Coding challenge 5: data wrangling

Sarah Batten

2025-03-20

Coding challenge 5

Question 1

Reading the data into R:

```
diversity <- read.csv("DiversityData.csv")
meta <- read.csv("Metadata.csv", na.strings = "na")
```

Question 2

Joining dataframes together:

```
library(tidyverse) # loading required package
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.0.4
## -- Conflicts ---------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become error
```

```
alpha <- left_join(diversity, meta, by = "Code")
```

Question 3

Finding Pielou's evenness:

```
alpha_even <- alpha %>%
  mutate(even = shannon/log(richness))
```

Question 4

Finding the mean and standard error evenness grouped by crop over time:

```
alpha_average <- alpha_even %>%
  group_by(Crop, Time_Point) %>%
  summarise(Mean.even = mean(even),
            n = n(),
            sd.dev = sd(even)) %>%
  mutate(std.error = sd.dev/sqrt(n))
```

```
## `summarise()` has grouped output by 'Crop'. You can override using the
## `.groups` argument.
```

Question 5

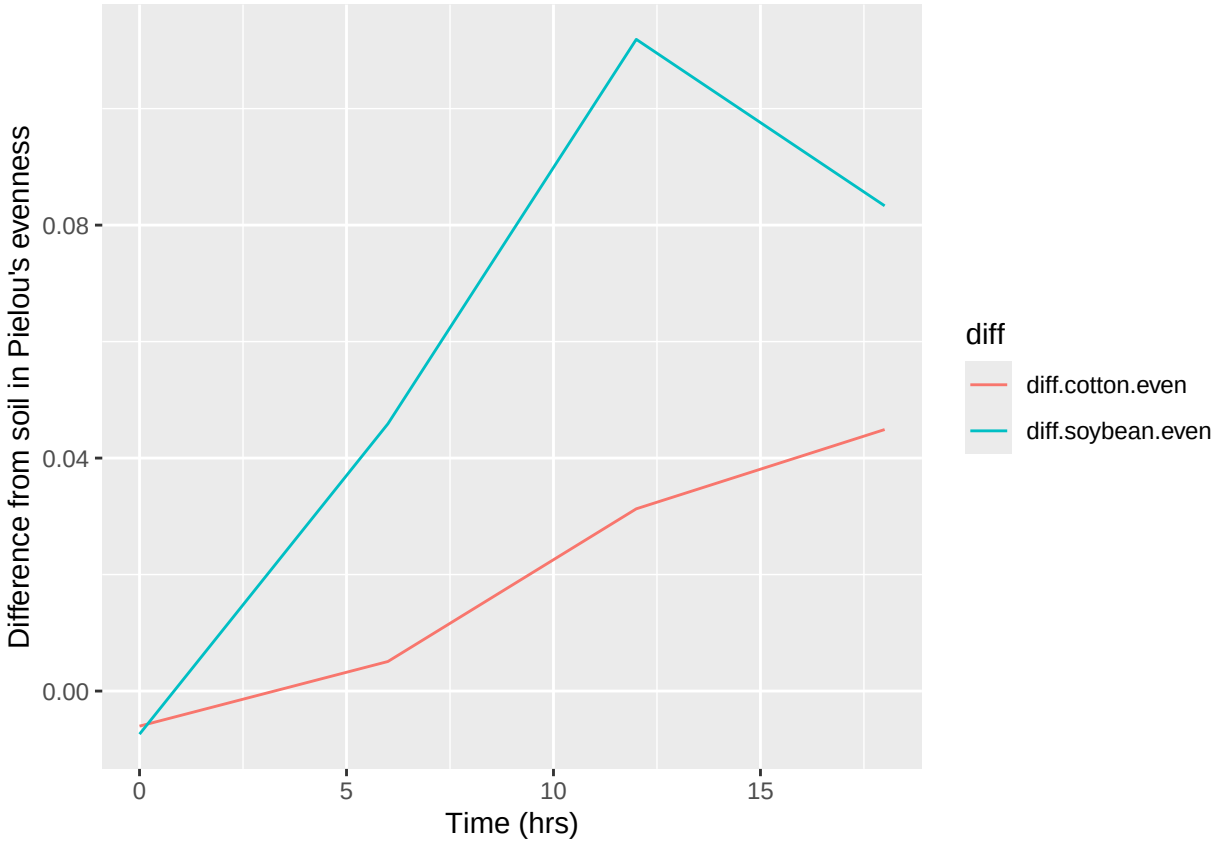
Calculating the difference between the soybean column, the soil column, and the difference between the cotton column and the soil column.

```
alpha_average2 <- alpha_average %>%
  select(Time_Point, Crop, Mean.even) %>%
  pivot_wider(names_from = Crop, values_from = Mean.even) %>% # pivot to wide format
  mutate(diff.cotton.even = Soil - Cotton) %>%
  mutate(diff.soybean.even = Soil - Soybean)
```

Question 6

Making plots:

```
alpha_average2 %>%
  select(Time_Point, diff.cotton.even, diff.soybean.even) %>%
  pivot_longer(c(diff.cotton.even, diff.soybean.even), names_to = "diff") %>%
  ggplot(aes(x = Time_Point, y = value, color = diff)) + # changing color to 'diff' variable
  geom_line() +
  xlab("Time (hrs)") +
  ylab("Difference from soil in Pielou's evenness")
```



Question 7

Link to GitHub: Coding Challenge 5 Directory