Modeling with R and the Tidyverse

Matt McDonald
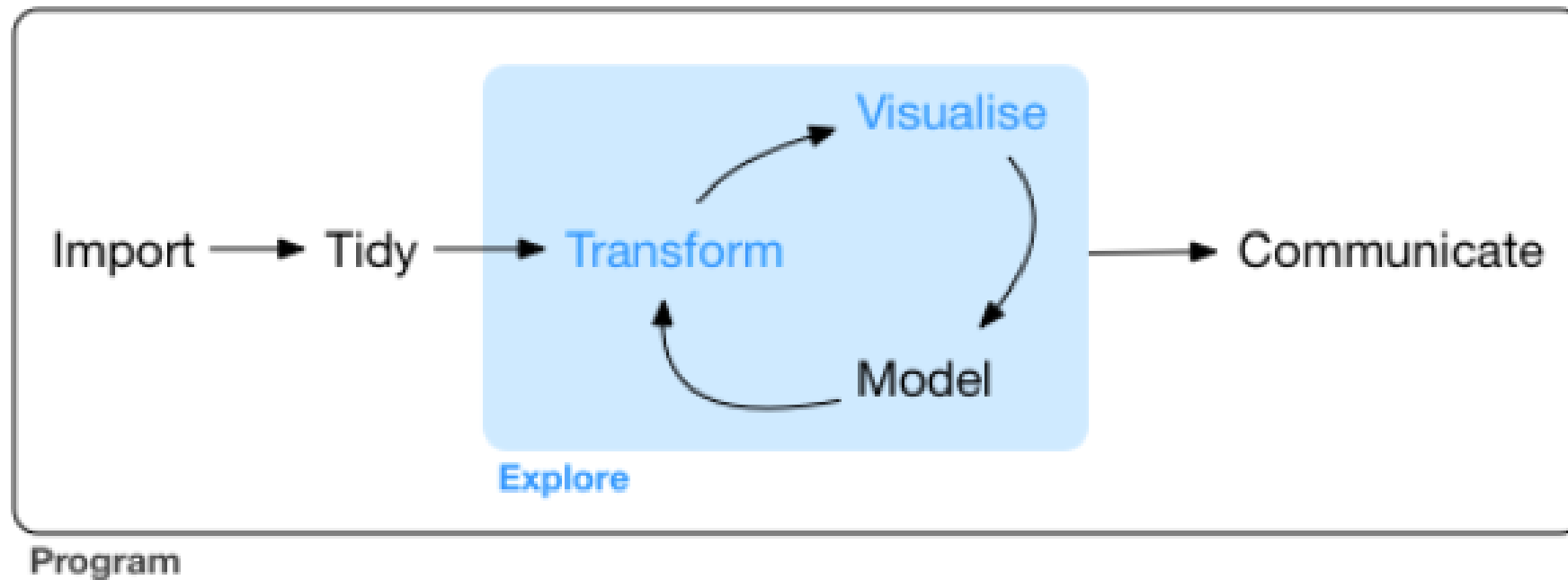
University of Connecticut

FNCE 5352

Materials

- Materials are available on github
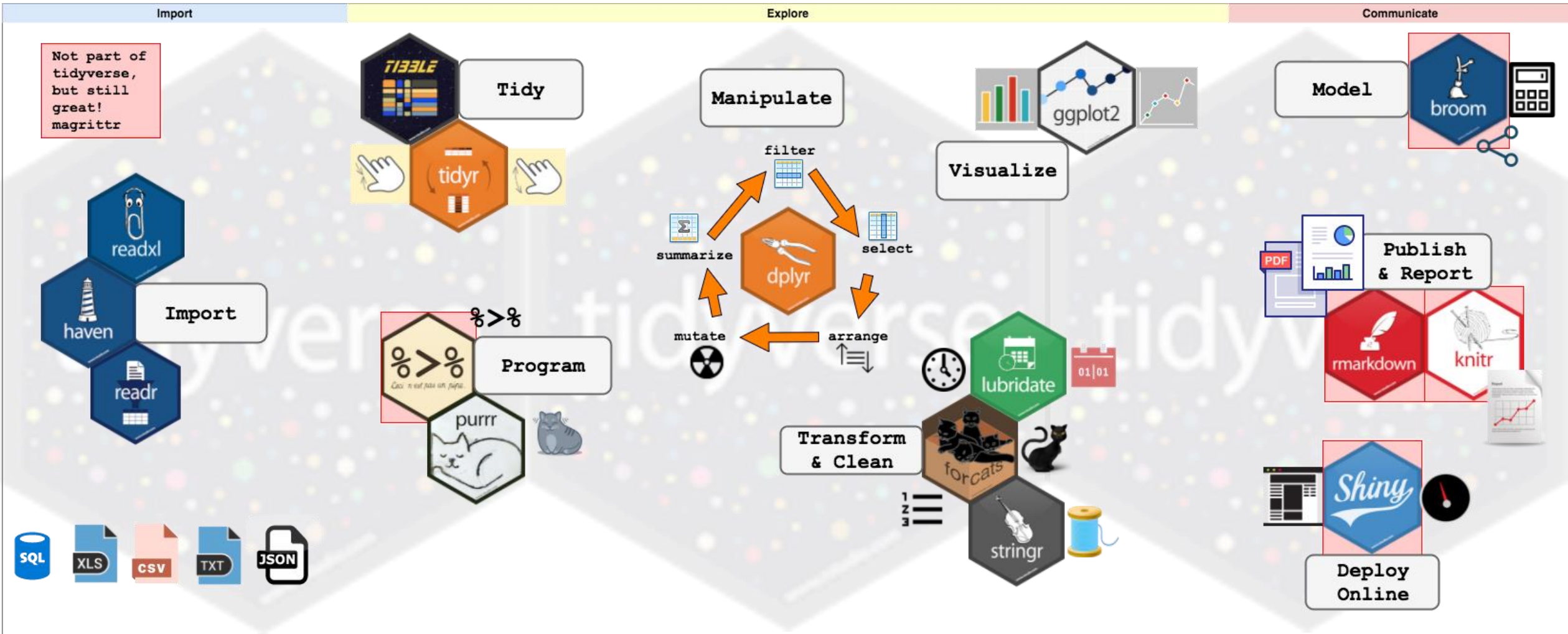 - LINK: https://github.com/mattmcd71/fnce5352_spring2020

Modeling Workflow

Focus of this course

- Model Workflow
- Tools available to support modeling workflow
- Getting things done!

Tidyverse

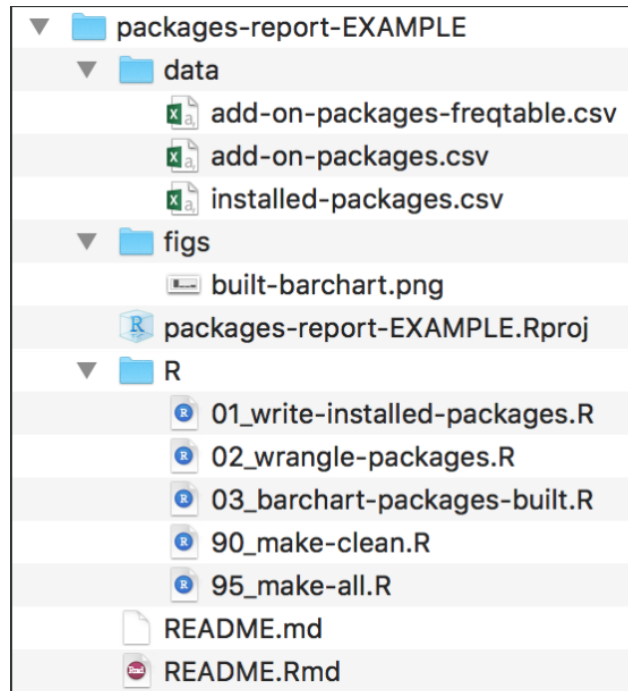# Project-Based Workflow

Why?

- work on more than 1 thing at a time
- collaborate, communicate, distribute
- start and stop

Project-Based Workflow

How?

- dedicated directory
- RStudio **P**roject
- Git repo, probably syncing to a remote

Project-Based Workflow


```
▼ 📁 packages-report-EXAMPLE
   ▼ 📁 data
        📊 add-on-packages-freqtable.csv
        📊 add-on-packages.csv
        📊 installed-packages.csv
   ▼ 📁 figs
        🖼 built-barchart.png
      Ⓡ packages-report-EXAMPLE.Rproj
   ▼ 📁 R
        Ⓡ 01_write-installed-packages.R
        Ⓡ 02_wrangle-packages.R
        Ⓡ 03_barchart-packages-built.R
        Ⓡ 90_make-clean.R
        Ⓡ 95_make-all.R
      📄 README.md
      Ⓡ README.Rmd
```

>>> file salad
+ an out-of-date README

Project-Based Workflow

```
smell-test.R
wrangle.R
model.R           >>>       everything.R
make-figs.R
report.Rmd
```

Help on the Tidyverse

- Rstudio Cheat Sheets!
  - https://www.rstudio.com/resources/cheatsheets/
- R for Data Science
  - http://r4ds.had.co.nz/
- Feature Engineering and Selection
  - https://bookdown.org/max/FES/