

Project Organization

STAT 133

Gaston Sanchez

Department of Statistics, UC–Berkeley

`gastonsanchez.com`

`github.com/gastonstat`

Course web: gastonsanchez.com/teaching/stat133

Introduction

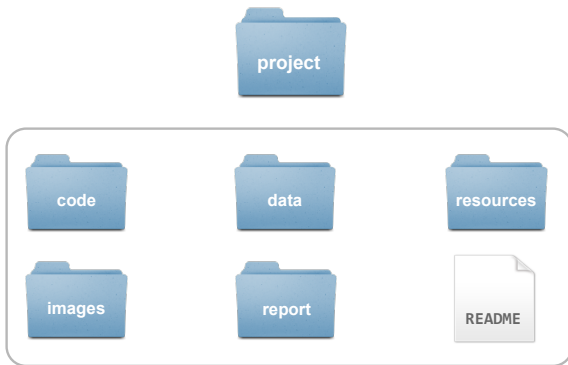
Things around a Research Project

Typically ...

- ▶ There is some interesting problem/phenomenon
- ▶ Giving raise to some research questions
- ▶ Data from experiments, surveys, observations, processes, etc
- ▶ Data cleaning, transformation, processing
- ▶ Exploratory Analysis (summaries, tables, plots)
- ▶ Study associations, relationships
- ▶ Perhaps some data modeling
- ▶ Reporting: white papers, slides, articles, etc

Organizing a Project

Anatomy of a Project (basic)



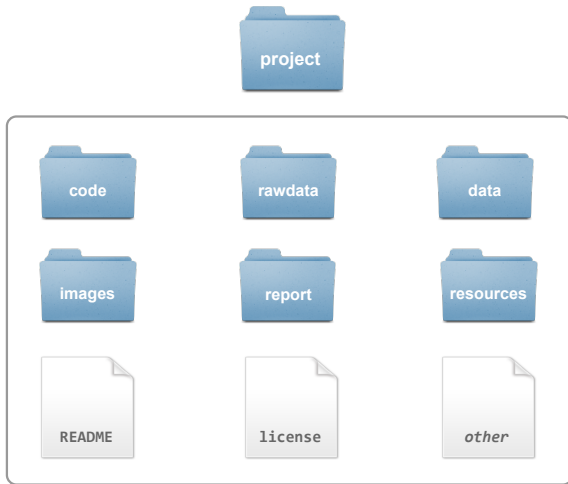
Every project has its own directory

Anatomy of a Project

Project Subdirectories

- ▶ README file: description of the project
- ▶ **code**: functions and scripts
- ▶ **data**: where the data files will live
- ▶ **images** (or figures): images, plots, figures
- ▶ **report**: final report, slides
- ▶ **resources**: articles, references, inspiring things

Anatomy of a Project (more options)



Anatomy of a Project

Project Subdirectories

- ▶ README file: description of the project
- ▶ **code**: functions and scripts
- ▶ **rawdata**: only raw data files
- ▶ **data**: clean data for analysis
- ▶ **images** (or figures): images, plots, figures
- ▶ **report**: final report, slides
- ▶ **resources**: articles, references, inspiring things
- ▶ license file: maybe you need a license
- ▶ other file: other required file?

Anatomy of a Project

code directory

All your scripts, functions, programs go here

rawdata directory

To store original data files. **DON'T touch this!**

data directory

To store cleaned and processed data files.

(these are the ones you use for your analysis, plots, etc)

Anatomy of a Project

images directory

To store all your plots, charts, graphics, illustrations, etc
(ideally produced from your code)

resources directory

References, papers, docs, supporting material, etc
(things that have helped with your project)

report directory

Your final report: exec summary, document, slides, poster, etc

Anatomy of a Project

README file

File describing what your project is about, and other important details (how are the files organized, authors, contact, etc)—It's an *About* file.

License file

Perhaps your project needs a license

e.g. <http://creativecommons.org/>

other file / extra directory

For things that don't fit in any of the previous files/directories

Project Organization

Considerations

- ▶ Get use to organize your projects
- ▶ You can develop your own system:
 - naming style, file-dirs structure
- ▶ Back your projects up
 - e.g. Dropbox, github, cloud storage
- ▶ Check how other people organize their projects

Sharing your Projects

Consider sharing your projects

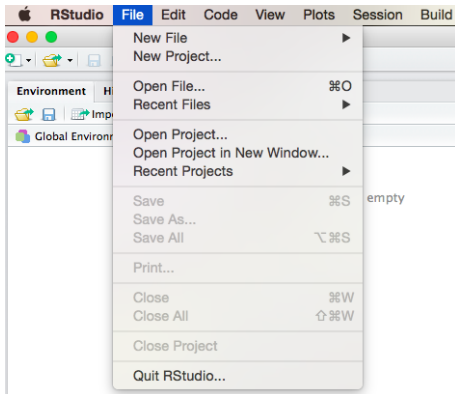
- ▶ Dare to share!
- ▶ Writing Code (scripts, functions, etc) implies a lot of work
- ▶ Most of the time this work never sees the “screen light”
- ▶ In many cases is like writing papers
- ▶ Opportunity to give something back (you’ve benefitted from others’ code)
- ▶ Free peer review
- ▶ Not as bad as it may seem/sound

RStudio Projects

RStudio Project

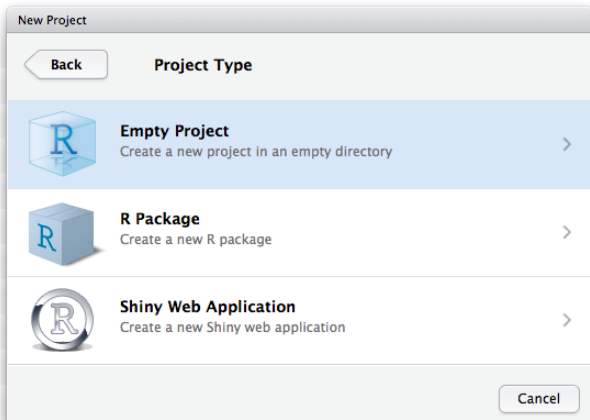
- ▶ You can use RStudio to organize a project
- ▶ RStudio allows you to create *Projects*
- ▶ Can be version-controlled
- ▶ Facilitates working with relative paths

RStudio Project

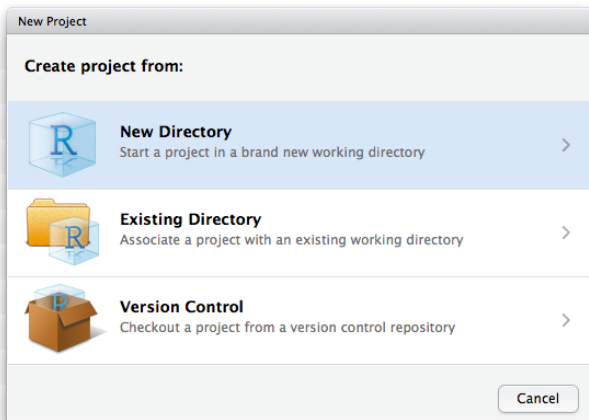


Start a **New Project**

RStudio Project




RStudio Project



RStudio Project

New Project

[Back](#) **Create New Project**



Directory name:

Create project as subdirectory of:
 [Browse...](#)

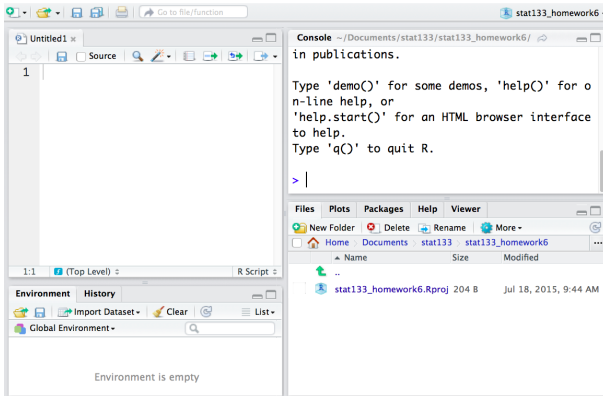
☐ Create a git repository

☐ Use packrat with this project

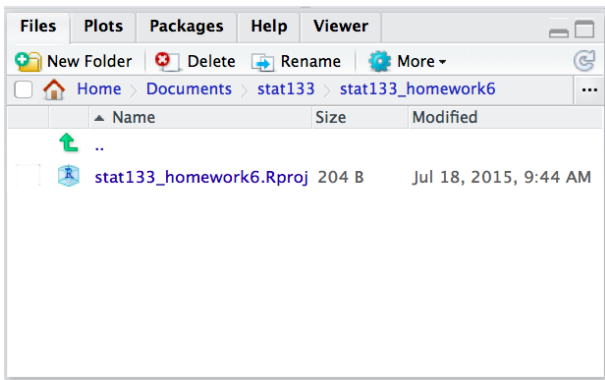
☐ Open in new window

[Create Project](#) [Cancel](#)

RStudio Project

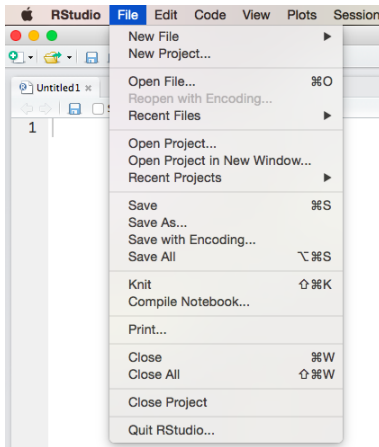


RStudio Project



RStudio associates an .Rproj file to your Project

RStudio Project



To close the Project go to the File menu bar and click **C**lose Project

RStudio Project

Create an RStudio project:

- ▶ Add a data file
- ▶ Add an R script
- ▶ Add an Rmd file
- ▶ Knit the document (either as HTML or pdf)
- ▶ Add a subdirectory

Shell for Windows

For those of you using Windows, you'll need to install either:

- ▶ **Git Bash**

<https://msysgit.github.io/>

- ▶ **PowerShell** (part of the Windows Management Framework 4.0)

<https://www.microsoft.com/en-us/download/details.aspx?id=40855>