Coding together week 5

Warm-up

- Open the coding together book and change it to dark mode, and change the font to serif
- Download the PDF of the book
- Find the part of the book describing `mutate()`: hint it's section 2.3.4
- Find the tidyr cheatsheet on the internet

pivot_longer()

Using an example from the tidyr website:

relig_income is a dataset that comes with tidyr that contains observations of the income bands for 18 religons.

Inspect it by typing: relig_income and pressing enter

Let's consider the income bands that are column headings as a variable. Pivot them into a single variable called "income" with the values as a variable called "count". Don't use the religon variable.

```
relig_income %>%
        pivot_longer(names_to = "income", values_to = "count" , -religion)
```

pivot_wider()

Using an example from the tidyr website:

fish_encounters is another dataset that come with tidyr contributed by Myfanwy Johnston, describes when fish swimming down a river are detected by automatic monitoring stations.

It only records when fish are detected, and not when they aren't. It has three variables: fish, station and seen

Pivot the table wider such that the station become variable names and the values in the new columns are from seen.

Pivot the table again, but this time fill in the missing values with zeros. Use ?pivot_wider to find out how to do this. Or google.

```
fish_encounters %>%
        pivot_wider(names_from = station, values_from = seen)

fish_encounters %>% pivot_wider(
  names_from = station,
  values_from = seen,
  values_fill = list(seen = 0)
)
```

Missing values

```
stocks <- read_csv("https://raw.githubusercontent.com/ab604/coding-together/master/exercises/stocks.csv
```

- Check if `stocks` has explicit missing values:

```
stocks %>% map_dfr(~ sum(is.na(.)))
```

- Check if `stocks` has implicit missing values

Complete

Use `complete()` to make implicit missing values in `stocks` explicit

```
stocks %>% complete(year,qtr)
```

Fill

Building on the last answer use `fill()` to replace `NA` in return with last non-missing value Does this make sense?

```
stocks %>% complete(year,qtr) %>% fill(return)
```

Coalesce

Complete the table again, but this time use `mutate` with `coalese()` to fill in the mean `return` value.

```
stocks %>%
        complete(year,qtr) %>%
        mutate(return = coalesce(return, mean(return, na.rm = TRUE)))
```

Joins

I've created a weather table from Portal data that has the average temperature and rainfall. Try out joins wit surveys_subset.

```
weather <- read_csv("https://raw.githubusercontent.com/ab604/coding-together/master/exercises/portal-wea
```

Inner join.

```
surveys_subset %>% inner_join(weather, by = "date")
```