

Best practices

DAY 2, SESSION 5

September 18, 2023



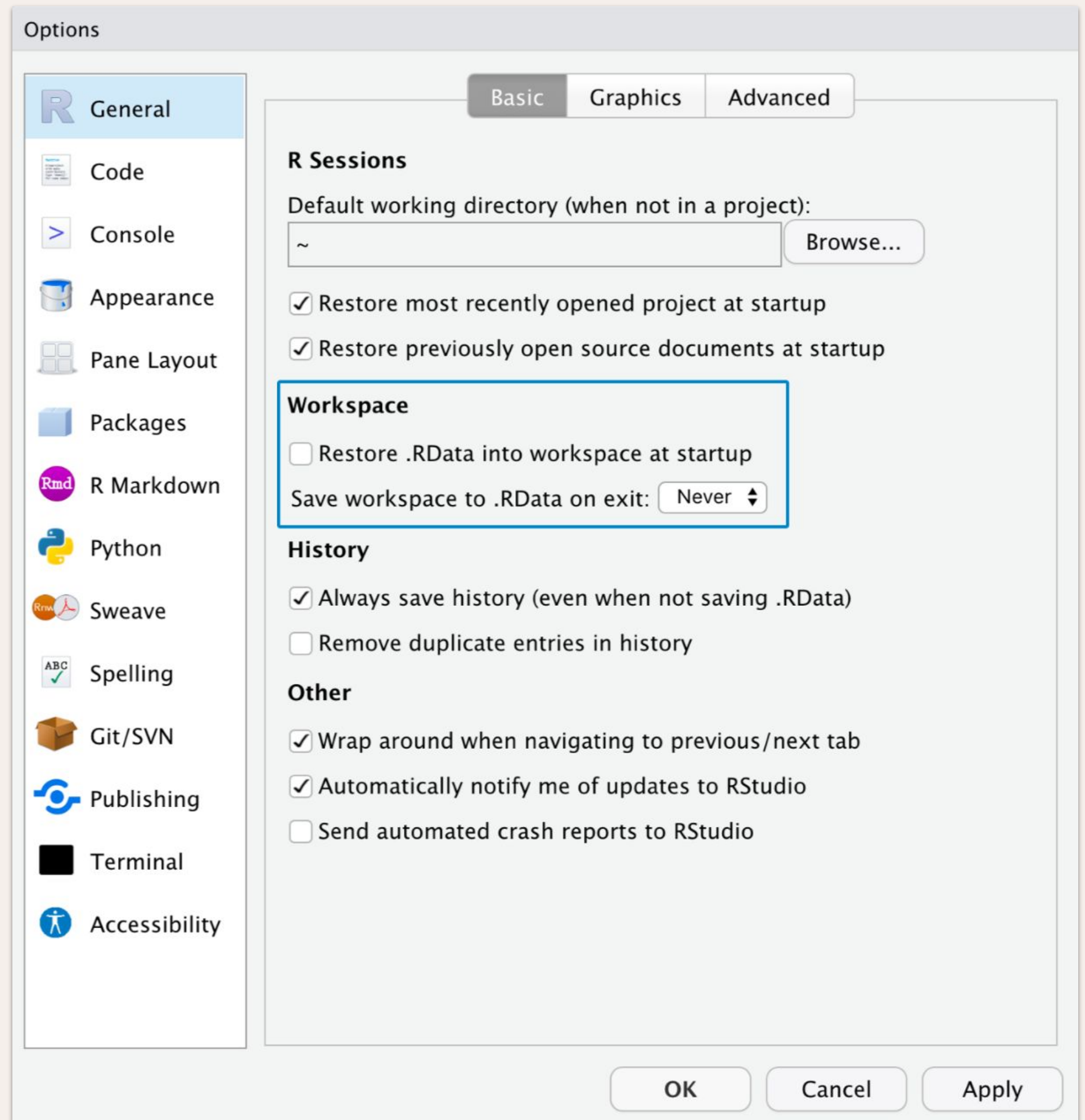
Agenda

1. Customize your workspace
2. Project-oriented workflows
3. Naming files
4. Reproducible environments
5. Code style
6. Version control
7. Asking for help

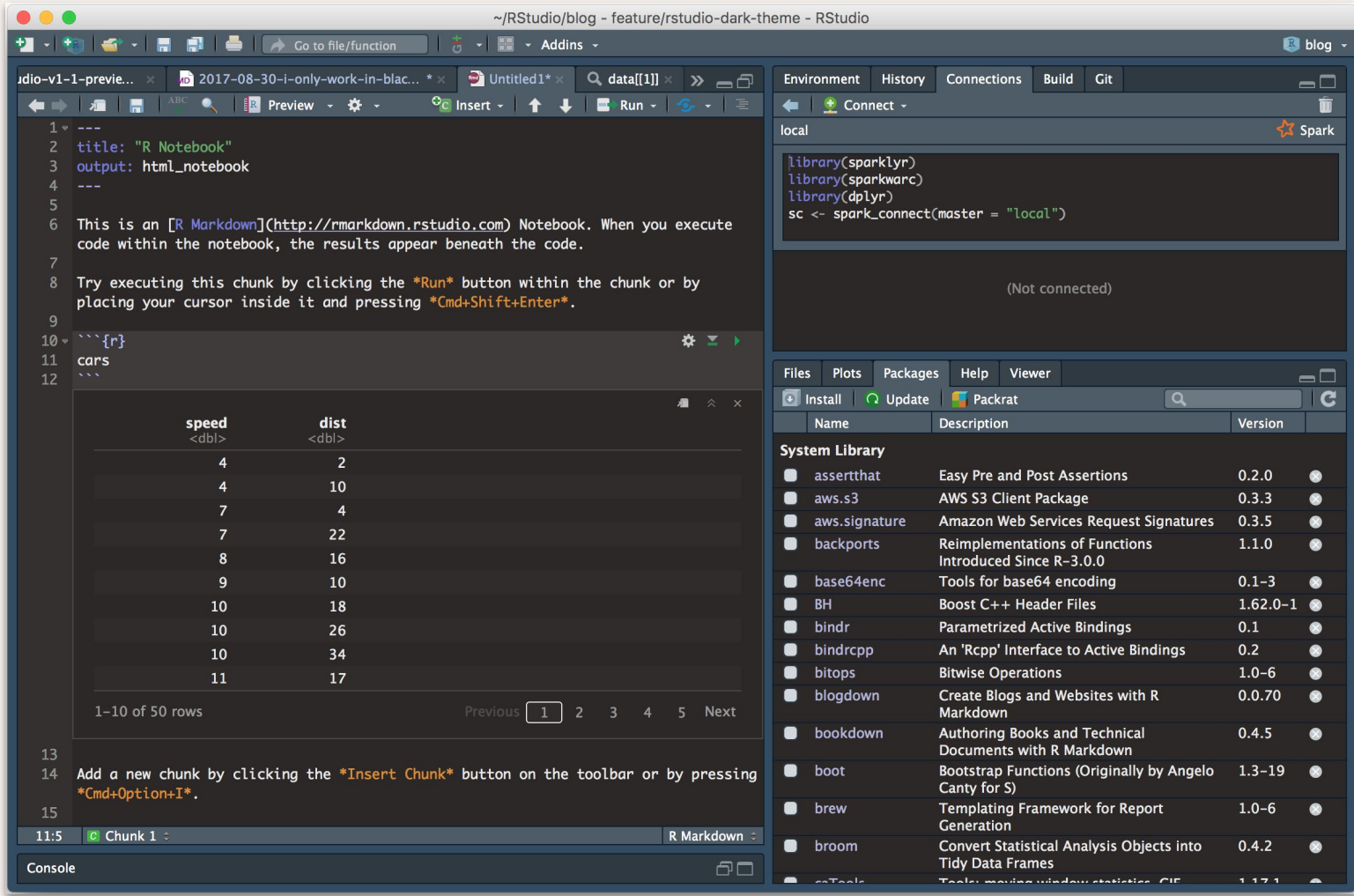
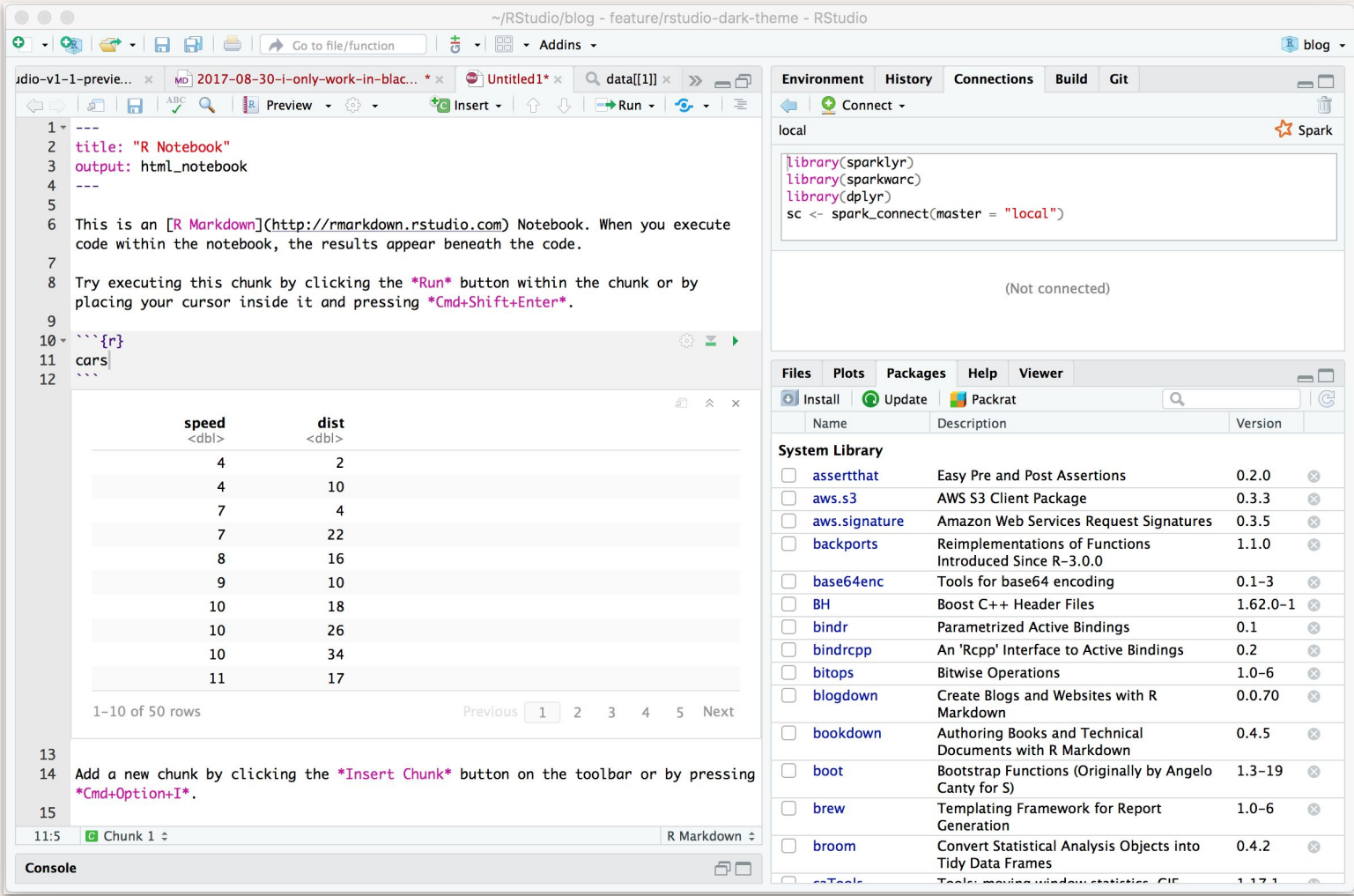
1. Customize your workspace

IDE settings

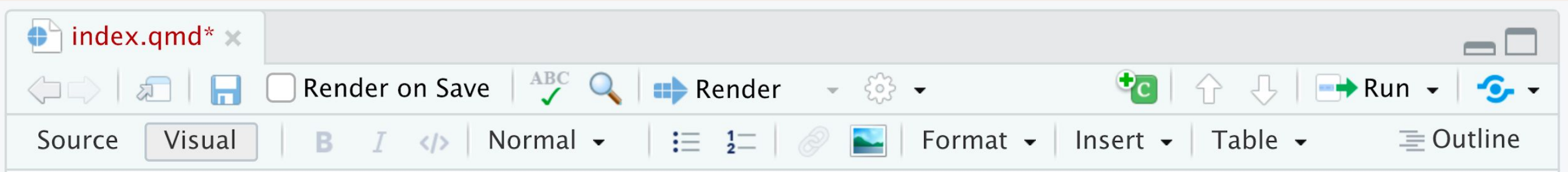
Tools > Global Setup



Themes



Visual editor



relational-data.qmd

Render on Save

Render

Visual

Normal

Format

Insert

Table

Outline

#filtering-joins

Filtering Joins

Filtering joins match observations in the same way as [mutating joins](#), but affect the observations, not the variables. There are two types:

| | | |
|------------------------------|---------------|---|
| <code>semi_join(x, y)</code> | $x \ltimes y$ | Keeps all observations in <code>x</code> that have a match in <code>y</code> |
| <code>anti_join(x, y)</code> | $x \rhd y$ | Drops all observations in <code>x</code> that have a match in <code>y</code> |

Graphically, a semi-join looks like this:

```
{r::join-semi}~
#l-echo:~false~
#l-out-width:~"6"~
~
knitr::include_graphics("diagrams/join-semi.png")
```

Only the existence of a match is important; it doesn't matter which observation is matched. This means that filtering joins never duplicate rows like mutating joins do:

Chunk 1: join-semi

Quarto

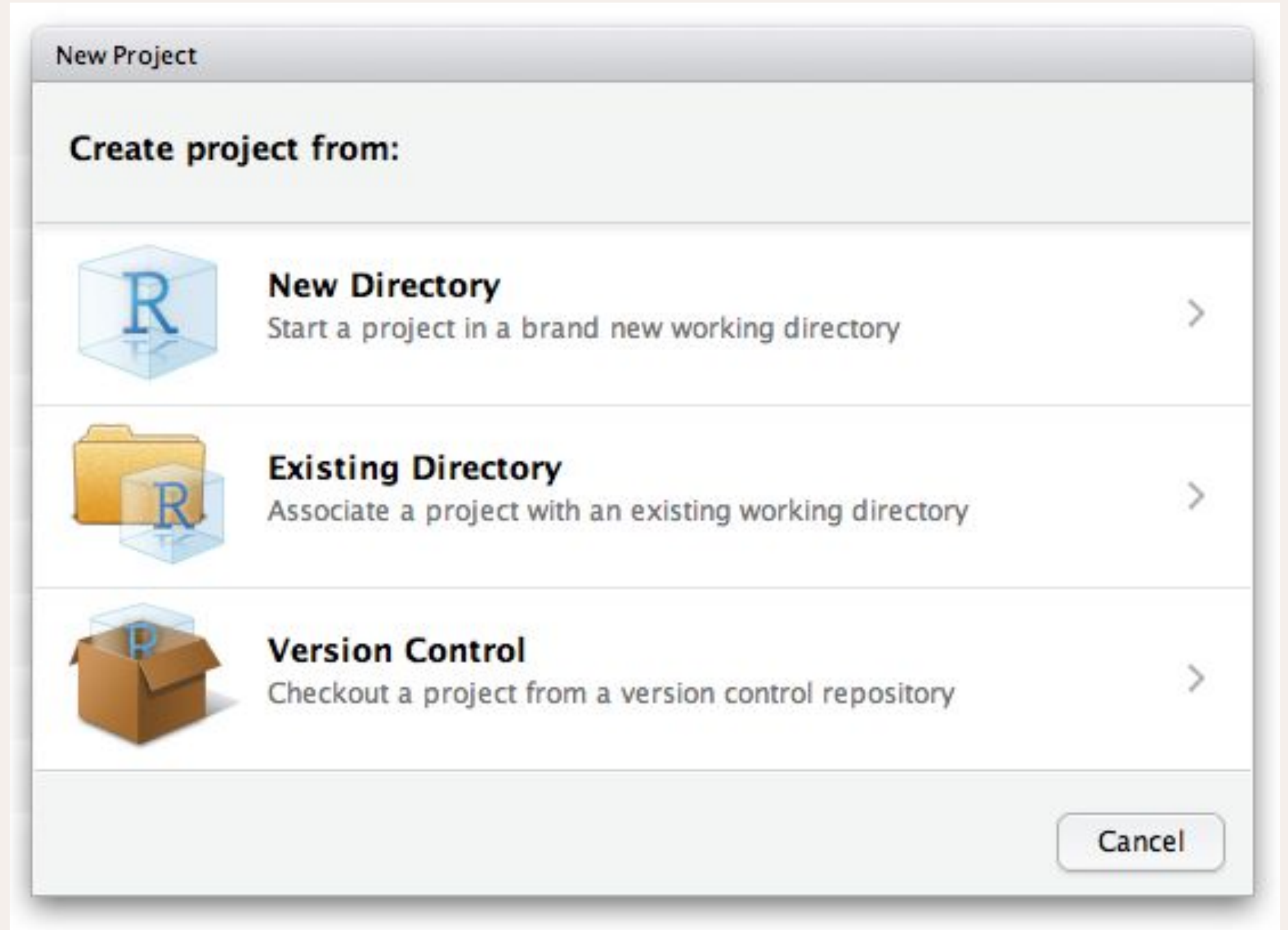
Resources

- [Customizing the RStudio IDE](#)
- [Using themes in the IDE](#)
- [Rainbow parentheses](#)
- [Visual editing in RStudio](#)

2. **Project-oriented workflows**

Projects

Use R projects to make your code more organized and shareable



here

Use the here package to
manage file paths



✗ `setwd("path/that/only/exists/on/my/computer")`

Resources

- [What They Forgot to Teach You About R Chapter 3: Project-oriented workflow](#)
- [R4DS Chapter 7: Script and Projects](#)
- The [here package](#)

3. **Naming files**

Naming files

- Machine readable
- Human readable
- Sorts/orders nicely

NO

myabstract.docx

Joe's Filenames Use Spaces and Punctuation.xlsx

figure 1.png

fig 2.png

JW7d^(2sl@deletethisandyourcareerisoverWx2*.txt

YES

2014-06-08_abstract-for-sla.docx

joes-filenames-are-getting-better.xlsx

fig01_scatterplot-talk-length-vs-interest.png

fig02_histogram-talk-attendance.png

1986-01-28_raw-data-from-challenger-o-rings.txt

Image from [Jenny Bryan](#)

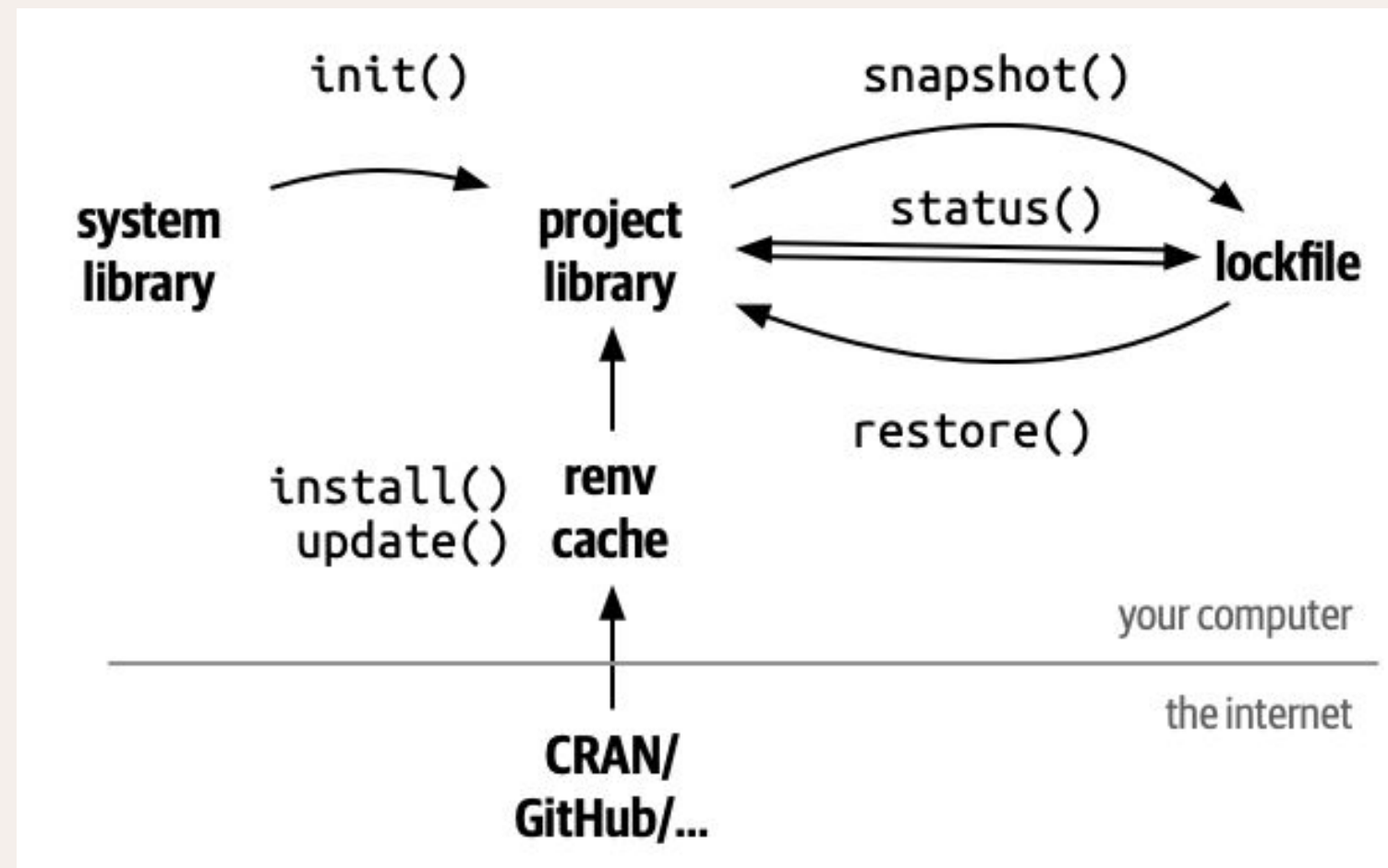
Resources

- [Naming things](#) (slides)
- [Tidyverse Style Guide Chapter 1: Files](#)

4. Reproducible environments

renv

Create reproducible
environments for your R
projects with the renv package



Resources

- The [renv package](#)
- [You should be using renv](#) (rstudio::conf(2022) talk)

5. **Code *style***

Code style

Consistent style makes your
easier to read & debug by your
future self and collaborators

Good

```
iris %>%  
  group_by(Species) %>%  
  summarize_if(is.numeric, mean) %>%  
  ungroup() %>%  
  gather(measure, value, -Species) %>%  
  arrange(value)
```

Bad

```
iris %>% group_by(Species) %>% summarize_all(mean) %>%  
ungroup %>% gather(measure, value, -Species) %>%  
arrange(value)
```

Image from the [Tidyverse Style Guide](#)

Resources

- [Tidyverse style guide](#)
- The [styler package](#)
- [R4DS Chapter 5: Code Style](#) (short primer)

6. **Version control**

Git & Github

Use version control to share and collaborate on code projects



Artwork by Allison Horst: <https://allisonhorst.com/git-github>

Resources

- [Happy Git With R](#)

7. Asking for help

Tips for searching

- Google the error message!
- Add “R” to your query
- Add a package name (e.g. “dplyr”) to narrow your results
- Try [Posit Community](#) and [Stack Overflow](#)

Reprex

Create a **reproducible example** when asking for help



Resources

- [R4DS Chapter 9: Getting Help](#)
- The [reprex package](#)

Questions?