https://intro2r.com/ Chapter 1

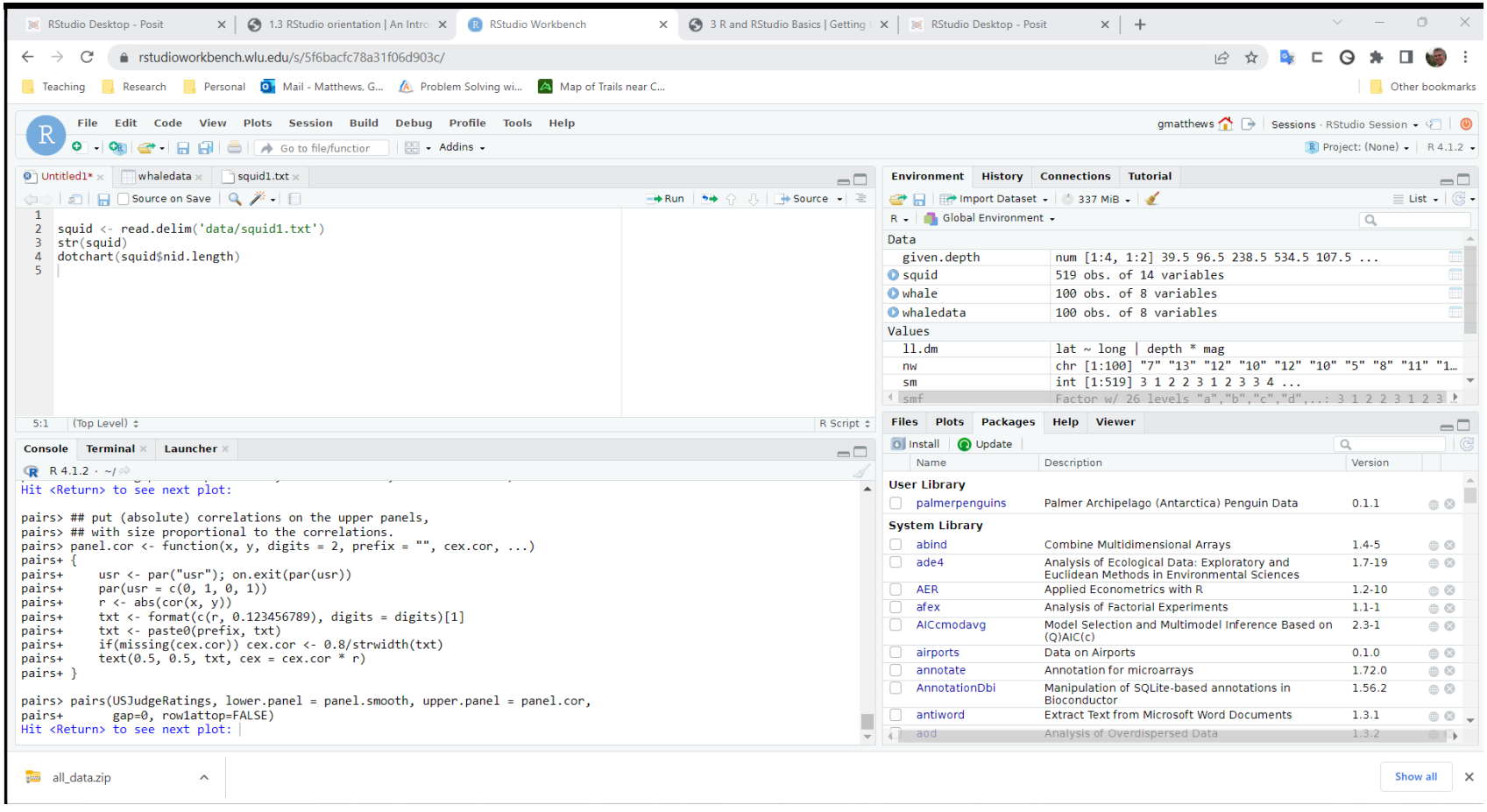
CSCI 297b, Spring 2023

April 19, 2023

Advantages of R

- R is open source and freely available.
- R is available for Windows, Mac and Linux operating systems.
- R has an extensive and coherent set of tools for statistical analysis.
- R has an extensive and highly flexible graphical facility capable of producing publication quality figures.
- R has an expanding set of freely available 'packages' to extend R's capabilities.
- R has an extensive support network with numerous online and freely available documents.

RStudio panels



RStudio panels left side

- Source window
  - Enter and edit commands for the R interpreter.
  - Not sourced into the R interpreter until you hit Source button.
  - Can be sourced with control+enter (windows) or command+enter (mac)
- Console/Terminal/Launcher window
  - Type commands directly into the R interpreter.
  - Use as a calculator

RStudio panels right side

- Environment/History/Connections/Tutorial window
  - Environment tells you what names you have defined.
  - History gives list of commands you've entered in console.
- Files/Plots/Packages/Help/Viewer window
  - File browser
  - Plot viewer
  - Help viewer
  - Web viewer

RStudio console and source

- Mostly we use editor and console
- Can enter commands in console for quick answers
- Almost always better to enter commands into an R script
- Save your R scripts with a `.R` extension
- R scripts are plain text files that can be edited with any text editor

R packages

- Large collection already installed on server.
- CRAN packages:

```
install.packages('remotes', dependencies=
    TRUE)
```

- Using packages:

```
library(remotes)
```

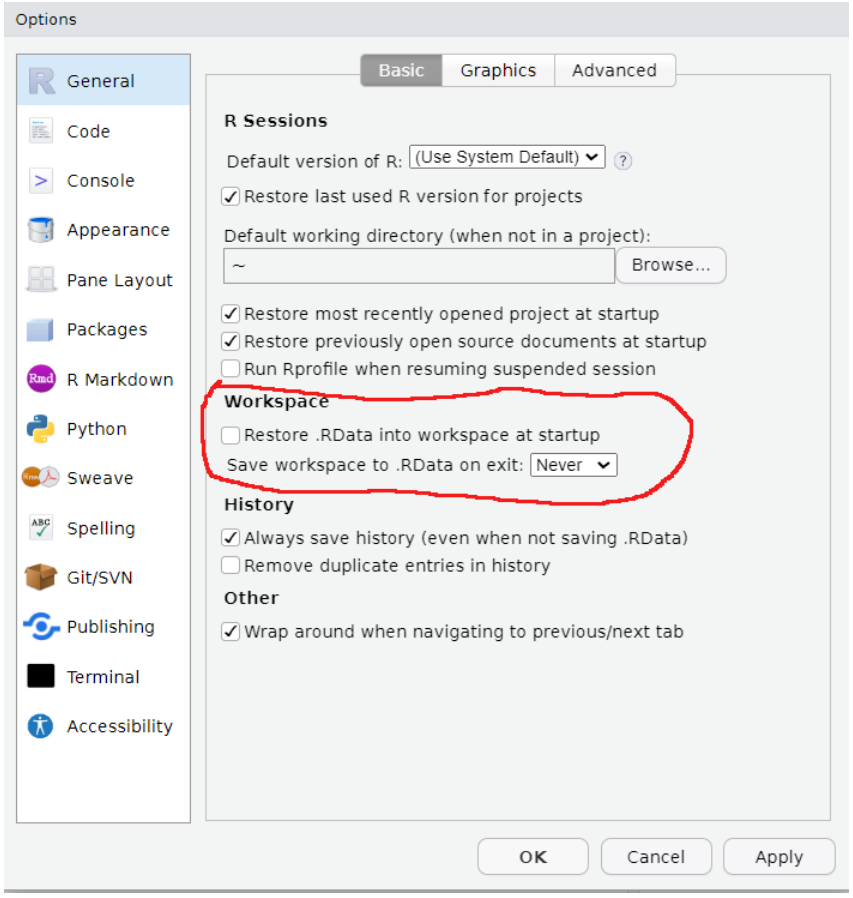
- GitHub packages

```
library(remotes)
install_github('tidyverse/dplyr')
```

- You can install additional packages on your own computer.
- You can install additional packages on the server, if they are just for your own use.

Global options

- Tools → Global Options ...
- Turn off `.RData` save and restore
- This prevents new R sessions from being influenced by previous R sessions.



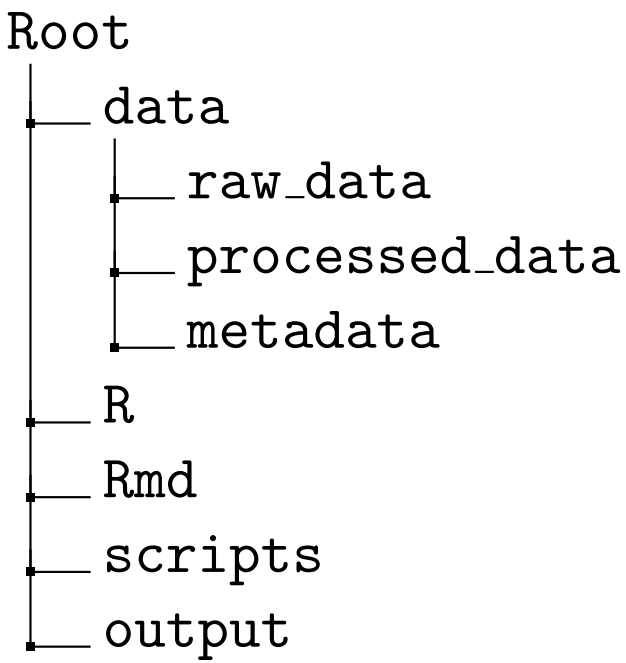
R projects

- Remembers settings, history, *etc.*
- Organizes all files, scripts, data, *etc.*
- File → New Project
- Usually use a new directory
- Create Project

Working directory

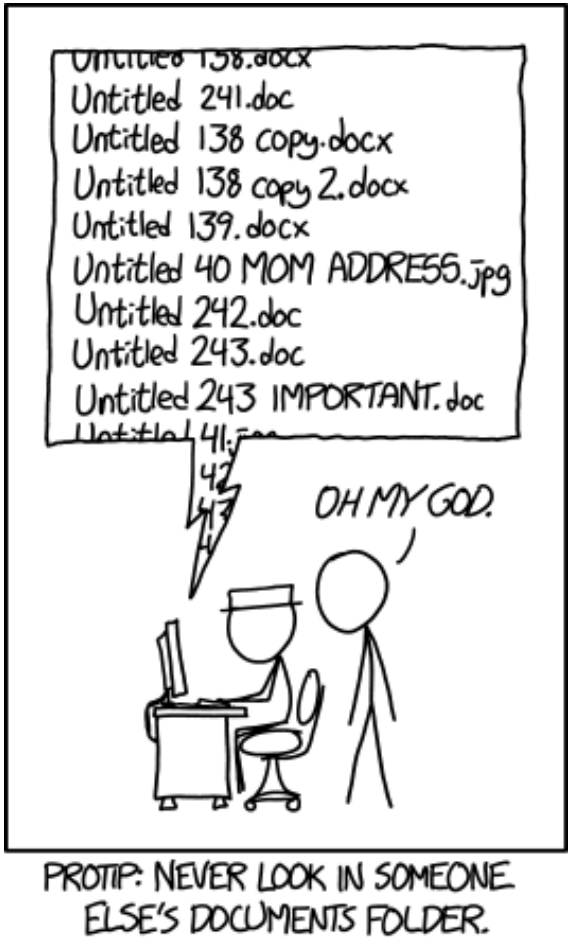
- Current working directory seen at top of console window
- Default folder R looks in when loading scripts, data, etc.
- Opening a project file automatically sets working directory
- Can use `getwd()` to identify working directory
- Can use `setwd()` to change working directory

Directory structure



- `Root` holds the `.Rproj` file
- `raw_data` holds data from the field that should not be edited
- `processed_data` holds data that has been cleaned up
- `metadata` holds text files describing how the data was collected, etc.
- `R` holds R functions of general utility
- `Rmd` holds Rmarkdown files
- `scripts` holds current work
- `output` holds plots, data summaries, etc.

File names



- Short
- Informative
- No spaces
- No special characters
  `!@#$%^&*()`
- Pad with zeros
  `file001 file002 file003`
- Use ISO 8601 date format
  `YYYY-MM-DD or YYYYMMDD`

R script meta information

```
# Title: Time series analysis of snouters

# Purpose : This script performs a time series analyses on
#           snouter count data.
#           Data consists of counts of snouter species
#           collected from 18 islands in the Hy−yi−yi
#           archipelago between 1950 and 1957.
#           For details of snouter biology see:
#           https://en.wikipedia.org/wiki/Rhinogradentia

# Project number: #007

# DataFile:'data/snouter_pop.txt'

# Author: A. Nother
# Contact details: a.nother@uir.ac.uk

# Date script created: Mon Dec 2 16:06:44 2019 ————————
# Date script last modified: Thu Dec 12 16:07:12 2019 ——

# package dependencies
library(PopSnouter)
library(ggplot2)

print('put your lovely R code here')

# good practice to include session information

xfun::session_info()
```

Style guide

https://google.github.io/styleguide/Rguide.html

Citing R

```
> citation()

To cite R in publications use:

  R Core Team (2021). R: A language and environment for
  statistical computing. R Foundation for Statistical
  Computing, Vienna, Austria. URL https://www.R-project.org/.

A BibTeX entry for LaTeX users is

  @Manual{,
    title = {R: A Language and Environment for Statistical Computing},
    author = {{R Core Team}},
    organization = {R Foundation for Statistical Computing},
    address = {Vienna, Austria},
    year = {2021},
    url = {https://www.R-project.org/},
  }

We have invested a lot of time and effort in creating R,
please cite it when using it for data analysis. See also
'citation("pkgname")' for citing R packages.
```

Exercise 1