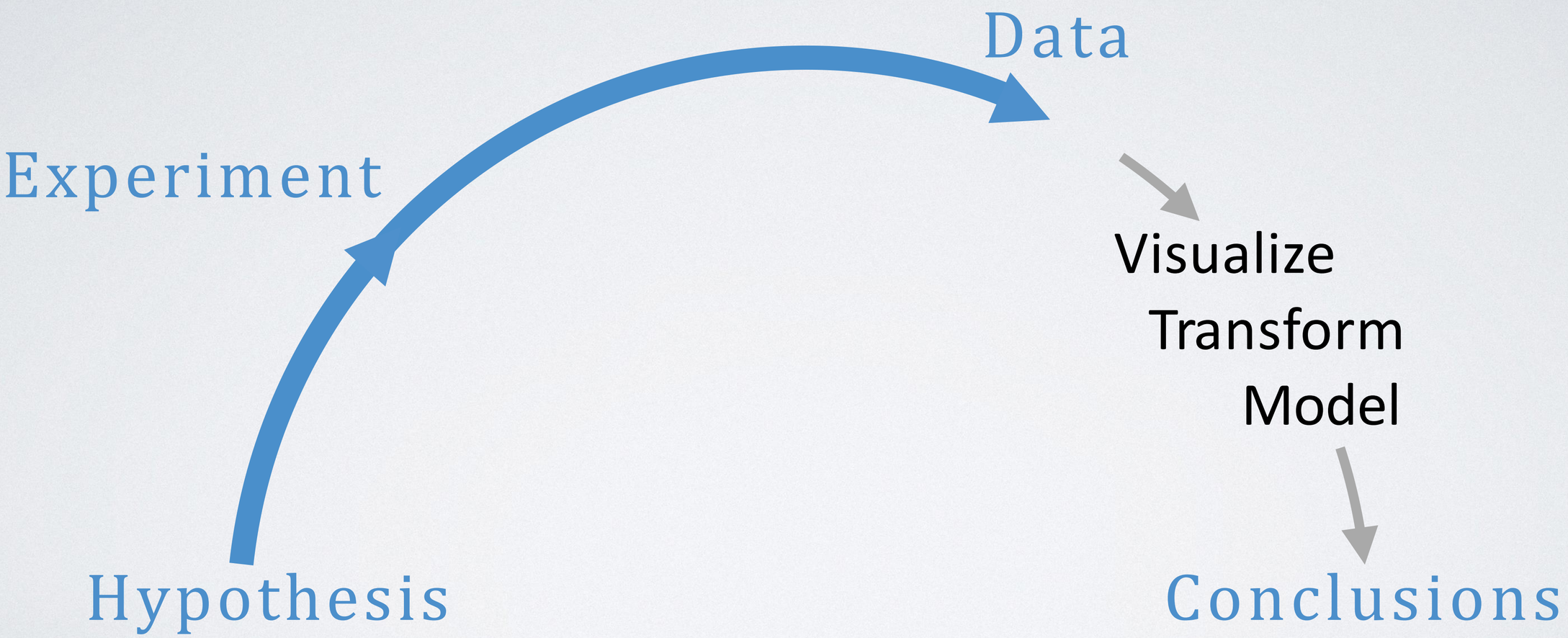
CC BY-SA RStudio

Data Science



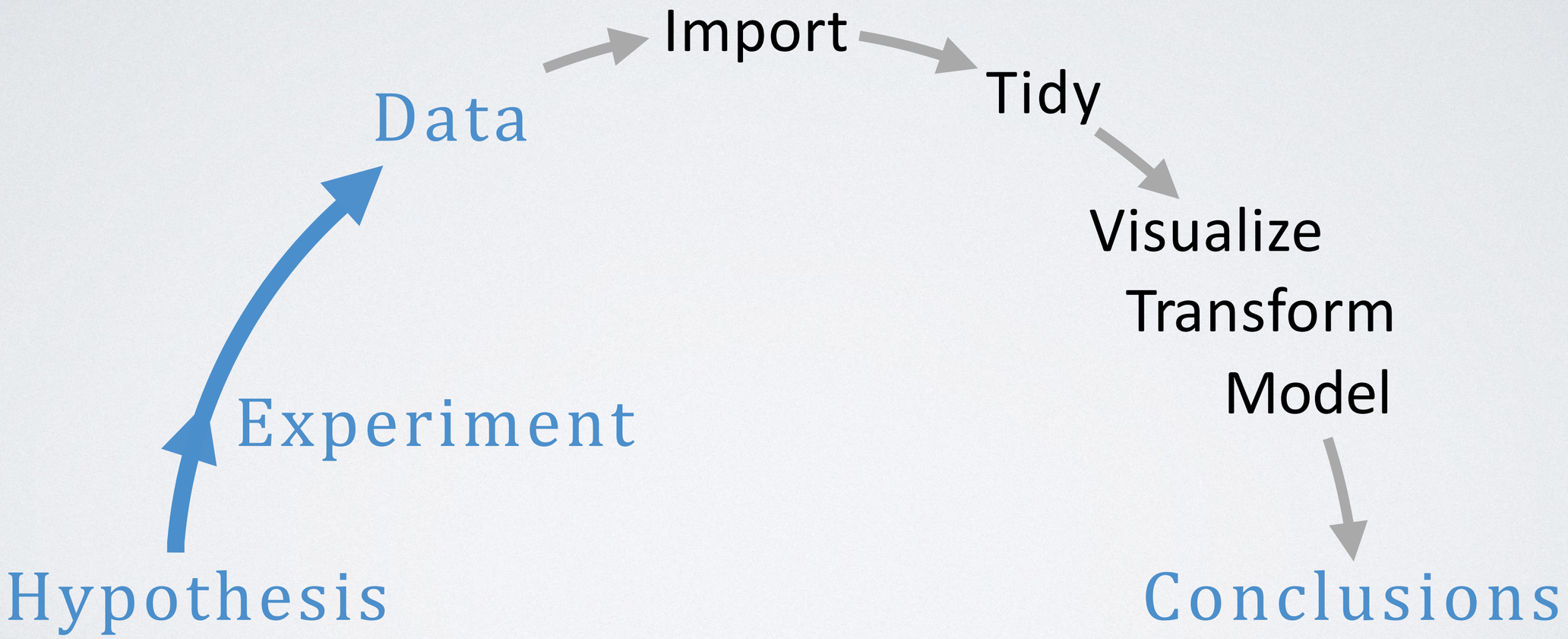
CC BY-SA RStudio

Data Science



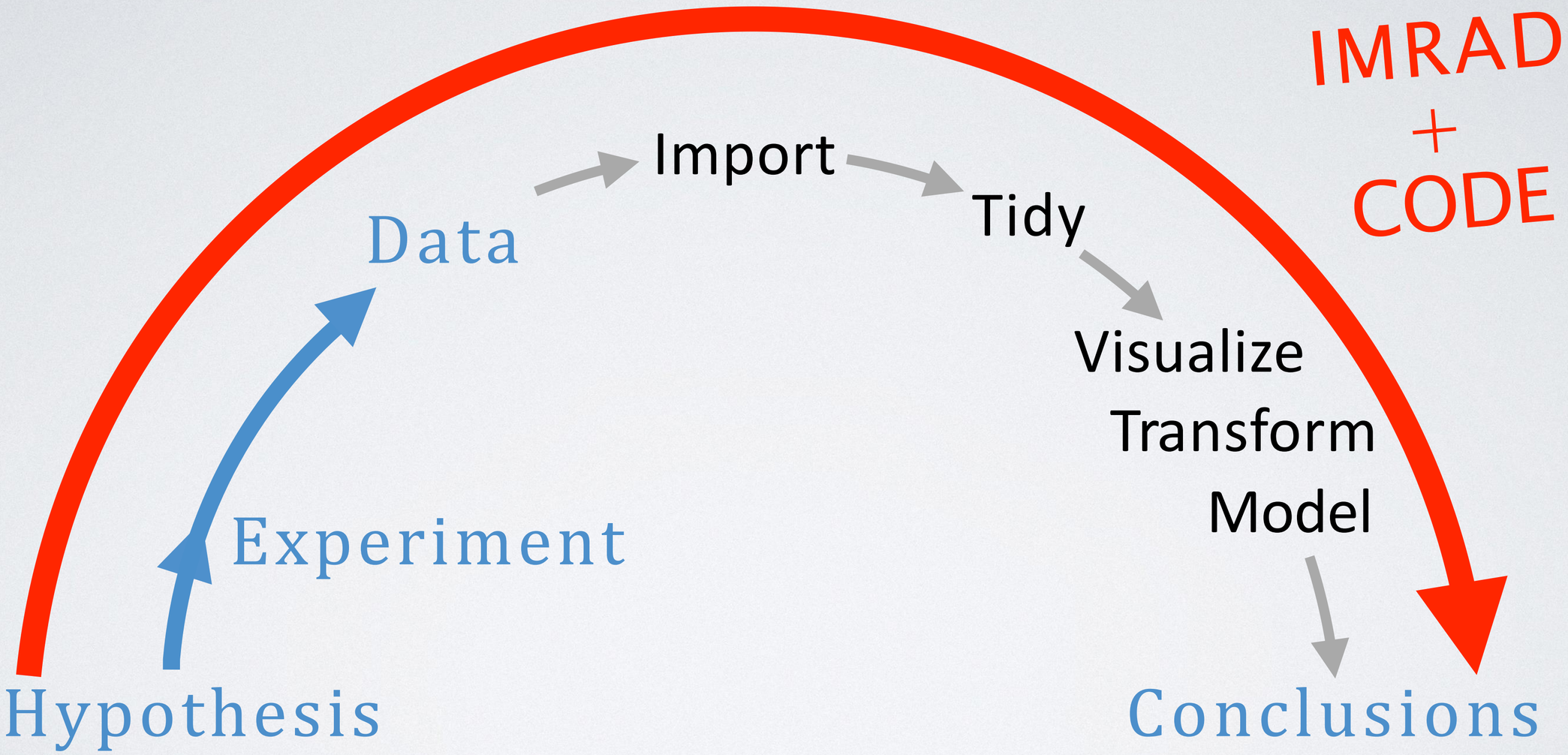
CC BY-SA RStudio

Data Science



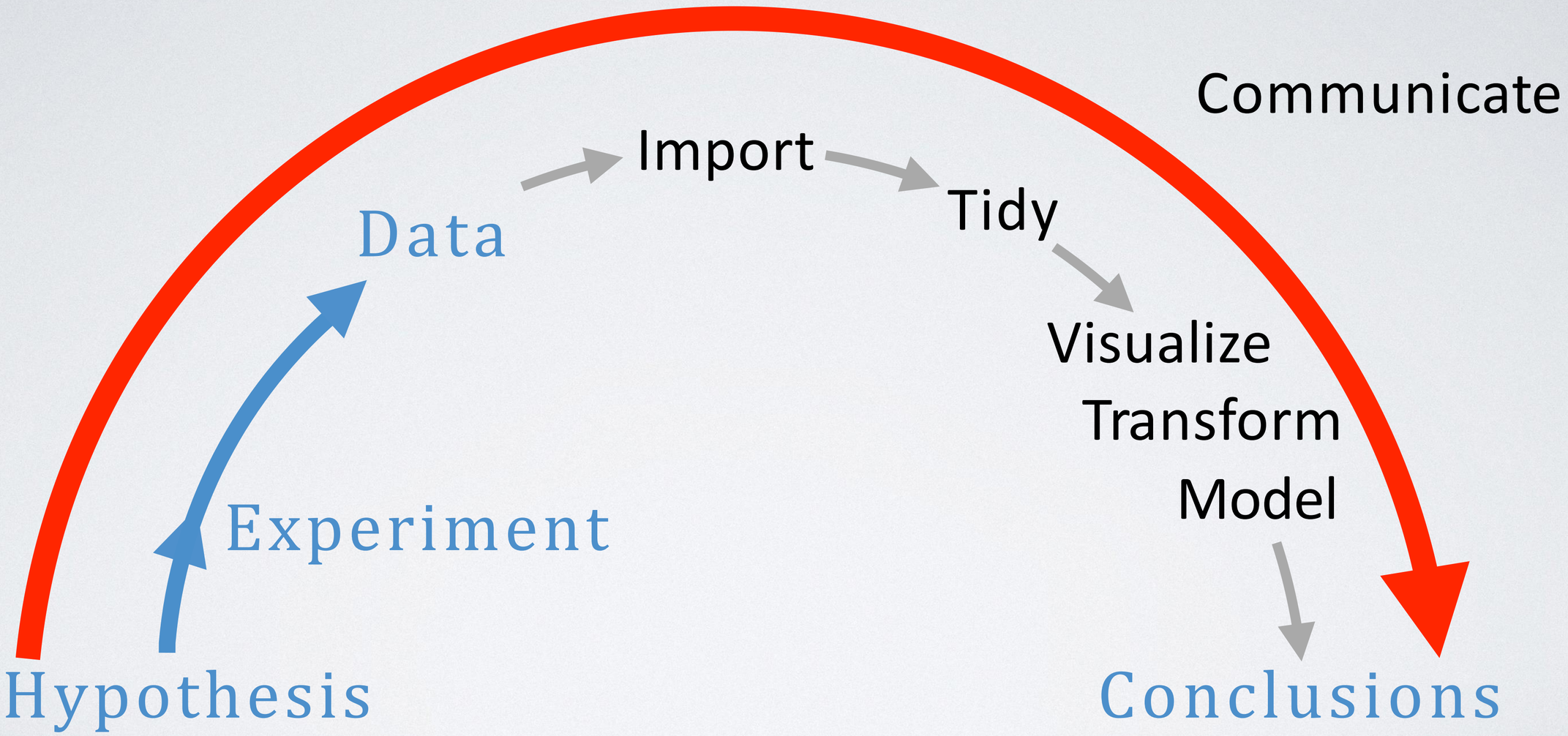
CC BY-SA RStudio

Data Science



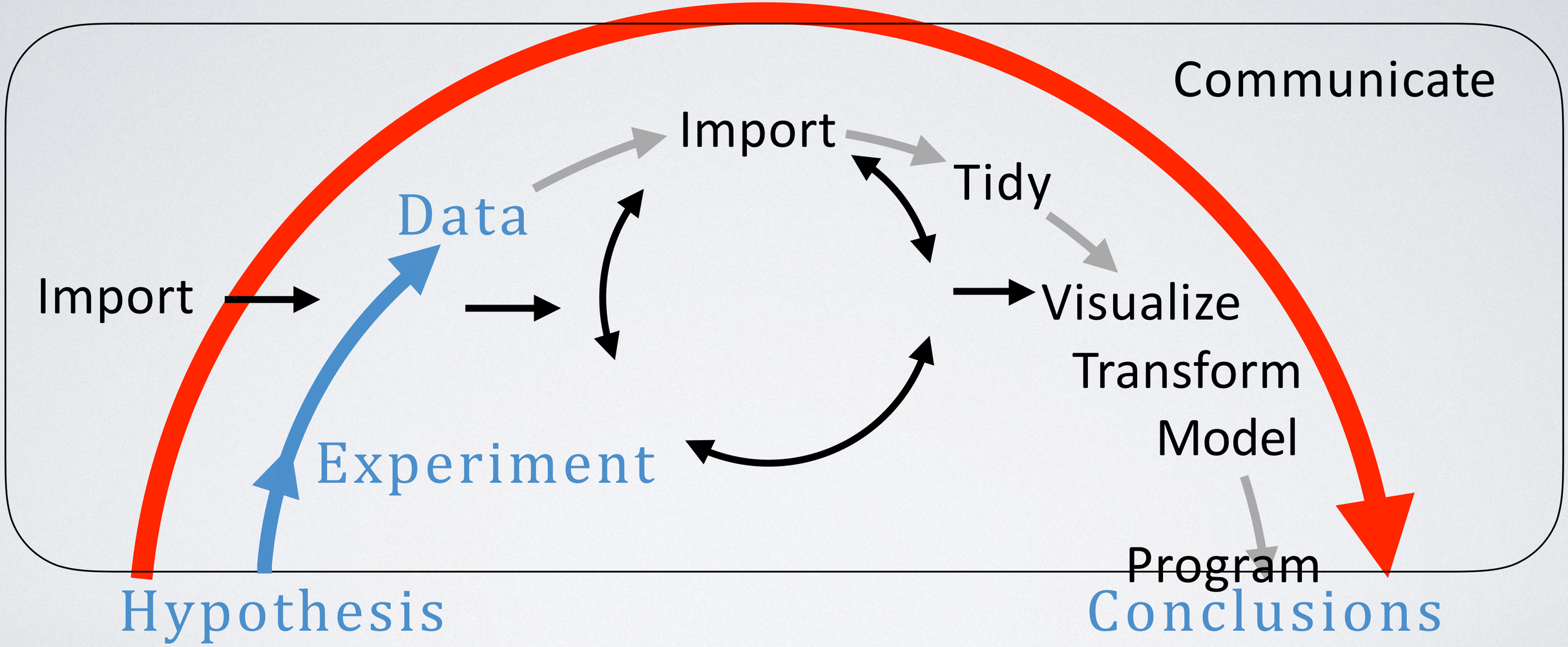
CC BY-SA RStudio

Data Science



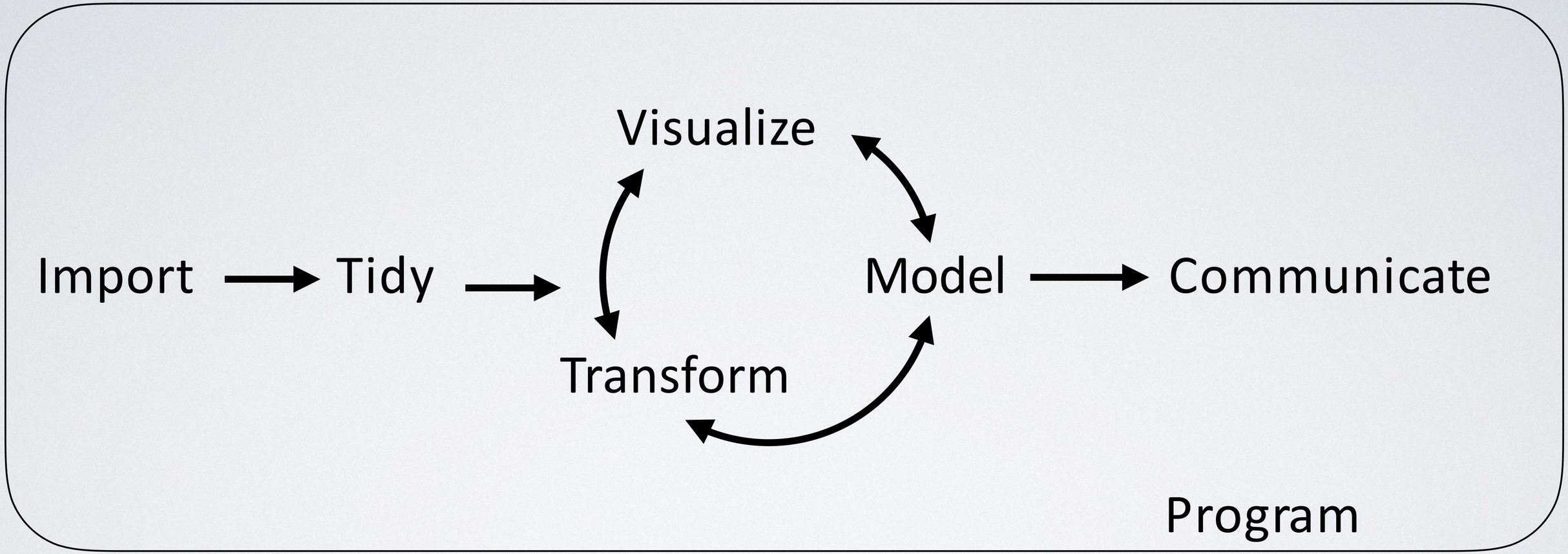
CC BY-SA RStudio

Data Science



CC BY-SA RStudio

(Applied) Data Science



A language

A language

Writing software

R the language

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

R the language

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

Objects - x <- 22/7

A name without quotes

< followed by - (it looks like an arrow)

A value, object, or function result

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 c()

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 function

Arguments: values, objects, or function results

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"	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)**

Other Data Types

- Datetime (POSIXct, date, etc)
- Factor
- Numeric (real, decimal, complex/scientific)
- Integer
- Logical (boolean)
- Vectors (lists)
  - Ie: `c(1, 2, 3, 4)` or `1:4`

R Function Syntax

R Functions

Functions- round(x, digits = 3)

A name without quotes

followed by () to run the function

Arguments: values, objects, or function results

Base R Functions

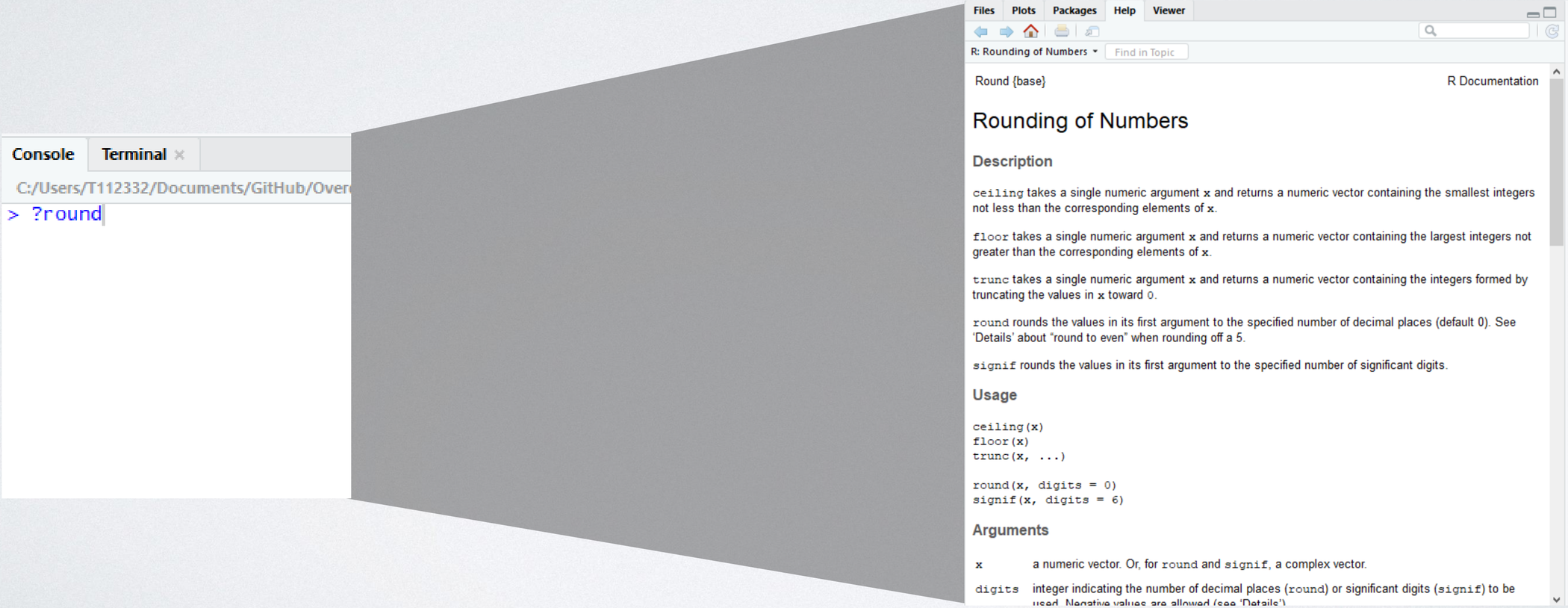
```
rnd <- round(x, digits = 3)
```

Define a stored value

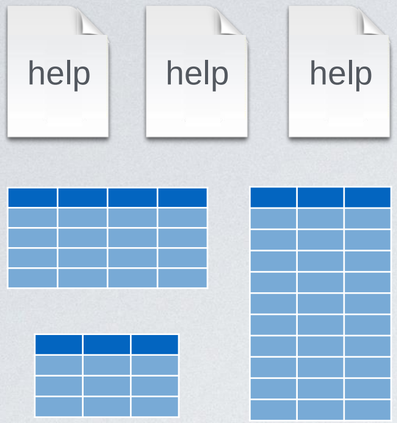
Assignment

How do you use them?

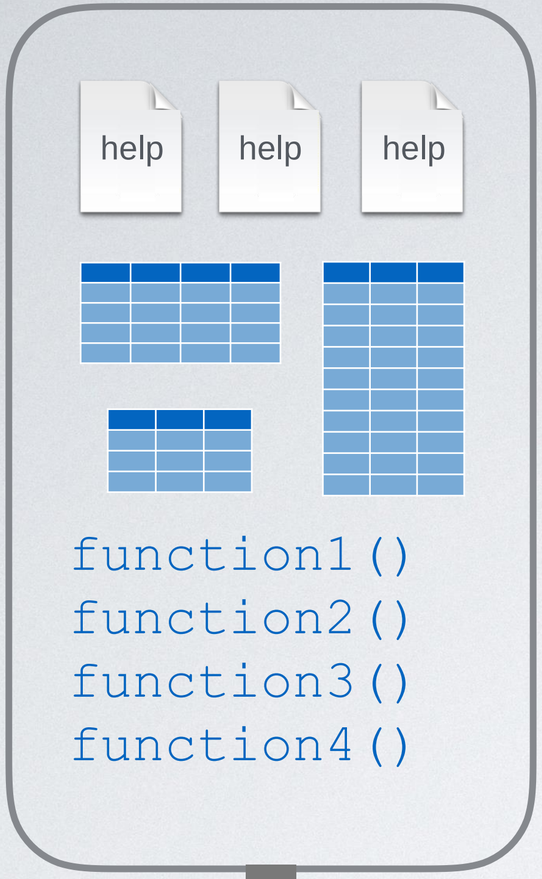
Check the Documentation!

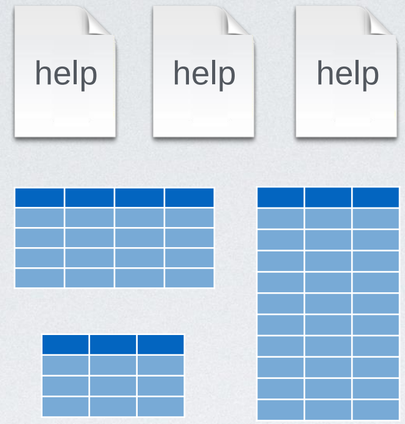


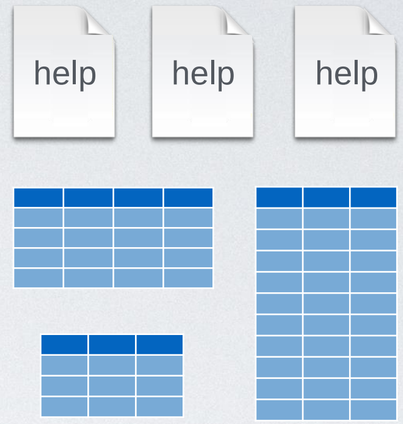
R Packages

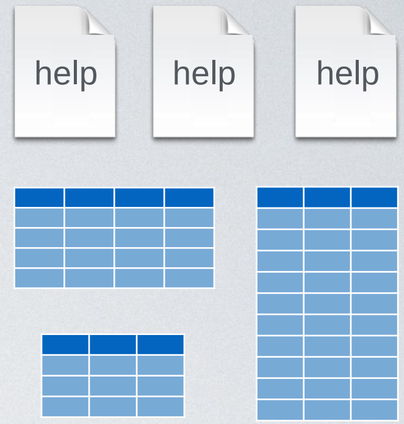



```
function1()
function2()
function3()
function4()
```

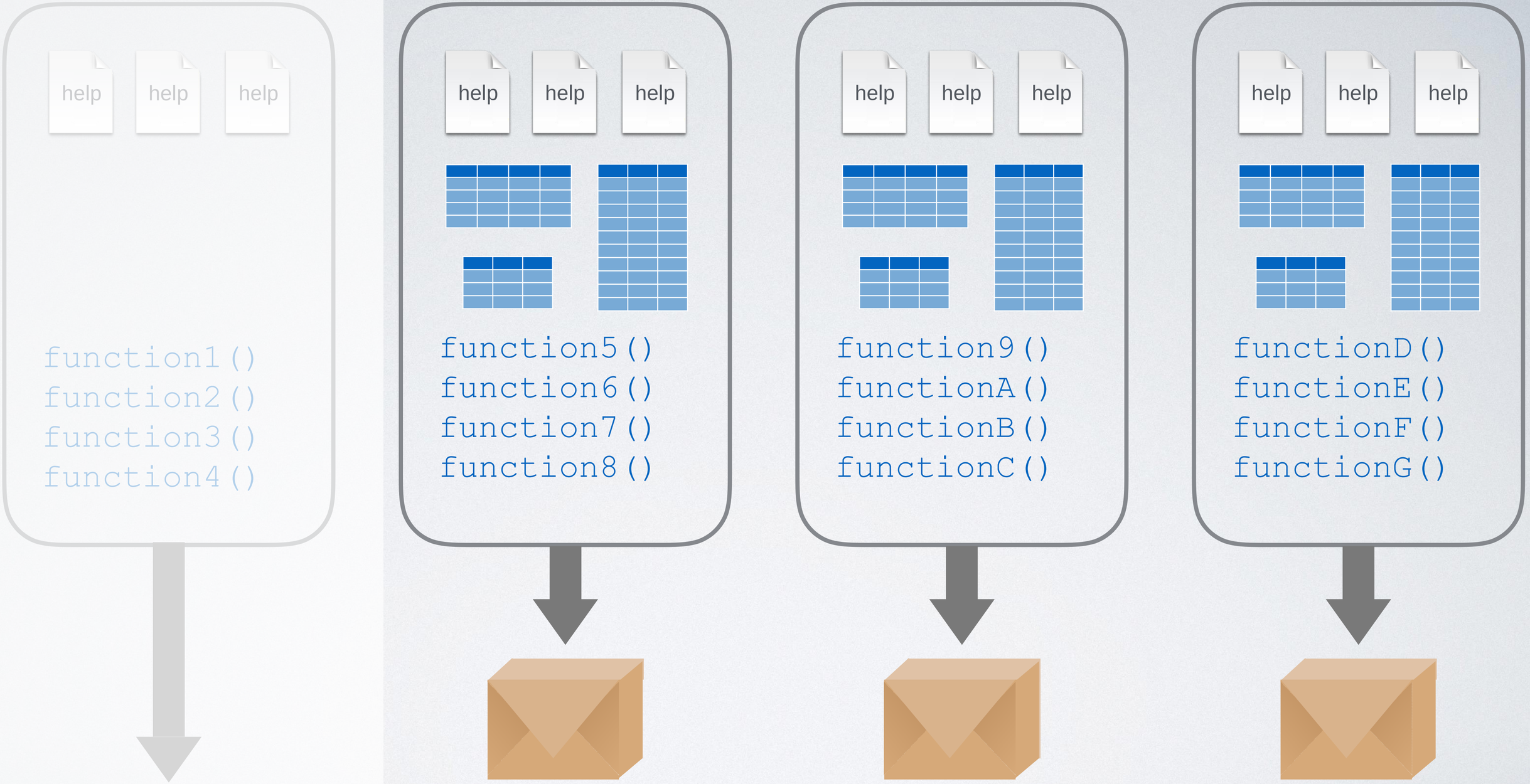

help
help
help
function1()
function2()
function3()
function4()


help
help
help
function5()
function6()
function7()
function8()


help
help
help
function9()
functionA()
functionB()
functionC()

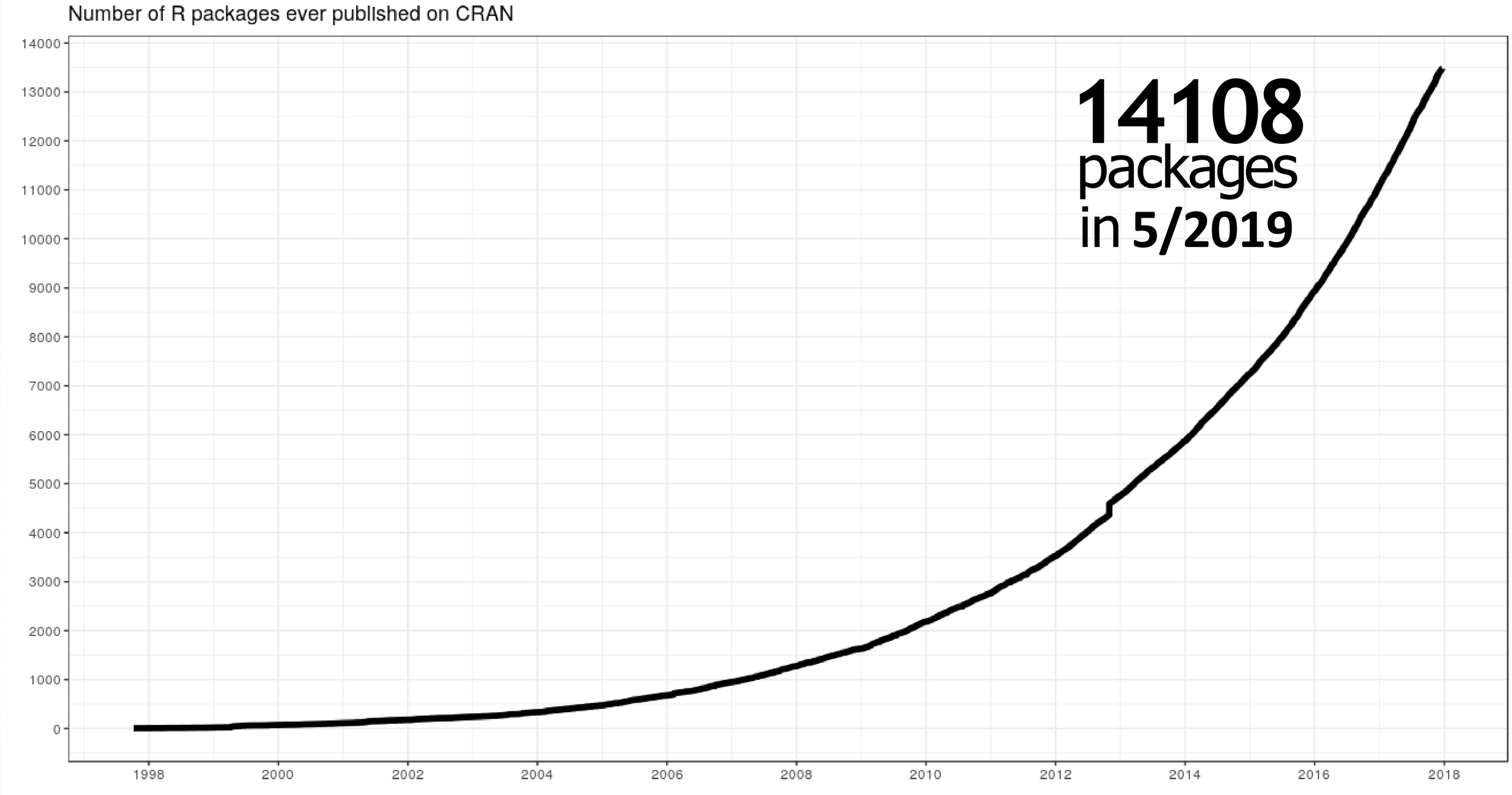

help
help
help
functionD()
functionE()
functionF()
functionG()

Base R



Base R

R Packages



Graph by https://besteamperu.org

The Comprehensive R Archive

Secure https://cran.r-project.org

CRAN
Mirrors
What's new?
Task Views
Search

About R
R Homepage
The R Journal

Software
R Sources
R Binaries
Packages
Other

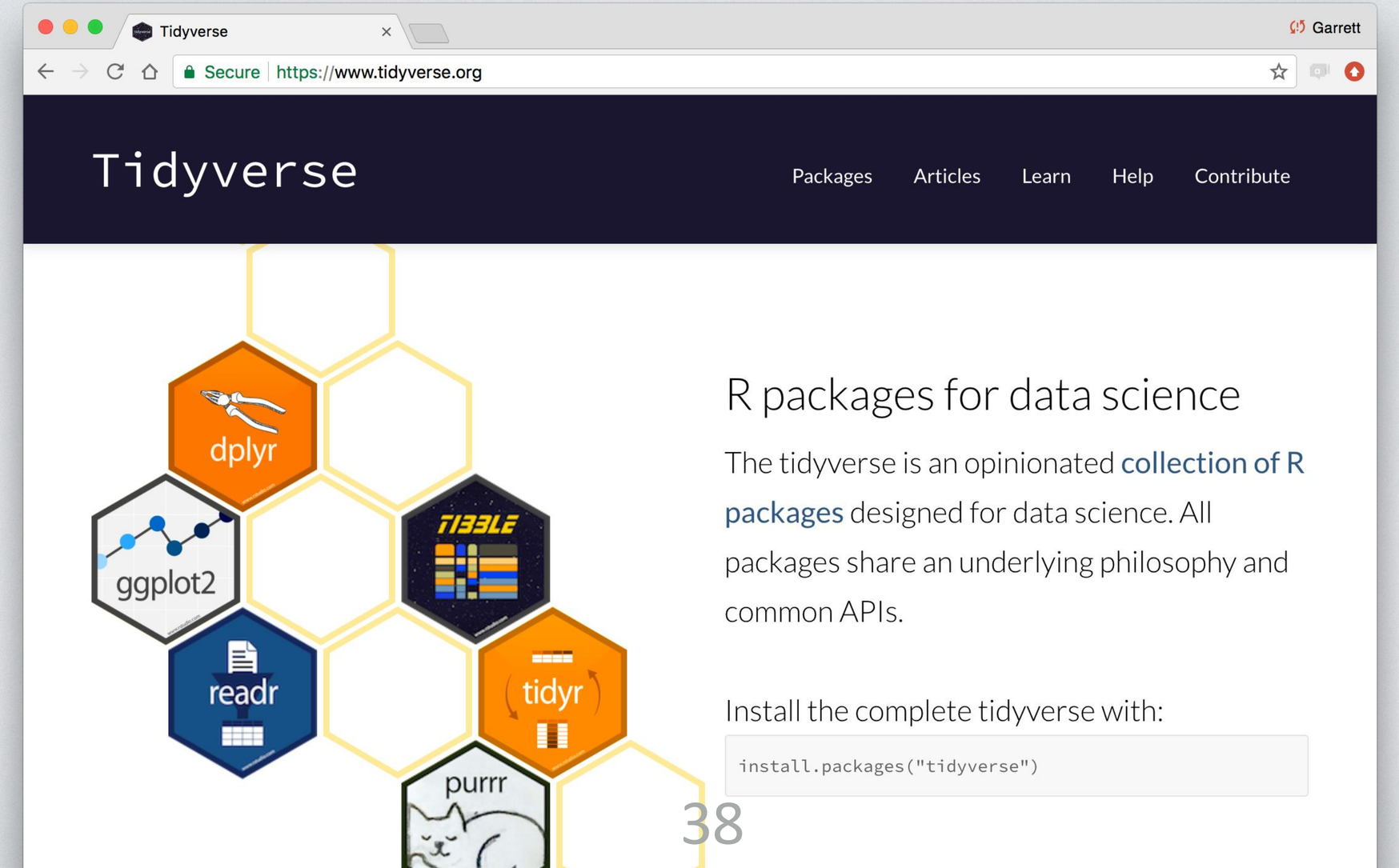
Documentation
Manuals
FAQs
Contributed

Available CRAN Packages By Name

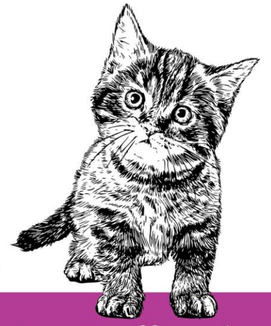
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

A3	Accurate, Adaptable, and Accessible Error Metrics for Predictive Models
abbyyR	Access to Abbyy Optical Character Recognition (OCR) API
abc	Tools for Approximate Bayesian Computation (ABC)
ABCanalysis	Computed ABC Analysis
abc.data	Data Only: Tools for Approximate Bayesian Computation (ABC)
abcdeFBA	ABCDE_FBA: A-Biologist-Can-Do-Everything of Flux Balance Analysis with this package
ABCoptim	Implementation of Artificial Bee Colony (ABC) Optimization
ABCp2	Approximate Bayesian Computational Model for Estimating P2
ABC.RAP	Array Based CpG Region Analysis Pipeline
abcrf	Approximate Bayesian Computation via Random Forests
abctools	Tools for ABC Analyses
abd	The Analysis of Biological Data
abf2	Load Gap-Free Axon ABF2 Files
ABHgenotypeR	Easy Visualization of ABH Genotypes
abind	Combine Multidimensional Arrays
abjutils	Useful Tools for Jurimetrical Analysis Used by the Brazilian Jurimetrics Association
abn	Modelling Multivariate Data with Additive Bayesian Networks
abodOutlier	Angle-Based Outlier Detection

tidyverse.org



How to actually learn any new programming concept



Essential

Changing Stuff and Seeing What Happens

O RLY?

@ThePracticalDev

The internet will make those bad words go away

Essential

Googling the Error Message

O RLY?

The Practical Developer
@ThePracticalDev

Cutting corners to meet arbitrary management deadlines

Essential

Copying and Pasting from Stack Overflow

O'REILLY®

The Practical Developer
@ThePracticalDev

Software can be chaotic, but we make it work

Expert

Trying Stuff Until it Works

O RLY?

The Practical Developer
@ThePracticalDev

Does it run? Just leave it alone.

Writing Code that Nobody Else Can Read

The Definitive Guide

O RLY?

@ThePracticalDev

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!

Your Turn

Go to Rstudio

Go to the console and type in:

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

```
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")
```

I've already installed all of the packages you need

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")
```

R Markdown

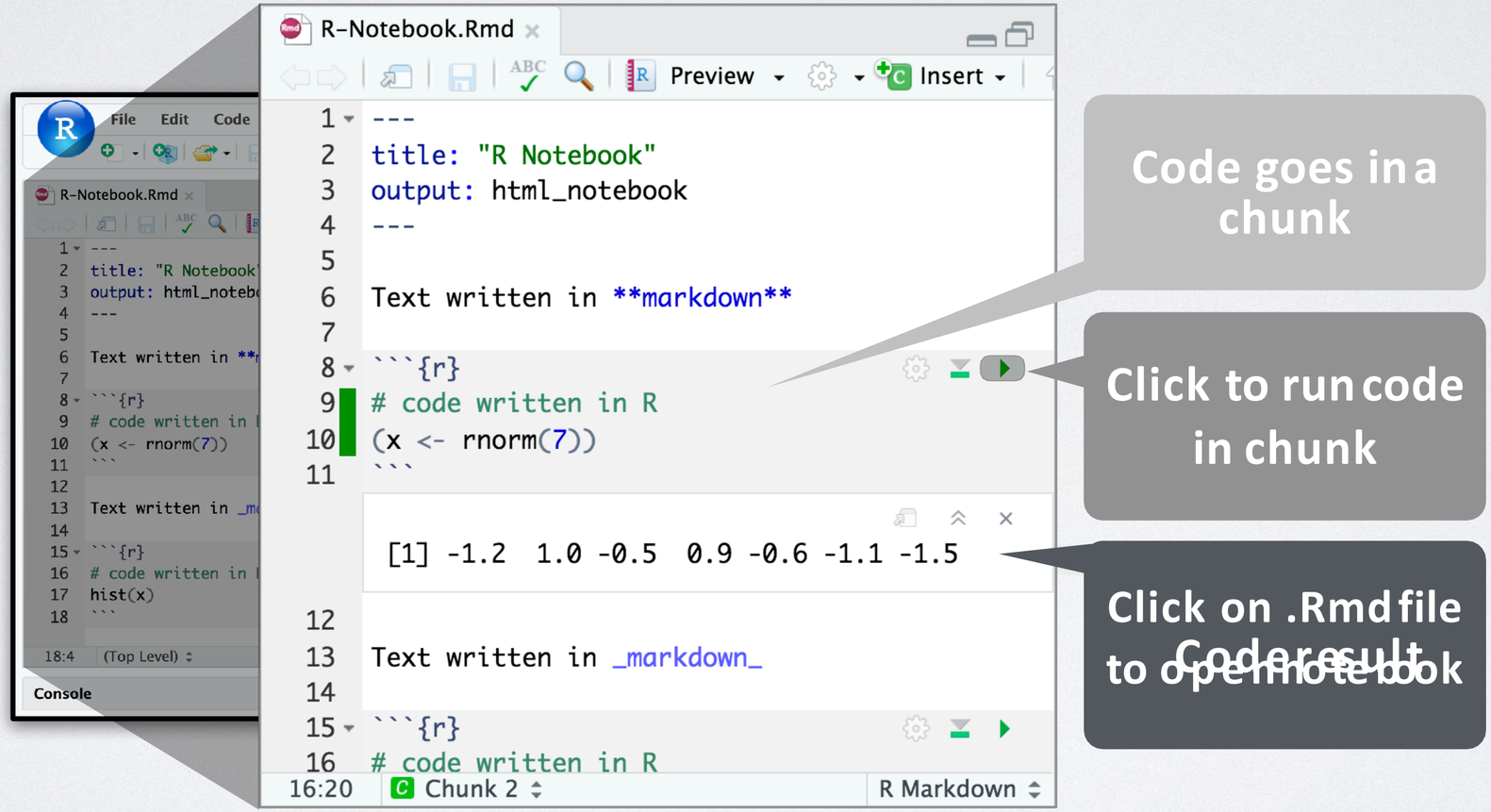
Your Turn

Go to **01-Introduction**.

Open **01-RMarkdown-Exercises.Rmd**. Read through the file and do everything it tells you to do.

R Markdown

An authoring format for Data Science.



Code chunks

Insert a chunk of R code with

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

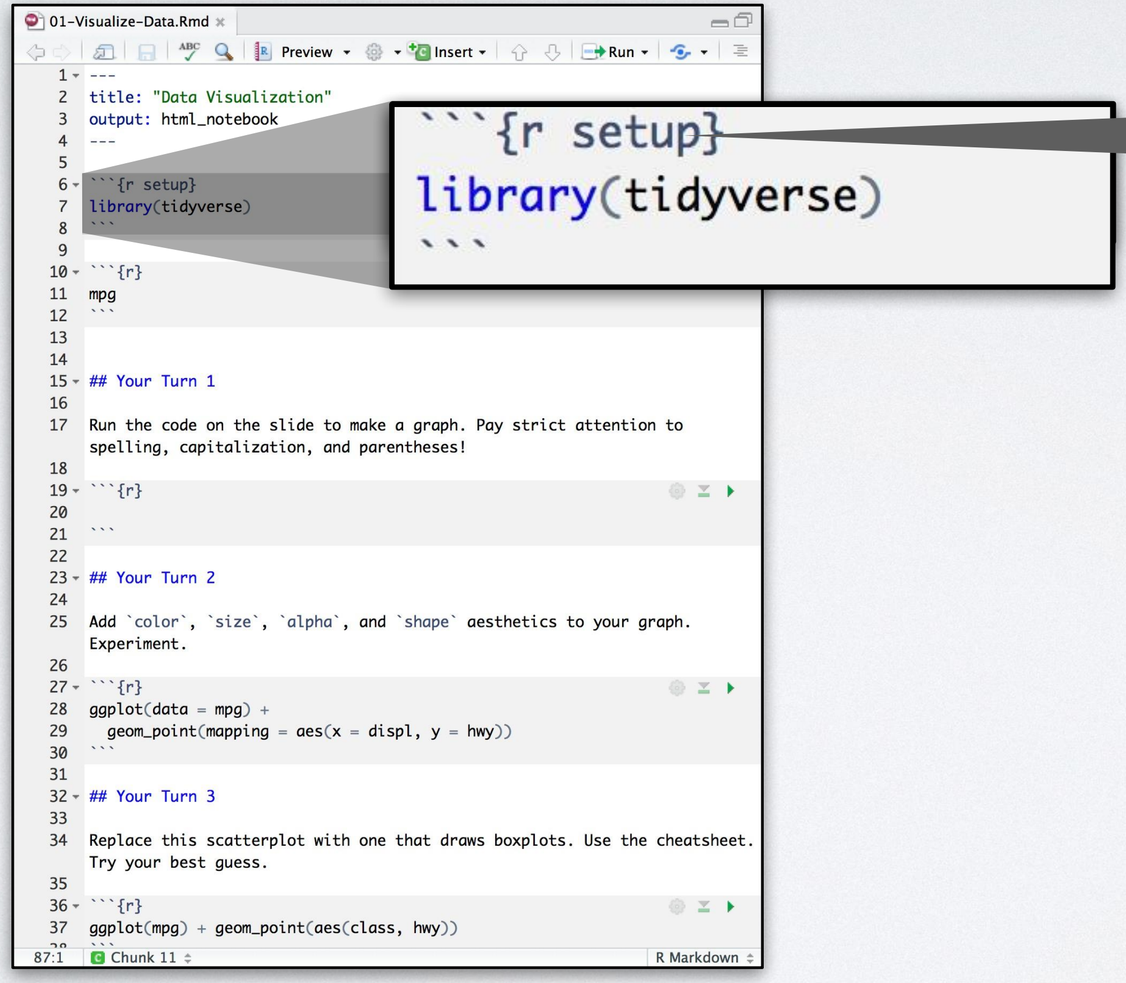
⌘ + Opt + i (Mac)

Ctrl + Alt + i (PC)

CC BY-SA RStudio

Setup

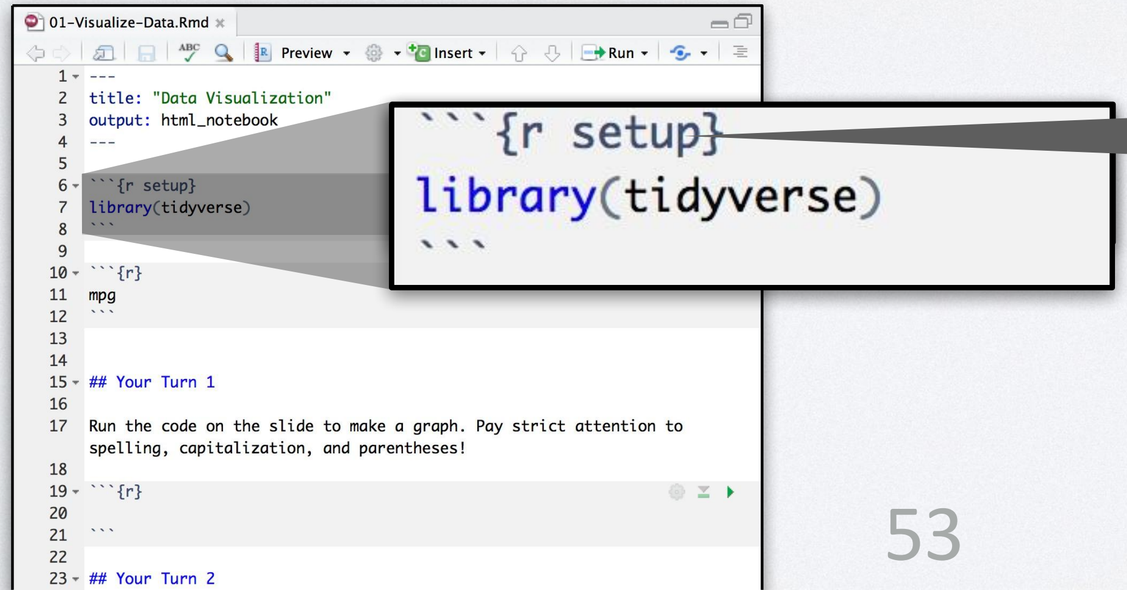
The setup chunk is always run once before anything else



chunk labels are optional, the setup label is special

Your Turn

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




```{r setup}
library(tidyverse)
```

chunk labels are optional, the setup label is special

Your Turn 0

Navigate up to the **02-Visualize** folder.

Open **02-Visualize-Exercises.Rmd**

Add a setup chunk that loads the tidyverse packages.