Week 4 Day 2: Lab Warm-up

## Lab 3 Debrief

1. Why do you need to use a .x inside ~ mean() for the across() function?
2. What does it mean for a function to be “superseded”?
3. What does it mean for a function to be “deprecated”?
4. How would you know if the function you are using is superseded or deprecated?
5. What should you do if the function you are using is superseded or deprecated?

## Relational Data

Do counties with higher property taxes also have higher childcare costs? Has this relationship changed over time?



### inner\_join()

Keeps observations when their keys are present in **both** datasets.

inner\_join(x = counties,
 y = ca\_tax\_rev,
 by = join\_by(county\_name == entity\_name)
 )

1. Why am I using join\_by() to join these two datasets?
2. What counties will remain in the joined dataset?

### left\_join()

Keeps only (and all) observations present in the left data set

left\_join(x = counties,
 y = ca\_tax\_rev,
 by = join\_by(county\_name == entity\_name)
 )

1. What counties will remain in the joined dataset?

### right\_join()

Keeps only (and all) observations present in the right data set

right\_join(x = counties,
 y = ca\_tax\_rev,
 by = join\_by(county\_name == entity\_name)
 )

1. What counties will remain in the joined dataset?

### NAs in Joins

1. What might be the cause of NA values when joining two datasets together?

## Filtering Joins

Filter *observations* based on values in new dataframe.

### anti\_join()

Removes observations when their keys are present in both datasets, and only keeps variables from the first dataset

military |>
 anti\_join(cont\_region,
 by = join\_by(Country == Region))

1. How is this approach different from what we did in the Practice Activity?

## Factor Variables

1. What are the two general purposes for factor variables (in R)?
2. What function have you used previously to create a factor variable? What function inside the **forcats** package does this?
3. Can you specify the order you want the levels (of a factor) to be in?

### fct\_reorder()

ca\_tax\_rev <- ca\_tax\_rev |>
 mutate(entity\_name = fct\_reorder(.f = entity\_name,
 .x = sales\_and\_use\_taxes,
 .fun = mean)
 )

1. What does this code do?
2. In what setting would you want to use code like this?

### fct\_reorder2()

1. What does this function do?
2. When would you want to use it?

### fct\_collapse()

1. What does this function do?
2. When would you want to use it?

## Code Formatting

Don’t forget, writing “tidy” and “well documented” code are two of the learning targets for this course. As such, I would strongly encourage you to use every opportunity to practice these skills.

As you are writing code for this assignment, make sure your code follows the [tidyverse style guide for dplyr code](https://style.tidyverse.org/pipes.html). Specifically, your code should:

* use whitespace liberally
	+ before & after every = sign
	+ after every ,
	+ before every |>
* use new lines liberally
	+ after every |>
	+ after , when needed (if code is more than 80 characters in length)