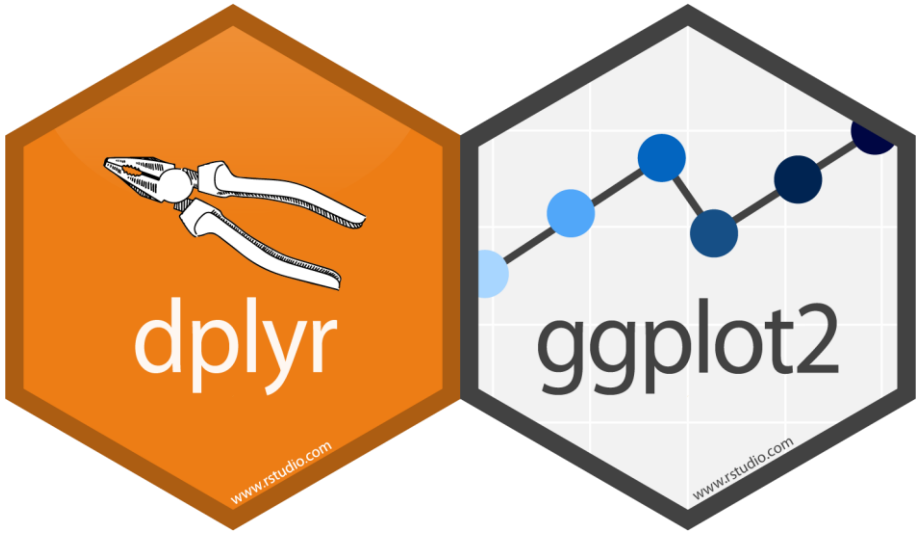
Introduction to R - Day 3:

Transform and Visualise with



James Morgan

@jmorgan

@mogranjm

jamesa.morgan@petermac.org

Adapted from “Exploring the Tidyverse” CC by Hadley Wickham

Introduction

HELLO

my name is

James

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

Green 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/

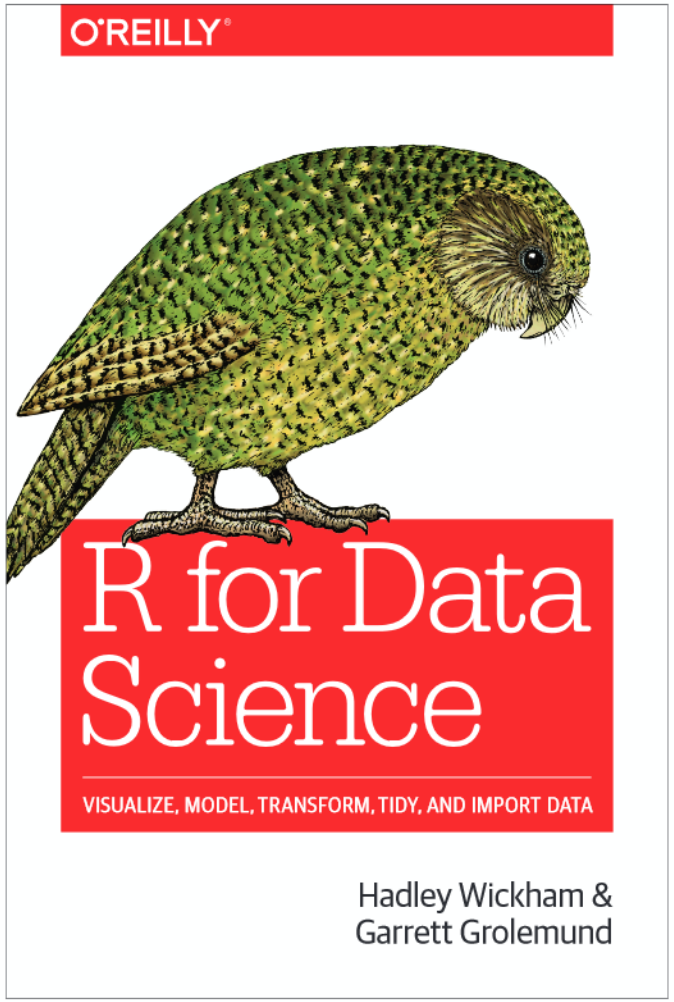
Chapter 3: Data Visualisation

Chapter 5: Data Transformation

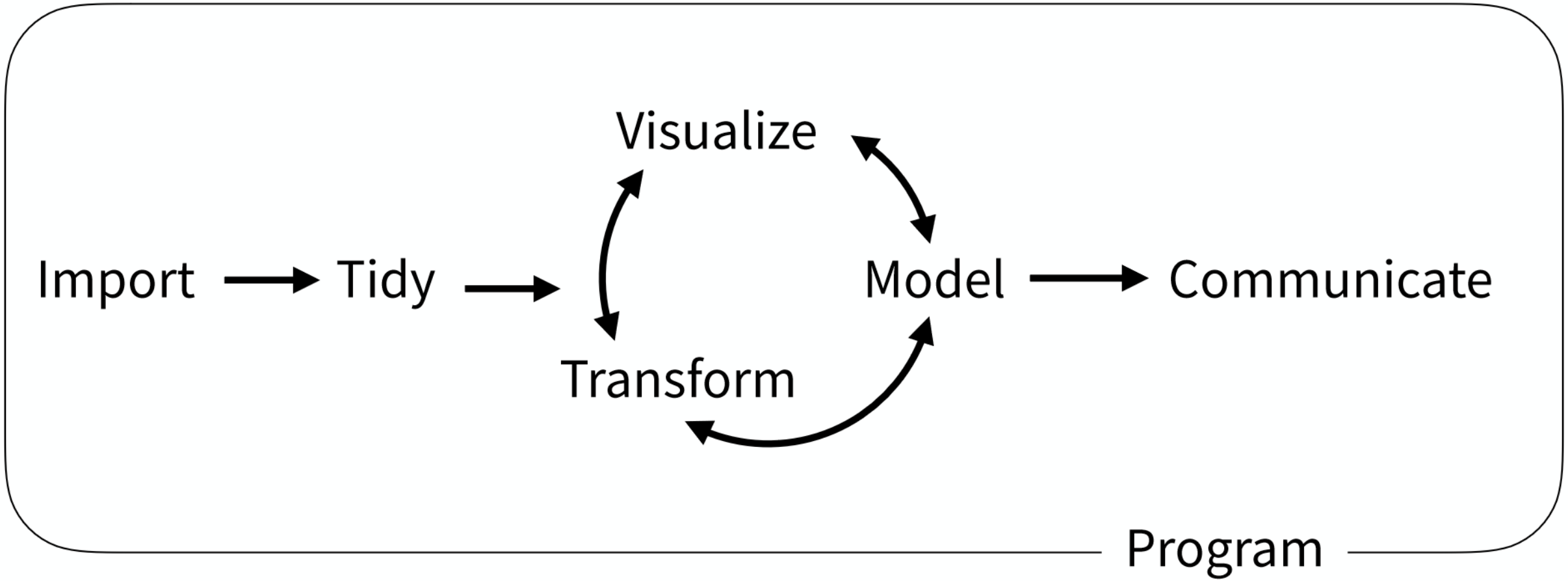
and

Exploring the Tidyverse

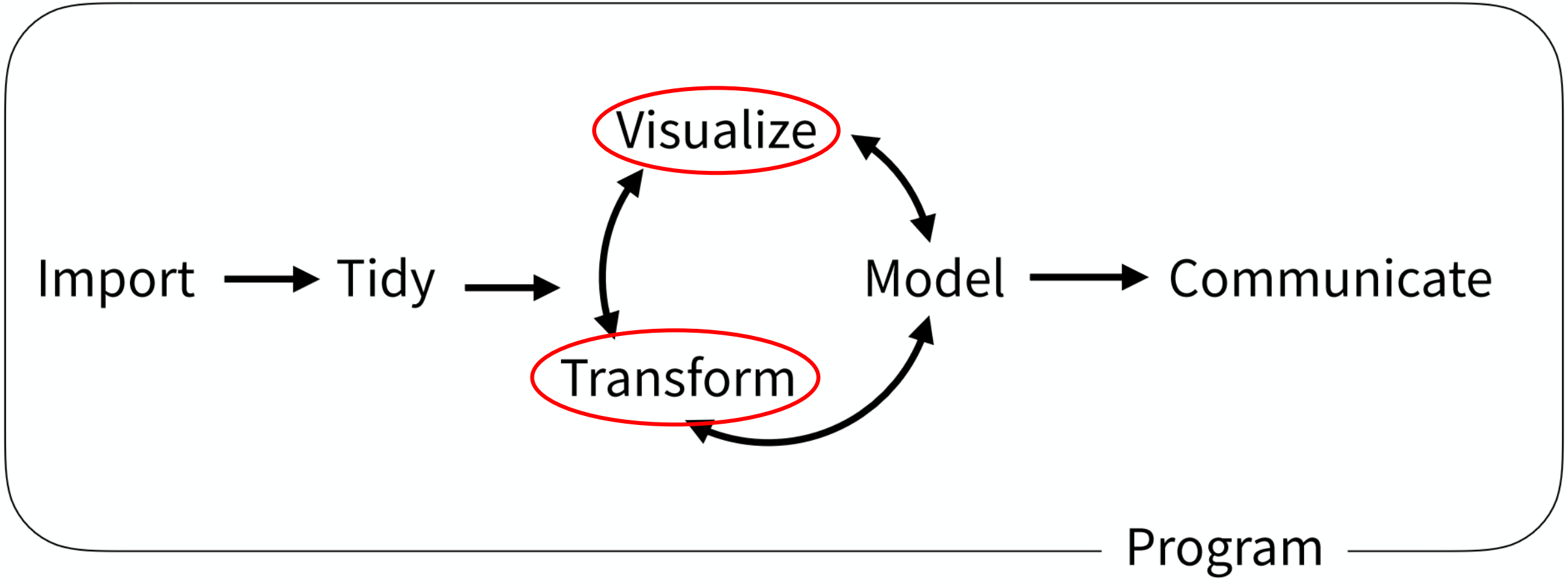
https://github.com/hadley/data-science-in-tidyverse/

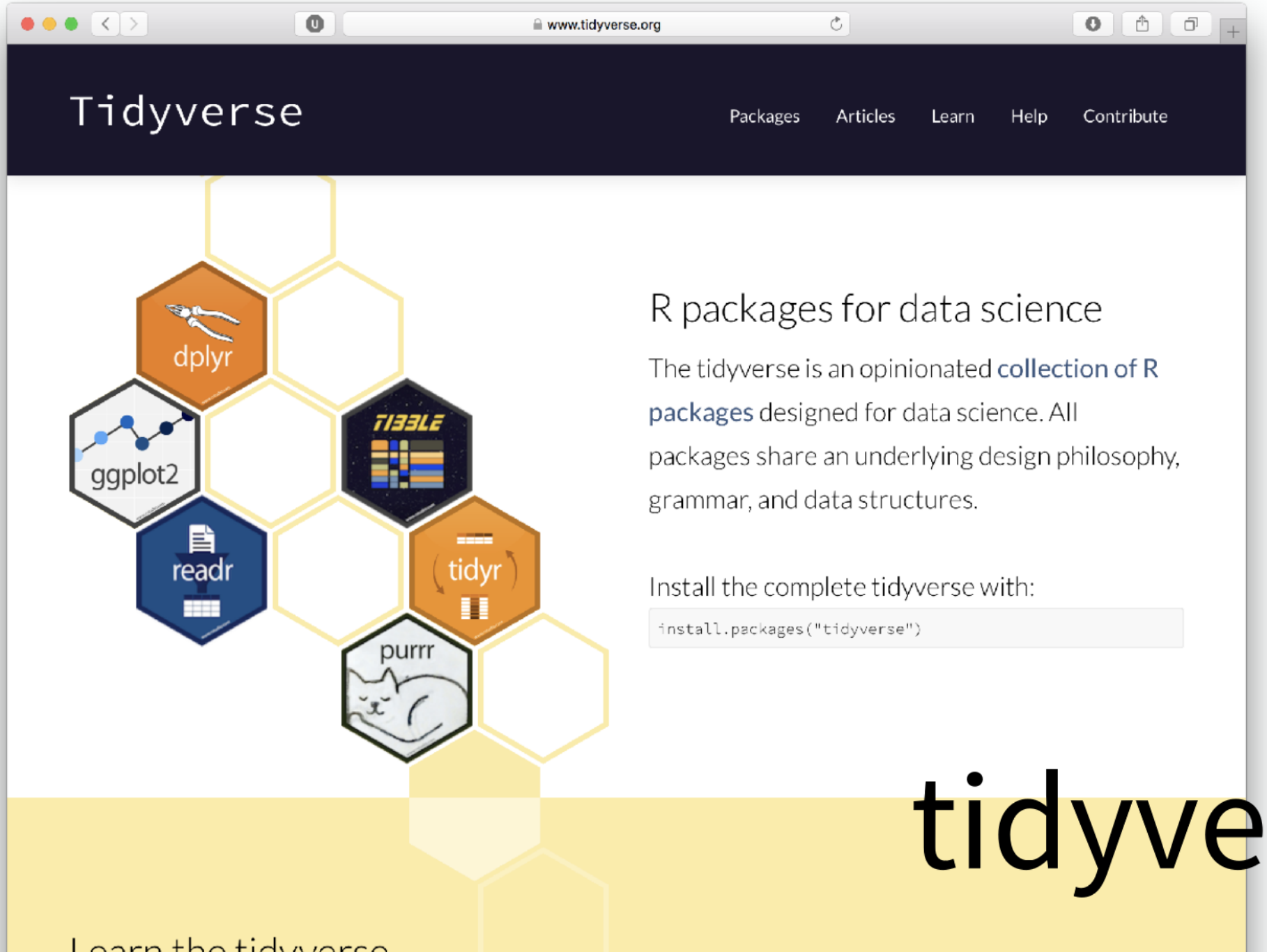


(Applied) **Data Science**



(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")
```

```
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")
```

```
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")
```

Tidy Data: Structuring datasets for analysis

Rectangular tables:
rows x *columns*

country	year	cases	population
Afghanistan	1999	745	19987071
Afghanistan	2000	2666	20595360
Brazil	1999	37737	172006362
Brazil	2000	80488	174504898
China	1999	212258	1272915272
China	2000	213766	1280428583

table1

Columns == Variables

variables

Rows == Observations

observations

Wickham, H. (2014). Tidy Data. *Journal of Statistical Software, 59*(10), 1 - 23.
doi: http://dx.doi.org/10.18637/jss.v059.i10

Getting Started

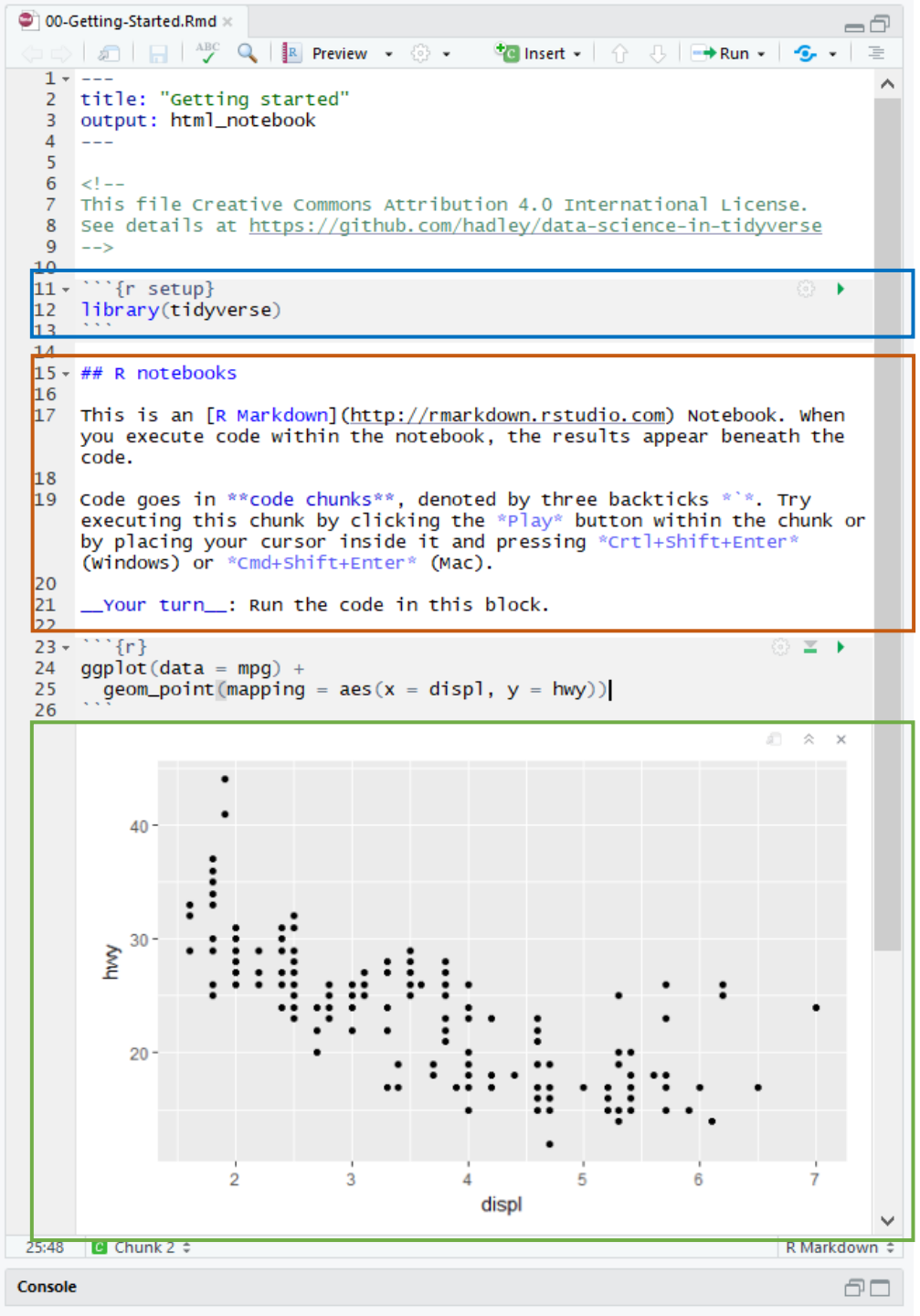
R Notebooks

Authoring format for Data Science

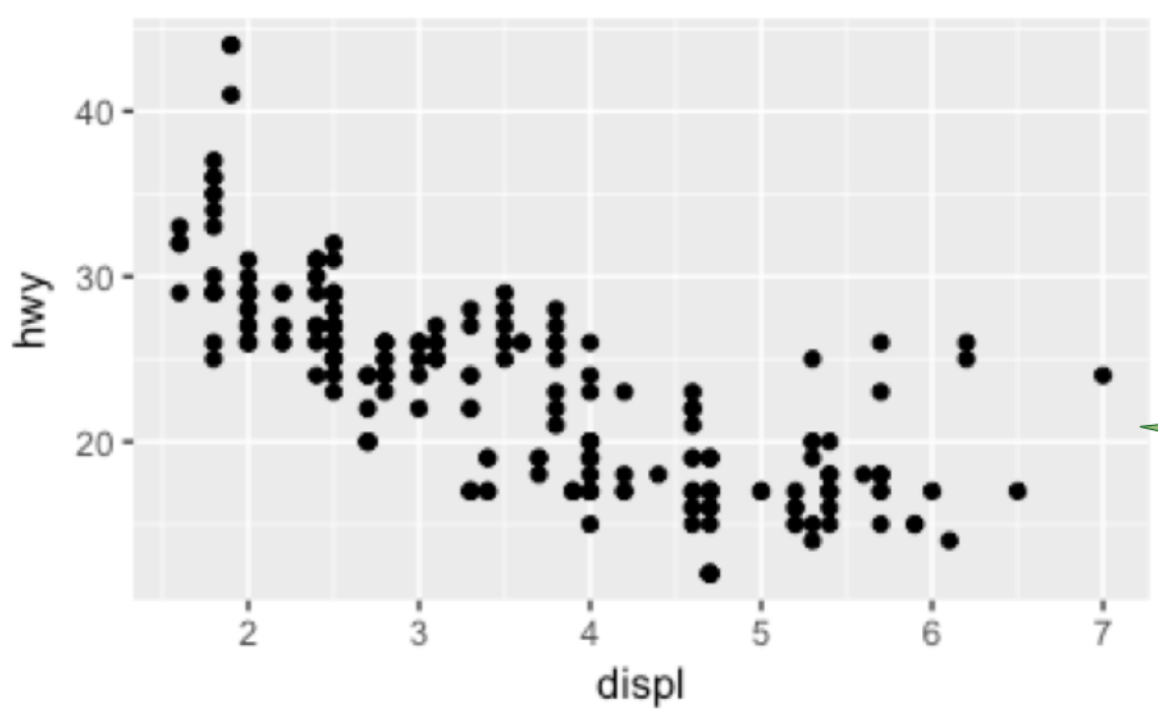
00-Getting-Started.Rmd is an R Notebook

It contains:

- Code
- Text
- Output



```{r}
ggplot(data = mpg) +
  geom_point(aes(x = displ, y = hwy))
```



Click to run code in chunk

Click to run all code chunks above

Code result

Your Turn

From Rstudio run the following (or navigate with the GUI):

```
> setwd("your/filepath/here/r-intro/Day-3")
> file.edit("00-Getting-started.Rmd")
```

Follow the instructions.

Remember your sticky notes!