

# Lecture 7. Data Manipulation

Marcy Shieh

October 22, 2020

# Goals

Today, you will learn how to:

- ▶ Read data
- ▶ Use tidyverse for data manipulation

# Why do you need to know how to data manipulate data?

- ▶ Very few datasets are perfect because many of them have. . .
  - ▶ variable names need to be renamed
  - ▶ values that need to be recoded
- ▶ You may want to merge datasets
- ▶ You may want to create graphs that *look a certain way*

## here() vs. setwd()

- ▶ `here()` is helpful because you can put your project on GitHub, someone can download it, and run all the code on their computer without changing anything
- ▶ `setwd()` has been getting a bad rep lately in the world of R Projects but if you're doing data exploration work... for fun...? and don't plan to have it see the light of day ... then `setwd()` should be fine

# Think of Tidyverse as Excel

- ▶ Tidyverse is kinda like Excel but it's not point and click
- ▶ This comes with advantages, e.g., you can handle bigger datasets
  - ▶ Whenever you open large datasets in Excel, your computer will freeze, you'll need to force quit Excel, and life will suck lol
- ▶ You can filter, select, transpose, use pivot tables, etc. but Tidyverse has different names for those commands