Data Science in the tidyverse

Charlotte Wickham

cwick.co.nz

cwickham@gmail.com

@cvwickham

Introduction

Your Turn

Introduce yourselves to your neighbours:

Who are you?

What do you do with data?

How would you describe your R experience?

No sticky note: "I'm happily working on it"

Blue sticky note: "I'm all done and ready to move on"

Orange sticky note: "I'm stuck, can someone help me?"

Alternatively, flag one of us down

Hopefully, color-blind friendly, let me know if not.

This class is heavily based on

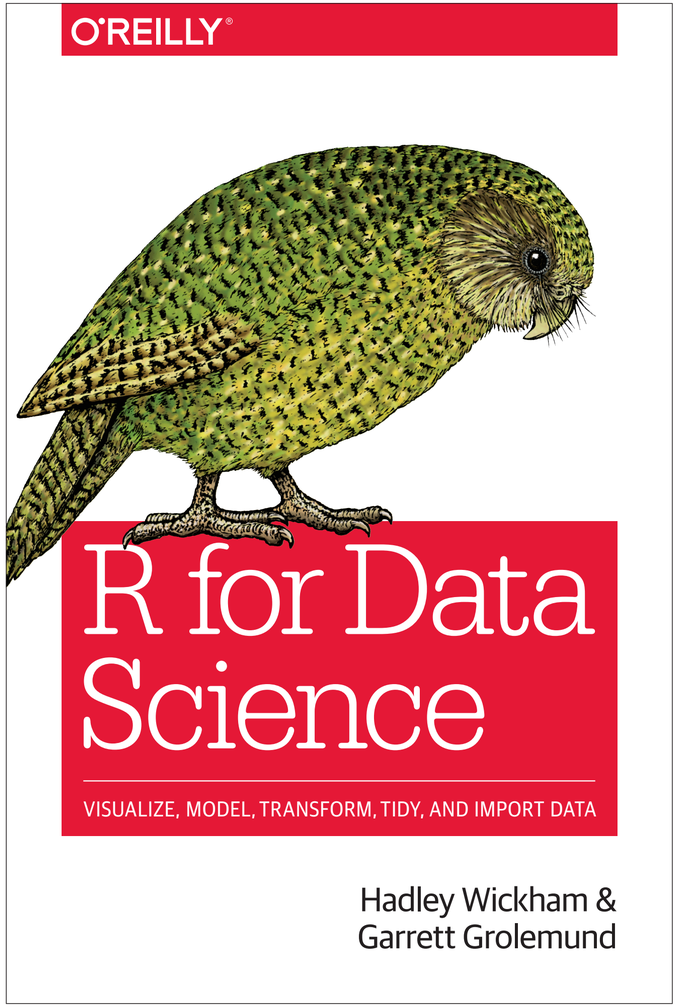
R for Data Science

http://r4ds.had.co.nz/

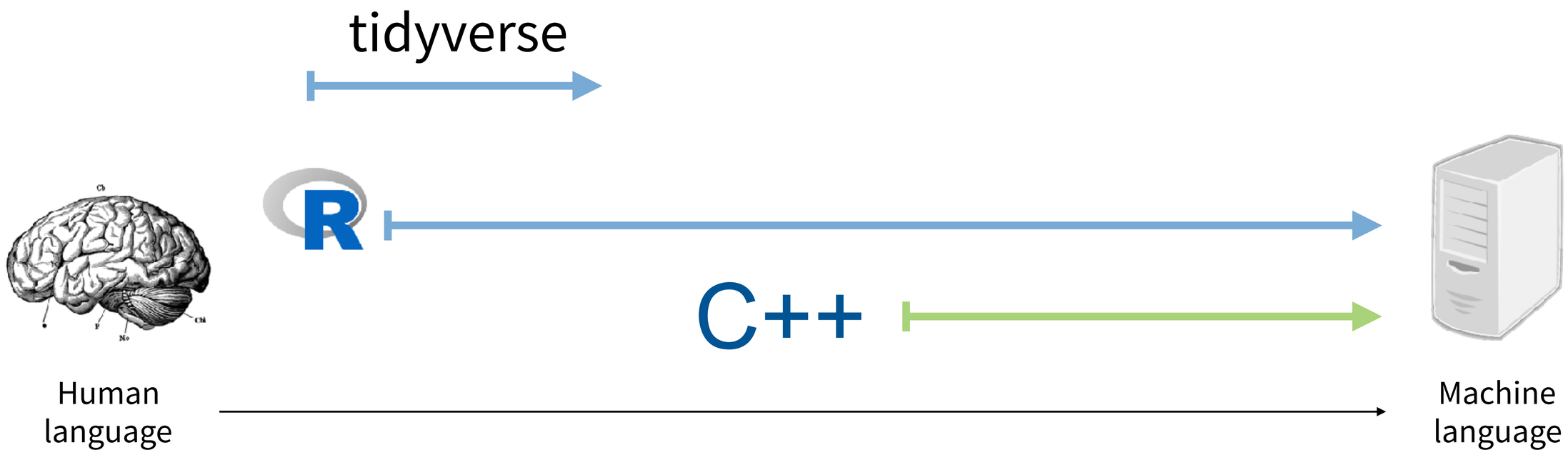
Links to the relevant sections of the book

In R4DS

Introduction



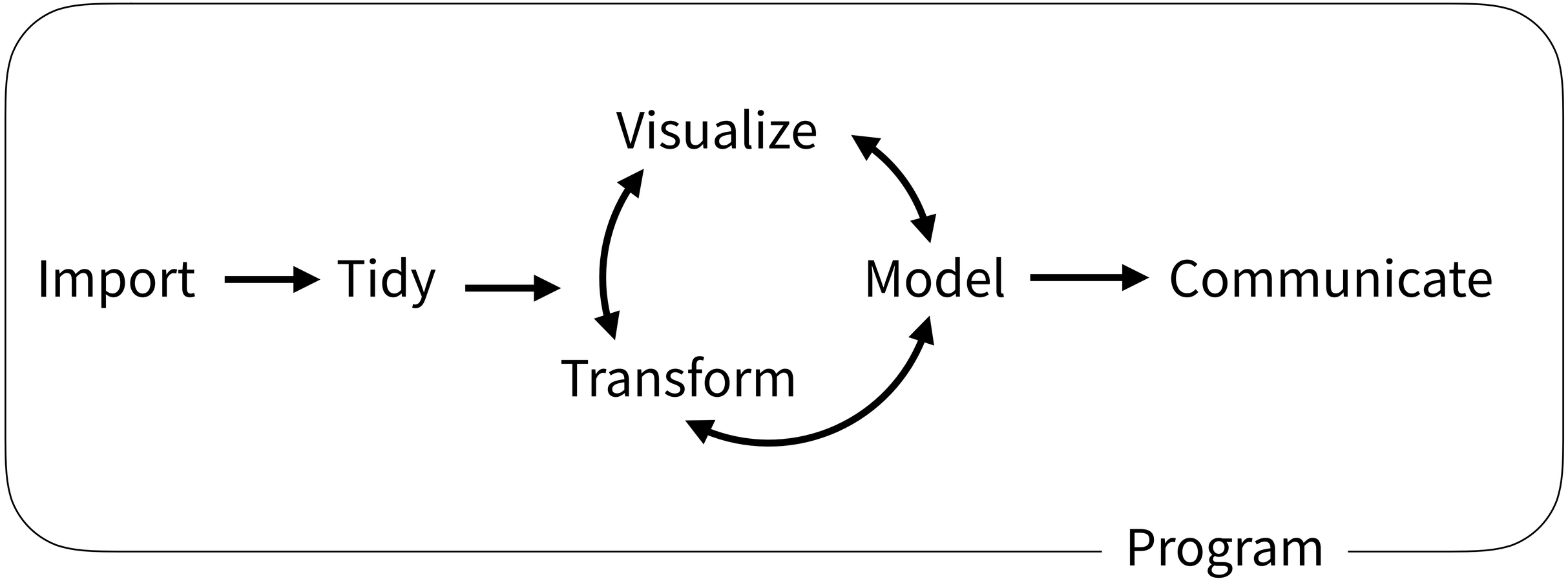
R - A computer language for scientists



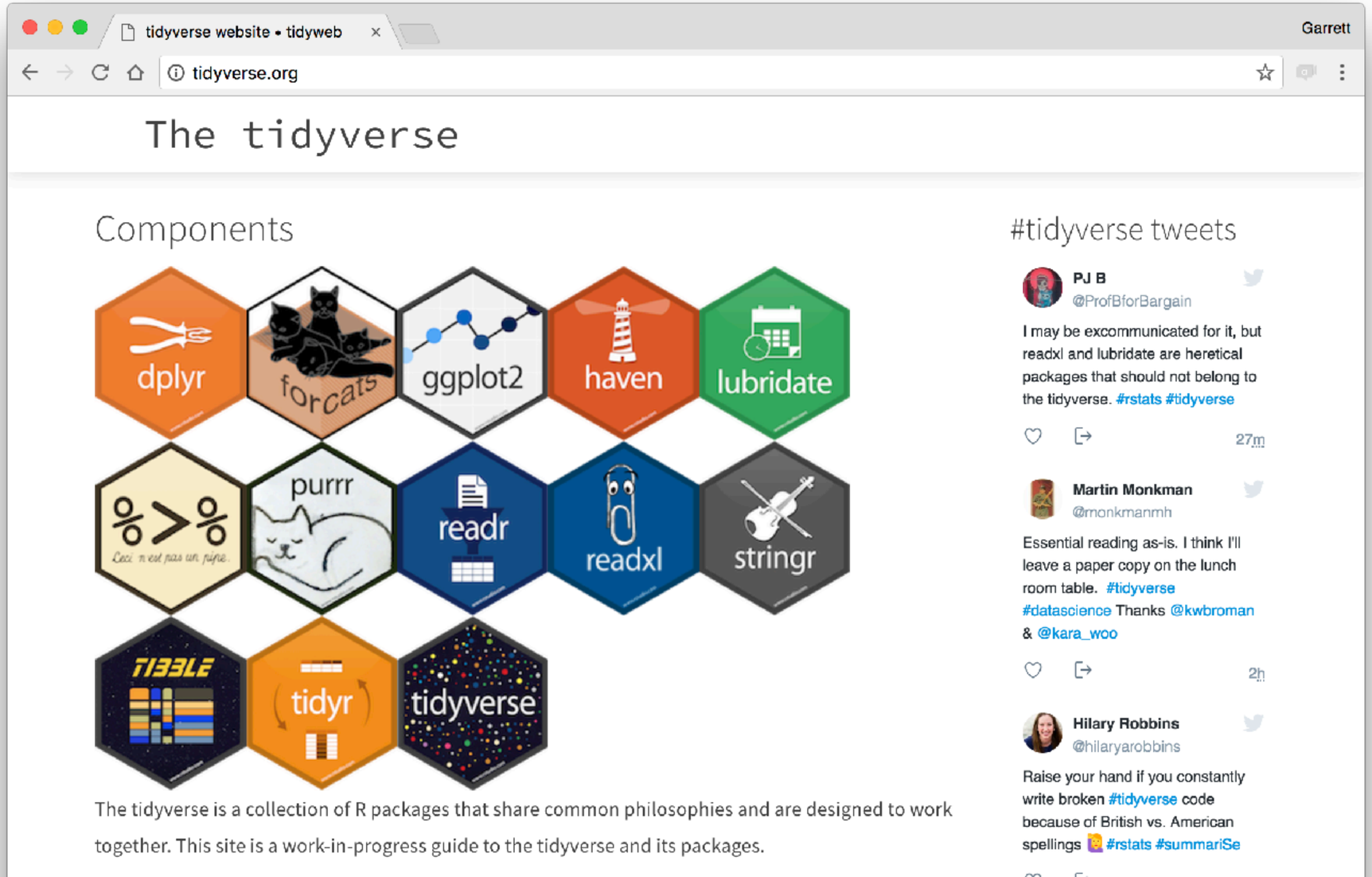
Adapted from 'Master the tidyverse' CC by RStudio

You spend less time thinking about code, and more time thinking about **data analysis**.

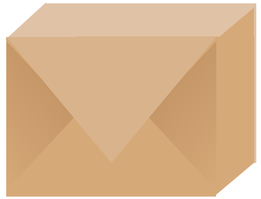
(Applied) **Data Science**



tidyverse.org



tidyverse



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

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

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

does the equivalent of

```r
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")
```

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

does the equivalent of

```r
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")
```

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

does the equivalent of

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

Day 1

Introduction and Visualize Data	9:00 - 10:30
Morning Break	10:30 - 11:00
Visualize Data/ Transform Data	11:00 - 12:30
Lunch	12:30 - 1:30
Transform Data	1:30 - 3:00
Afternoon Break	3:00 - 3:30
Tidy Data/ Case Study	3:30 - 5:00

Day 2

Data Types	9:00 - 10:30
Morning Break	10:30 - 11:00
Iteration	11:00 - 12:30
Lunch	12:30 - 1:30
Modelling	1:30 - 3:00
Afternoon Break	3:00 - 3:30
Organization with list columns	3:30 - 5:00

Getting Started

Your Turn

Instructions with screenshots at bit.ly/rstudio18-setup

First, if you haven't already

- Visit https://rstudio.cloud/project/10871
- Log In / Sign Up
- "Save a copy" of the project
- Open project/data-science-in-the-tidyverse.Rproj

When you have your copy of the project, let us know by **putting** up the **Blue** post-it.

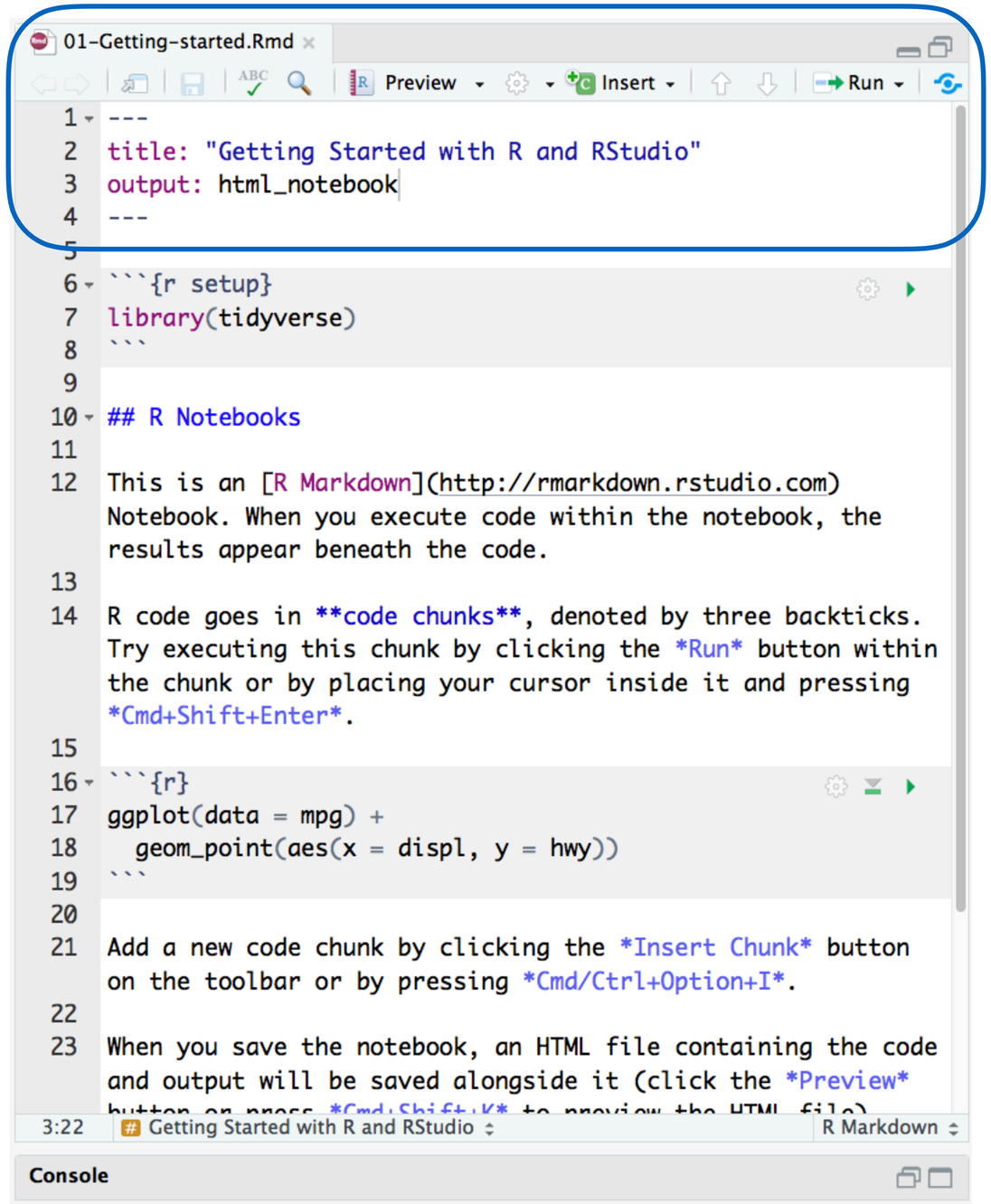
Then open 00-Getting-started.Rmd and take a look around!

rstudio.cloud

A bit like RStudio Server, but hosted for you.

Currently in *alpha.*

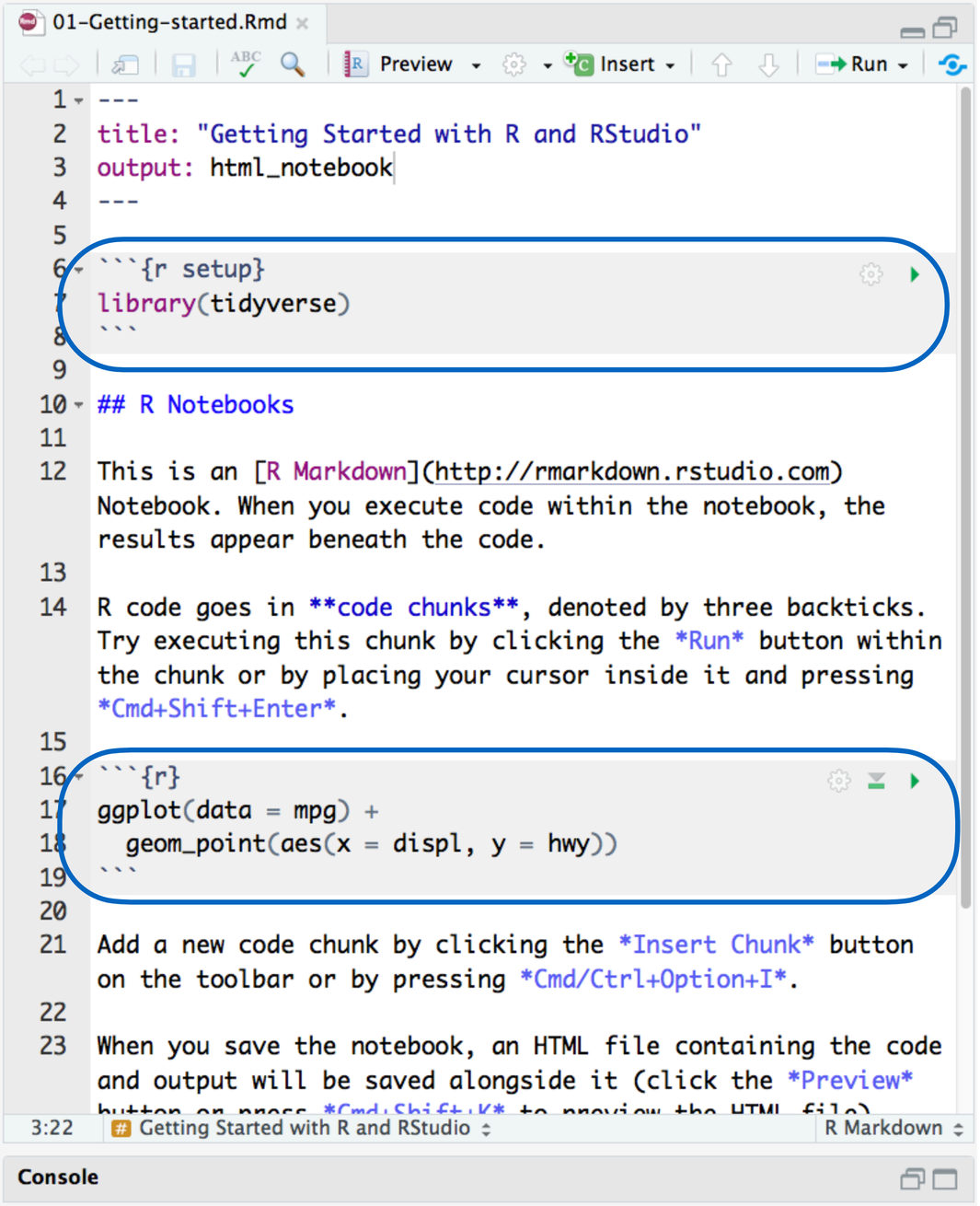
If you navigate away, go to rstudio.cloud, and you'll see your project in your Workspace.



R notebooks

An authoring format for Data Science

00-Getting-started.Rmd is an R notebook



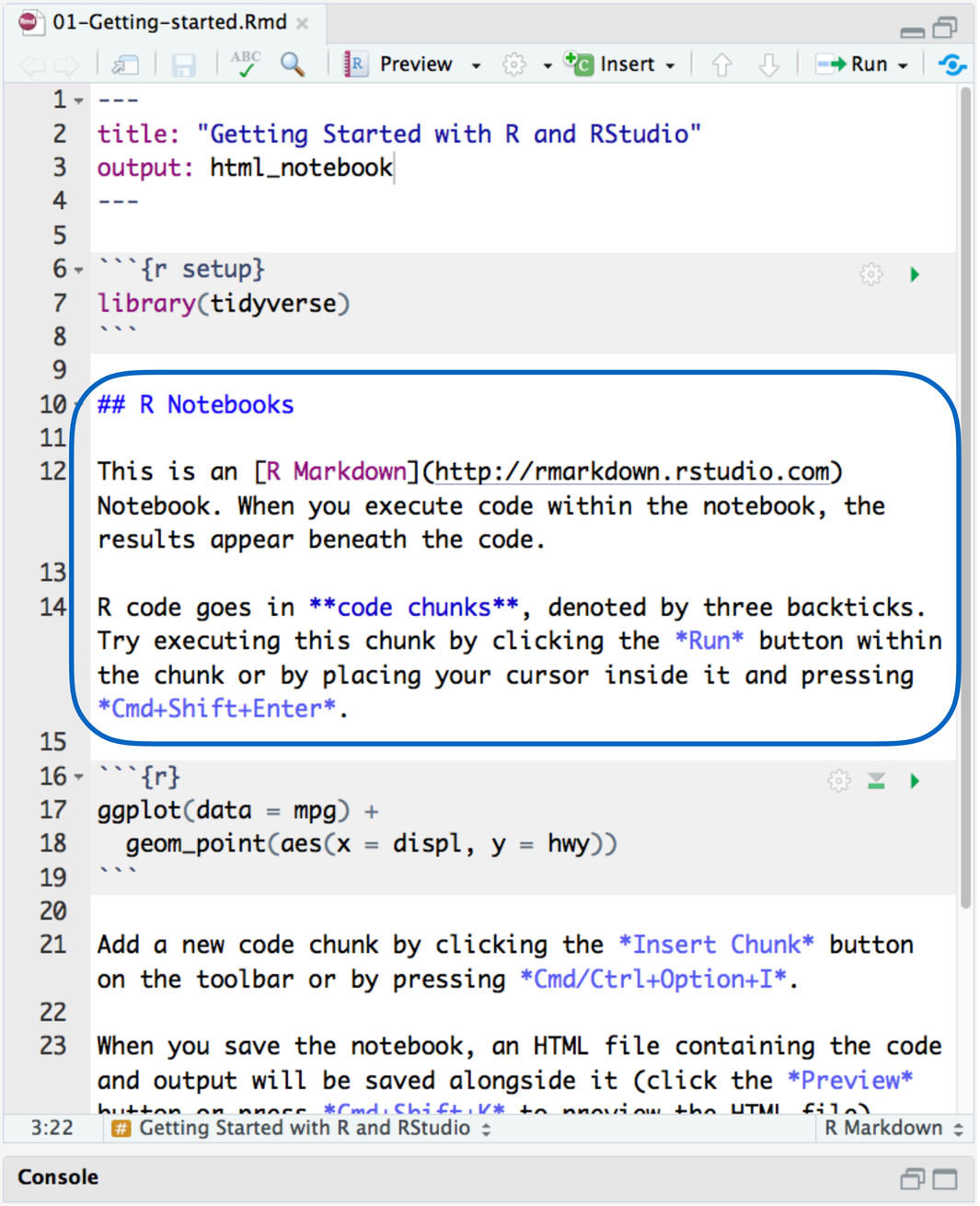
R notebooks

An authoring format for Data Science

00-Getting-started.Rmd is an R notebook

Integrates:

- Code



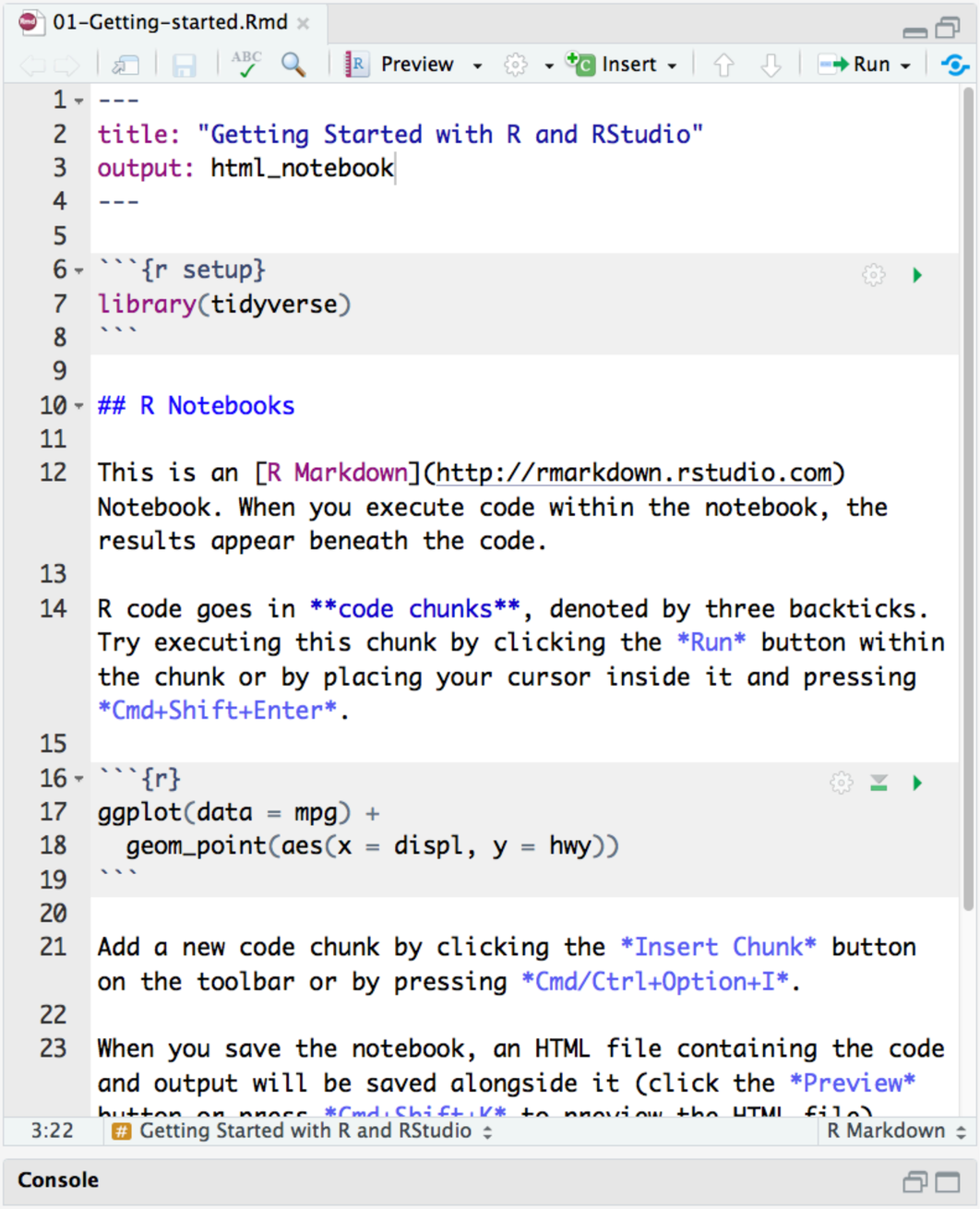
R notebooks

An authoring format for Data Science

00-Getting-started.Rmd is an R notebook

Integrates:

- Code
- Text



R notebooks

An authoring format for Data Science

00-Getting-started.Rmd is an R notebook

Integrates:

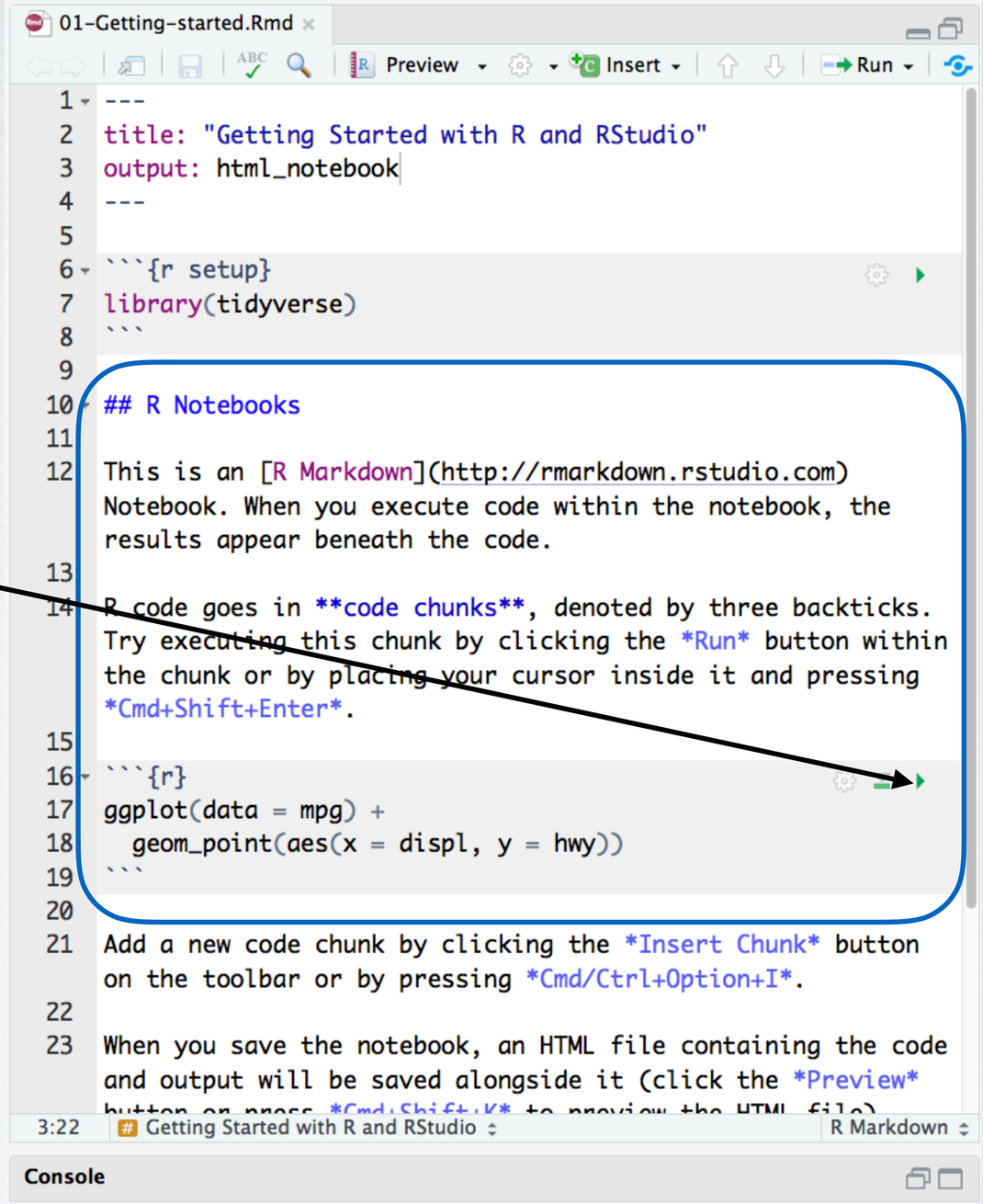
- Code
- Text
- Output

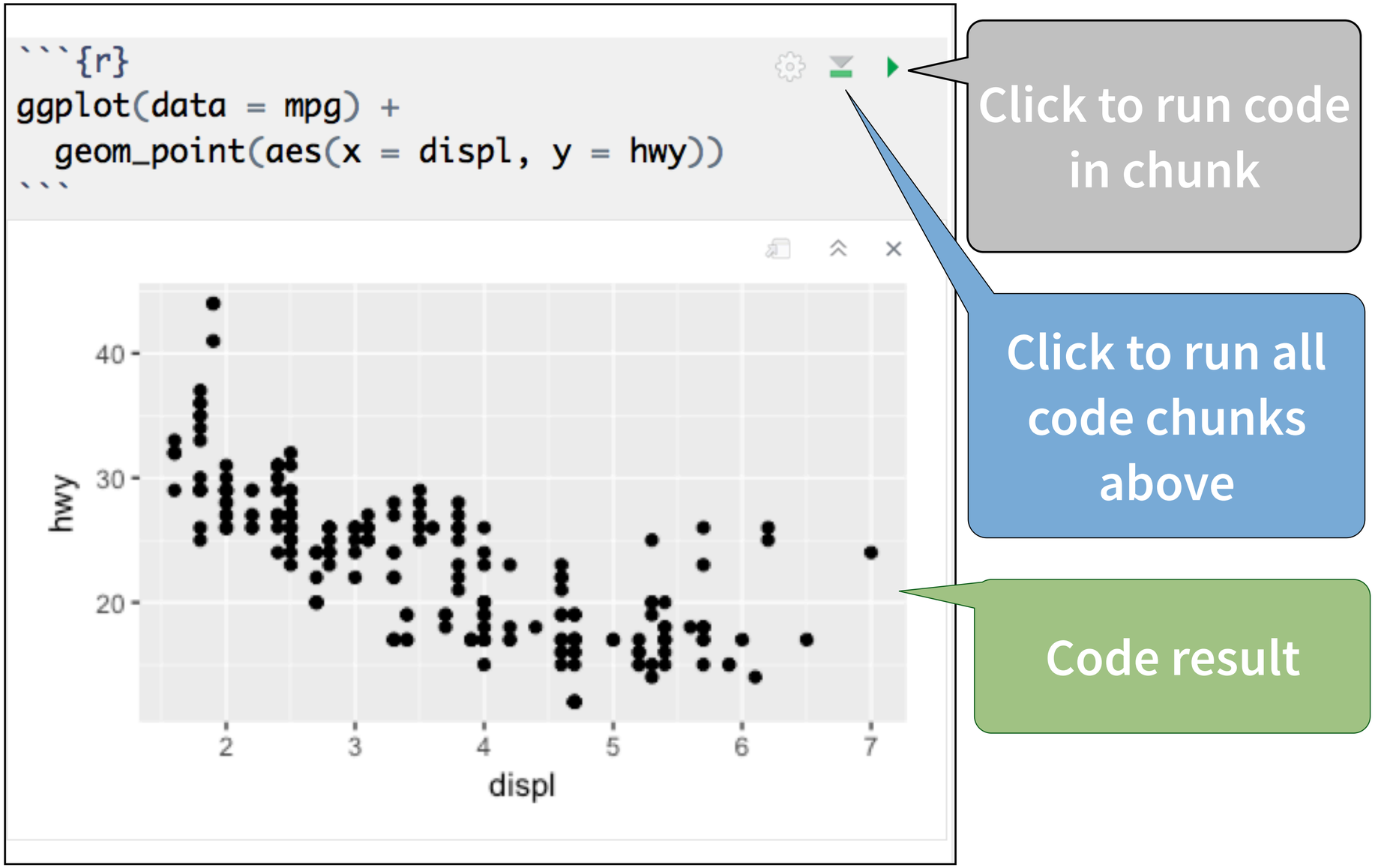
Your Turn

Read the instructions.

Run the code by hitting the play button,

or using the keyboard shortcut.





R Notebooks

An easy way to combine R code and narrative

Useful for us:

- I'll provide starter code
- You can complete "Your Turns"
- At the end, a useful record for you

Your Turn

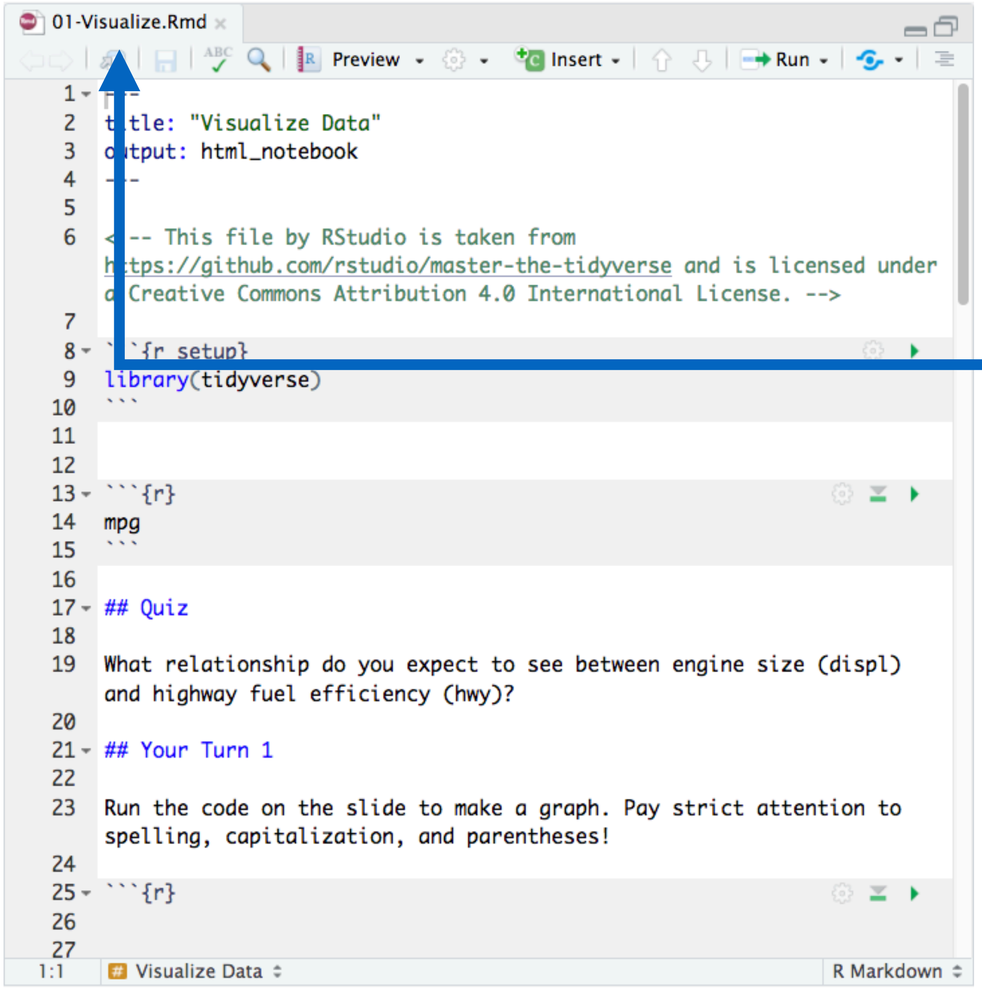
Open 01-Visualize.Rmd

I'm working on it

I'm stuck!

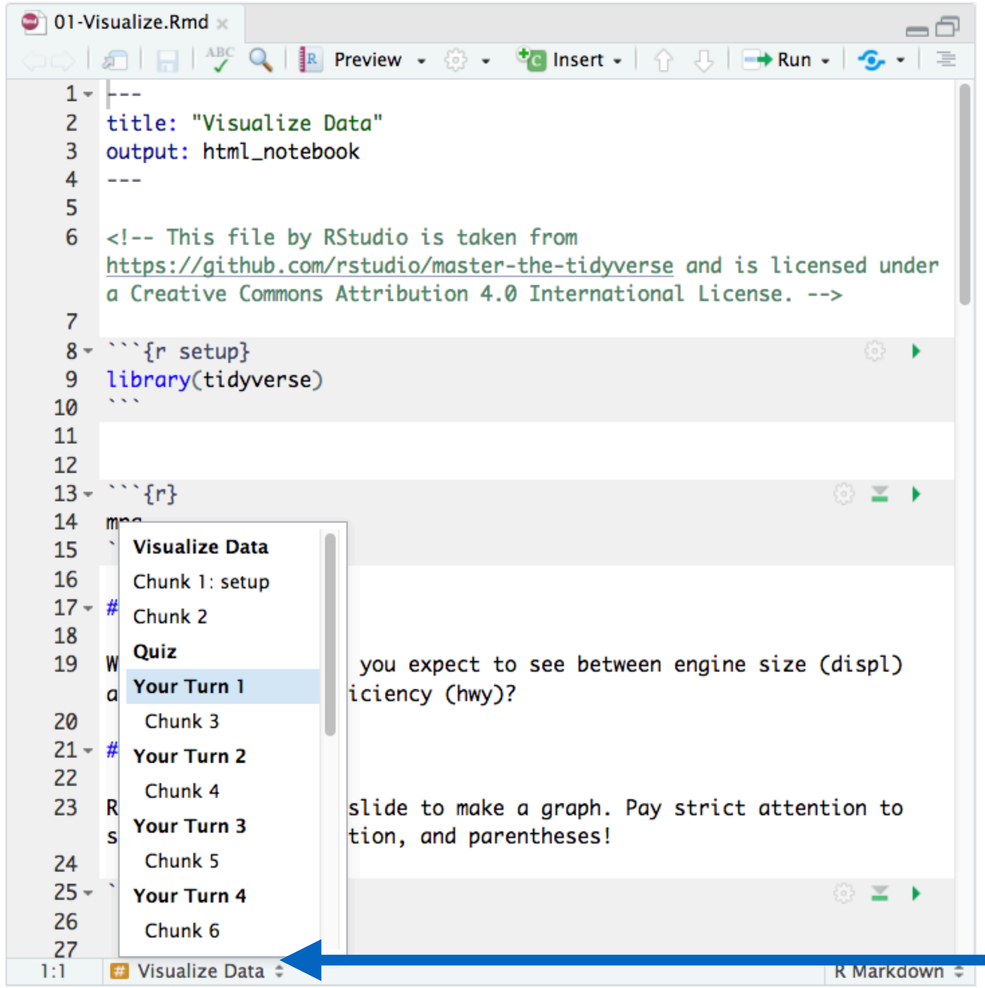
I'm done!

If you get lost or need to restart



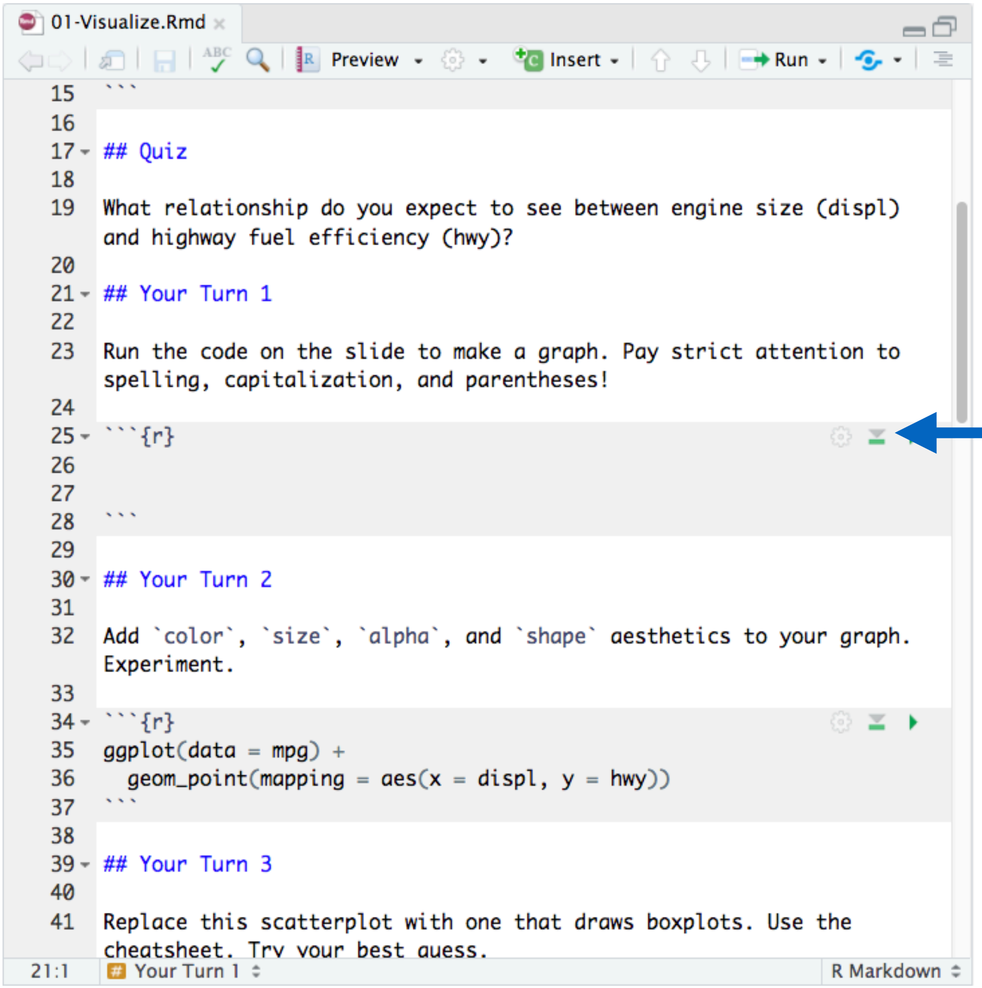
Check you are in the right file

If you get lost or need to restart



Use the section browser to quickly navigate to the right *Your Turn*

If you get lost or need to restart



Click to run all chunks before this one.

You should be ready to go.