

[github.com/MarketBridge/
tidyverse-training](https://github.com/MarketBridge/tidyverse-training)

Introduction to the Tidyverse

Ari Lamstein

Exercise: What do you already know?

- Split into groups of 2.
- Ask you partner: *What do you already know about the Tidyverse?*
- 5 minutes



Introduction

Main Ideas

Notes

Summary

How I can apply this to my work

Notes form - please take out

About Me

- Ari Lamstein
- Studied Math and CS
- Former "tech worker"
- Author of the *choroplethr* package
- Certified Tidyverse Instructor
- Now do R stuff at MB



Day 1

**Introduction and
Visualize Data**

9:00 - 10:30

Morning Break

10:30 - 11:00

Transform Data

11:00 - 12:00

Lunch

12:00 - 1:00

Transform Data

1:00 - 2:30

Afternoon Break

2:30 - 3:00

**Transform Data,
Import Data, Q&A**

3:00 - 5:00

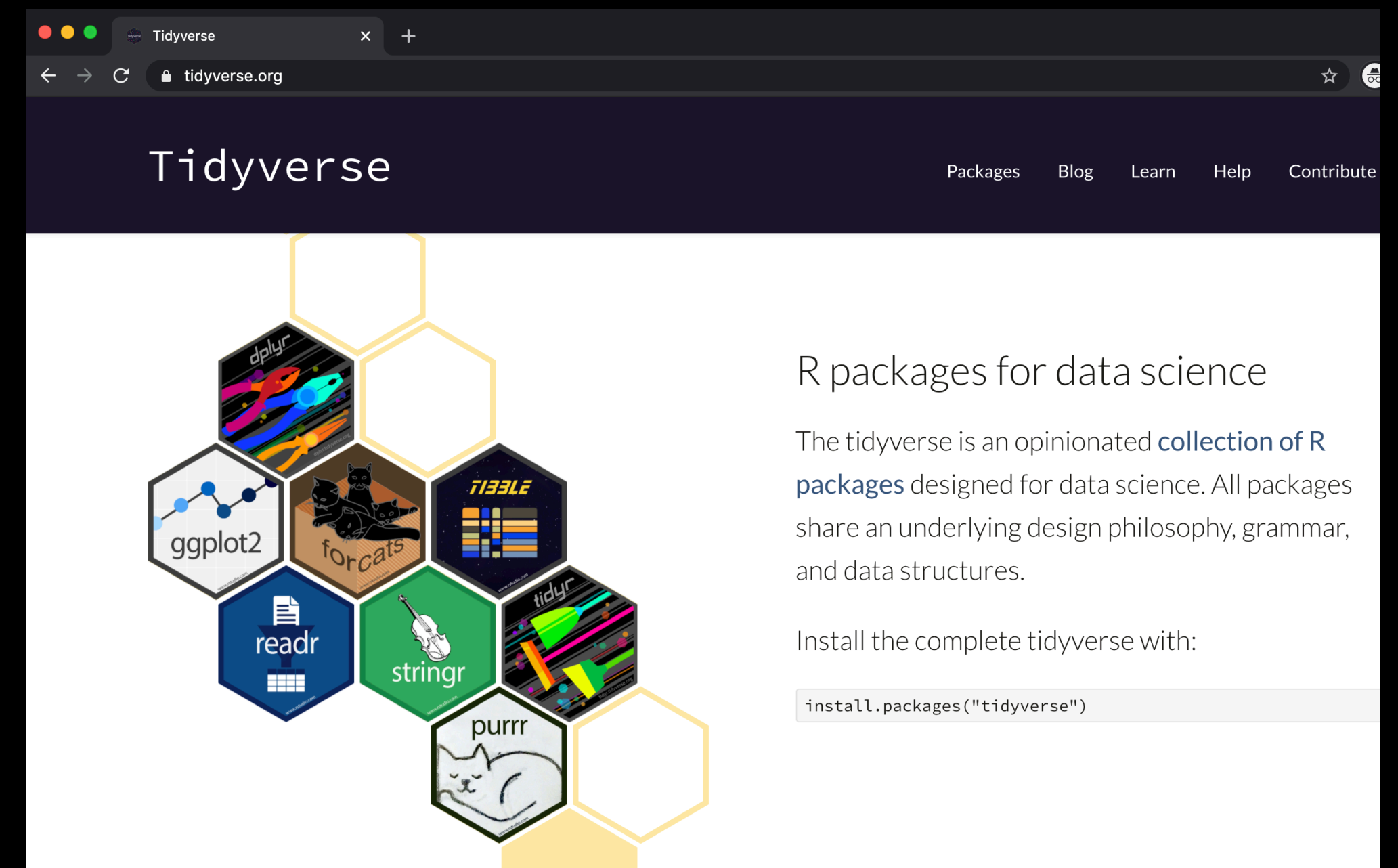
Learning Goals

- What is the Tidyverse (and why learn it?)
- Complete a project with R Notebooks

What is the Tidyverse?

Answer 1: A Marketing Term

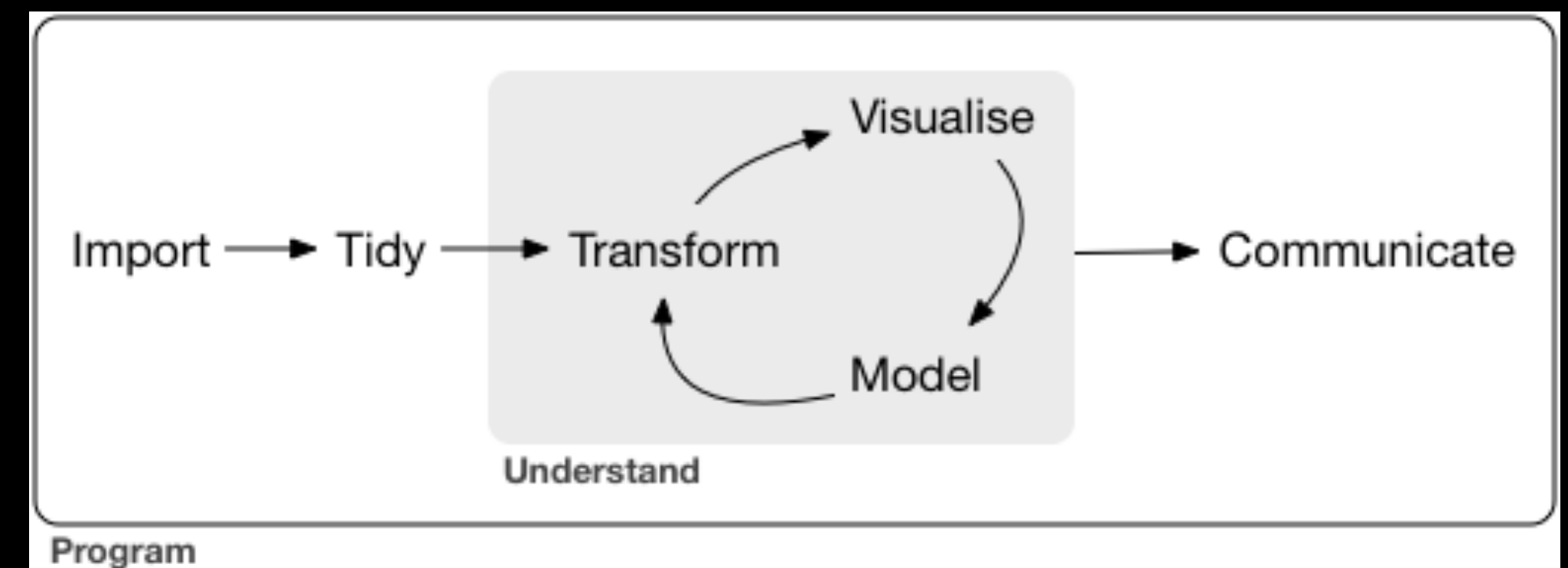
- Refers to dozens of packages created by Hadley Wickham (Chief Data Scientist at RStudio)
 - Replaces the "Hadley-verse"
- 2 main packages
 - **ggplot2** (data visualization)
 - **dplyr** (data manipulation)



What is the Tidyverse?

Answer 2: A "Dialect of R"

- Most R packages don't "agree" with each other
 - ?View vs. ?summary
- All Tidyverse packages are consistent
- Tidyverse packages cover all aspects of the "Data Science Workflow"



What is the Tidyverse?

Answer 3: A Package that Contains Packages

- "tidyverse" is a name of an actual R package
- The package itself does not do much
- Installing the package makes R install the most useful packages in the Tidyverse (such as ggplot2 and dplyr)

tidyverse: Easily Install and Load the 'Tidyverse'

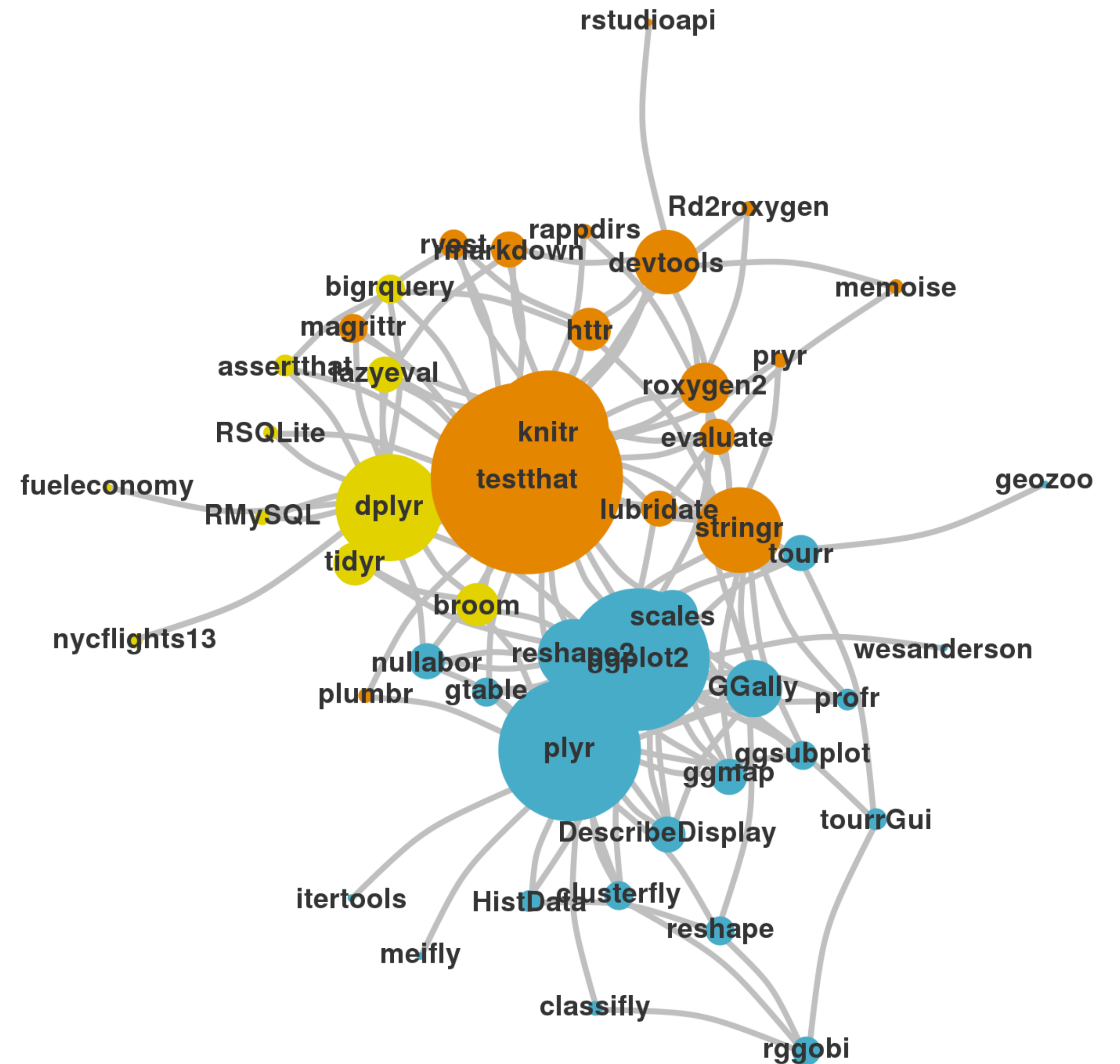
The 'tidyverse' is a set of packages that work in harmony because they share common design goals. It is possible to install all 'tidyverse' packages in a single step. Learn more about the 'tidyverse' at [<https://www.tidyverse.org>](#)

Version:	1.3.0
Depends:	R (≥ 3.2)
Imports:	broom (≥ 0.5.2), cli (≥ 1.1.0), crayon (≥ 1.3.4), dbplyr (≥ 1.6), lubridate (≥ 1.7.4), magrittr (≥ 1.5), modelr (≥ 0.1.0), rvest (≥ 0.3.5), stringr (≥ 1.4.0), tibble (≥ 2.1.3)
Suggests:	covr , feather , glue , knitr , rmarkdown , testthat
Published:	2019-11-21
Author:	Hadley Wickham [aut, cre], RStudio [cph, fnd]
Maintainer:	Hadley Wickham <hadley at rstudio.com>

```
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")
```



```
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")
```

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

does the equivalent of

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

What is the Tidyverse?

Answer 4: Packages that work with "Tidy Data"

- ggplot2 and dplyr only work with "tidy data"
- Tidy data = a data frame where:
 - Each **variable** is in its own **column**
 - Each **case** is in its own **row**
 - Each **value** is in its own **cell**

Subject (car)	Speed (mph)	Distance (ft)
1	25	85
2	24	120
3	24	93
4	24	92

Learning Goals

- What is the Tidyverse (and why learn it?)
- Complete a project with R Notebooks

R Notebooks

An authoring format for Data Science.

The screenshot displays the RStudio R Notebook interface. The main editor shows an R Markdown document with the following content:

```
1 ---
2 title: "R Notebook"
3 output: html_notebook
4 ---
5
6 Text written in markdown
7
8 ```{r}
9 # code written in R
10 (x <- rnorm(7))
11 ```
12
13 Text written in markdown
14
15 ```{r}
16 # code written in R
17 hist(x)
18 ```
```

Callouts highlight key features:

- Click to run all code chunks above:** Points to the 'Run All' button (a green play icon) in the toolbar.
- Click to run code in chunk:** Points to the 'Run Chunk' button (a green play icon) next to the first code chunk.
- Code result:** Points to the output of the first code chunk, which is a vector of seven random numbers: `[1] -1.2 1.0 -0.5 0.9 -0.6 -1.1 -1.5`.

The interface also shows a file explorer on the left, a console at the bottom, and a status bar indicating the current chunk and output format.

Exercise: Your First R Notebook

- Open up **00-Introduction.Rmd**
- Read through the notebook and follow all the instructions.
- Ask your partner (or me) if you have any questions.
- Share your result with your partner.
- 5 minutes

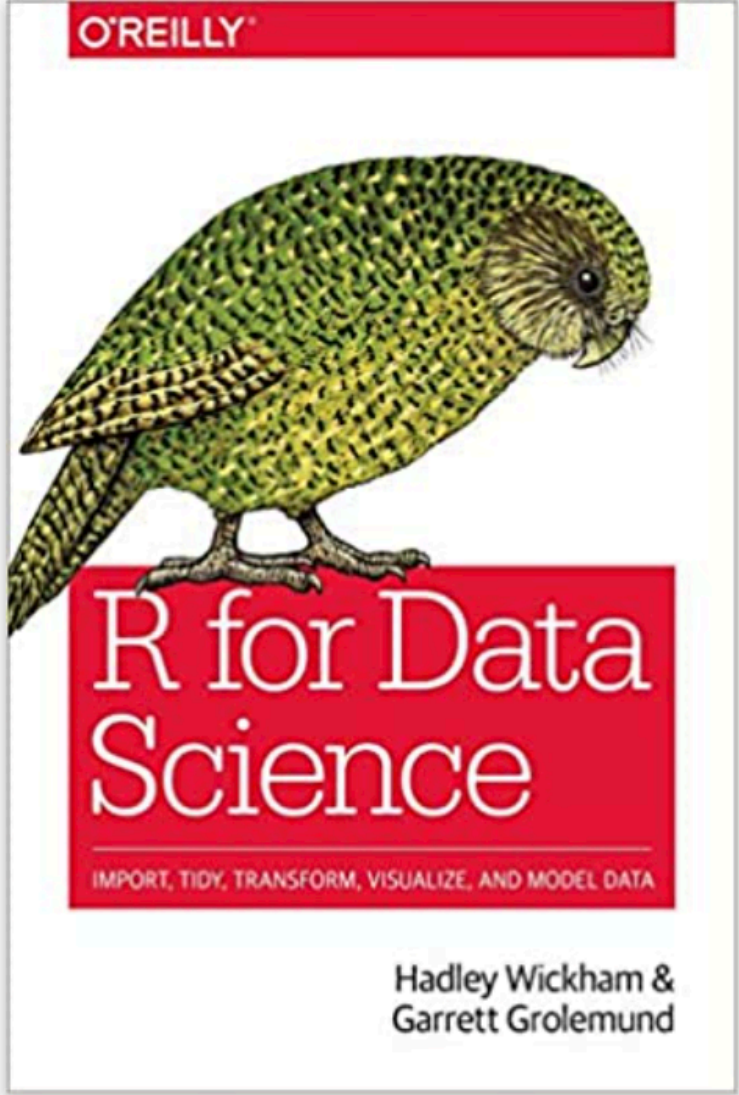


What next?

Using R after the workshop

★★★★★ 686 ratings

Look inside



ISBN-13: 978-1491910399
ISBN-10: 1491910399
[Why is ISBN important?](#)

Have one to sell?
Sell on Amazon

Add to List

Kindle \$24.99

Paperback
\$30.99 - \$43.73

Other Sellers
See all 2 versions

☐ Buy used: \$30.99


☒ Buy new: \$43.73


In Stock.
Ships from and sold by Amazon.com.
Available at a lower price from [other sellers](#) that may not offer free Prime shipping.
Arrives: **Wednesday, Nov 18** [Details](#)
Fastest delivery: **Tomorrow**
Order within 14 hrs and 34 mins [Details](#)


List Price: ~~\$49.99~~
Save: \$6.26 (13%)
34 new from \$37.22
& **FREE Shipping.** [Details](#)

Select delivery location

Qty: 1

**Enjoy fast, FREE delivery, exclusive deals and award-winning movies & TV shows with Prime**
[Try Prime](#) and start saving today with **Fast, FREE Delivery**

Add to Cart

Buy Now

Many Resources for Learning

datacamp

Learn ▾Features ▾PricingDataCamp For Business

Q

Sign In

Get Started

Introduction to the Tidyverse

Get started on the path to exploring and visualizing your own data with the tidyverse, a powerful and popular collection of data science tools within R.

Start Course For Free

⌚ 4 Hours

▶ 16 Videos

<> 50 Exercises

👤 153,289 Learners

📖 4150 XP

Create Your Free Account

Google

LinkedIn

Facebook

or

Email Address

Email address

Password

Password

Start Learning For Free

By continuing you accept the Terms of Use and Privacy Policy. You also accept that you are aware that your data will be stored outside of the EU and that you are above the age of 16.

DataCamp.com

Closing Exercise

- Fill out the "Summary" section
- Fill out the "How I can apply to to my work"
- Share with your neighbor!

02 : 00

Introduction

Main Ideas	Notes _____ _____ _____ _____ _____ _____ _____ _____ _____ _____
-------------------	--

Summary

How I can apply this to my work

Any Questions?