ENVS 193DS Homework 3

Jess Rousselle

2025-05-30

#### Setup

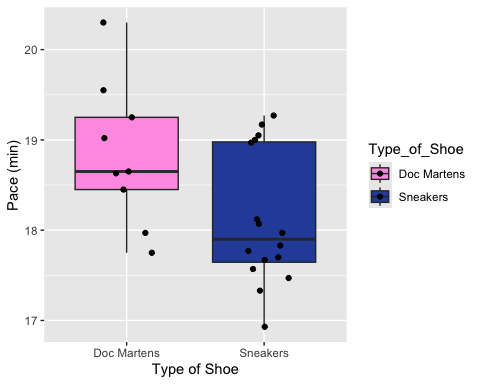
#loading in packages  
library(tidyverse)  
library(here)  
library(flextable)  
library(janitor)  
library(dplyr)

### A.

I could calculate the central tendencies for walking pace for both categories of shoe that I wear (docs or sneakers) to see if my shoes affect my walking pace. I think that docs make me walk slower since they are less comfortable.

### B.

setwd('/Users/jessrousselle/Desktop/github/ENVS-193DS\_homework-03')  
my\_data <- read\_csv("data/my\_data .csv") #reading in data  
ggplot(data = my\_data, #setting data frame to my\_data  
 aes(x = Type\_of\_Shoe, #setting x axis to type of shoe  
 y = Pace\_min,  
 fill = Type\_of\_Shoe)) + #setting y axis to pace  
 scale\_fill\_manual(values=c("#FF9FE5", "#2B50AA")) +   
 geom\_boxplot() + #creating a box plot  
 geom\_jitter(width = 0.2, #jittering the points horizontally  
 height = 0) +  
 labs(x = "Type of Shoe", #labels for x and y axis  
 y = "Pace (min)")



### C.

**Figure 1. Wearing docs tends to slow down walking pace.** Points represent observations of walking pace (min) for docs (n = 9) and for sneakers (n = 16). Box plots represent the central tendencies of the data and each point represents an observation of walking pace. The color represents the type of shoe (pink: docs, blue: sneakers).

### D.

#summarizing the data  
data\_summary <- my\_data |> #creating data\_summary to store values  
 group\_by(Type\_of\_Shoe) |> #grouping by Type\_of\_shoe  
 summarise(  
 median = round(median(Pace\_min), digits = 1), #calculating the median Pace\_min  
 minimum = round(min(Pace\_min), digits = 1), #calculating mimumum Pace\_min  
 maximum = round(max(Pace\_min), digits = 1), #calculating maximum Pace\_min  
 IQR = round(IQR(Pace\_min), digits = 1) #calculating interquartile range for Pace\_min  
 )  
ft <- flextable(data\_summary) #creating table from my\_data and storing it as an object called ft  
ft <- set\_caption(ft, caption = "Walking Pace in Docs vs Sneakers") #adding caption  
ft <- set\_header\_labels(ft, #creating more descriptive labels  
 median = "Median (min)",  
 minimum = "Minimum (min)",  
 maximum = "Maximum (min)",  
 IQR = "IQR (min)",  
 Type\_of\_Shoe = "Type of Shoe")  
ft <- theme\_box(ft) #changing theme of table  
ft

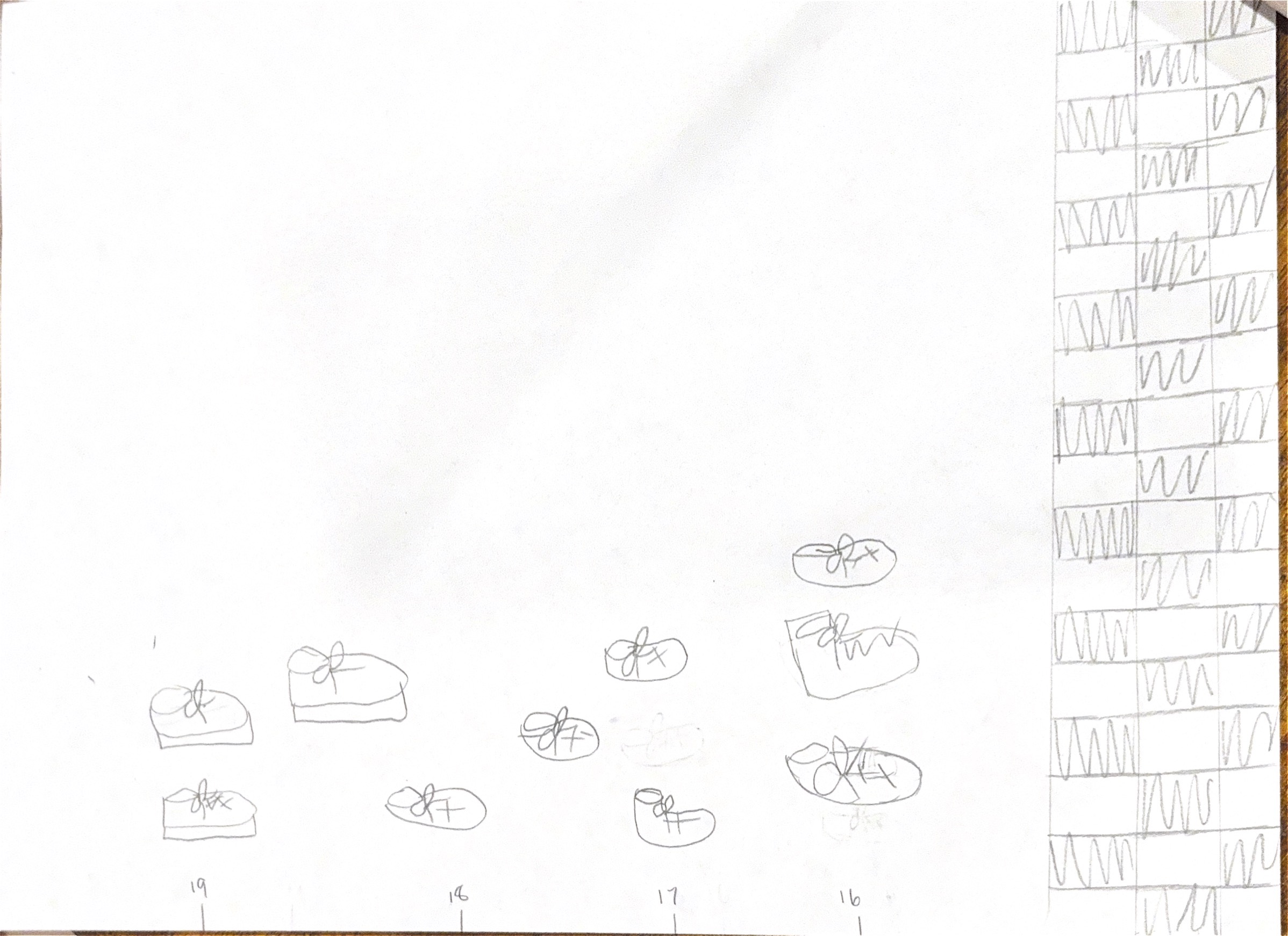
| **Type of Shoe** | **Median (min)** | **Minimum (min)** | **Maximum (min)** | **IQR (min)** |
| --- | --- | --- | --- | --- |
| Doc Martens | 18.6 | 17.8 | 20.3 | 0.8 |
| Sneakers | 17.9 | 16.9 | 19.3 | 1.3 |

#### Problem 2.

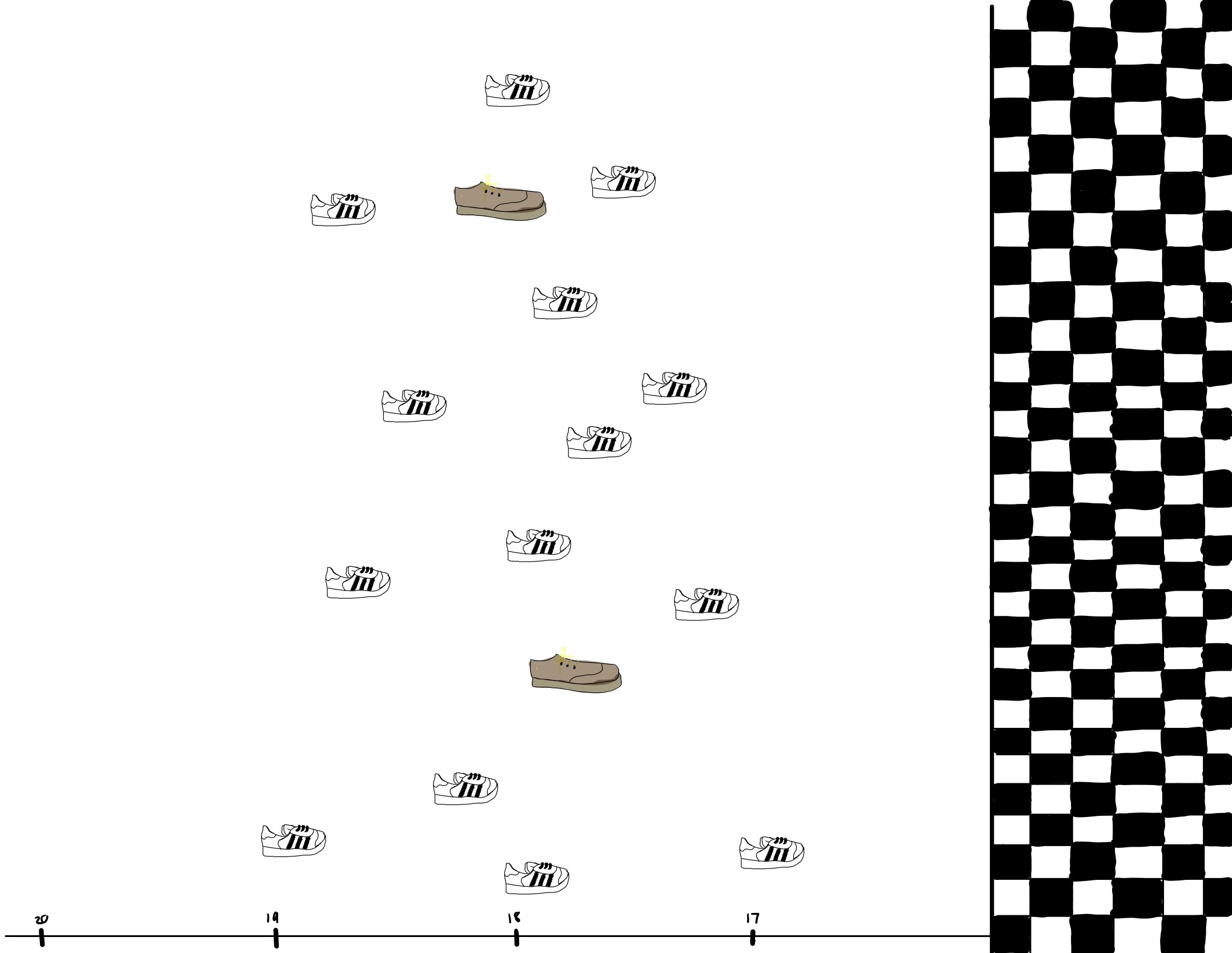
### A.

An affective visualization for my data could be a drawing of each shoe that I was wearing for each observation. These shoe drawings would be distributed on the page to show the walking pace for that observation. This could be done by drawing a finish line and drawing shoes with a faster pace closer to the line and shoes that resulted in a slower pace further from the line.

### B.



### C.



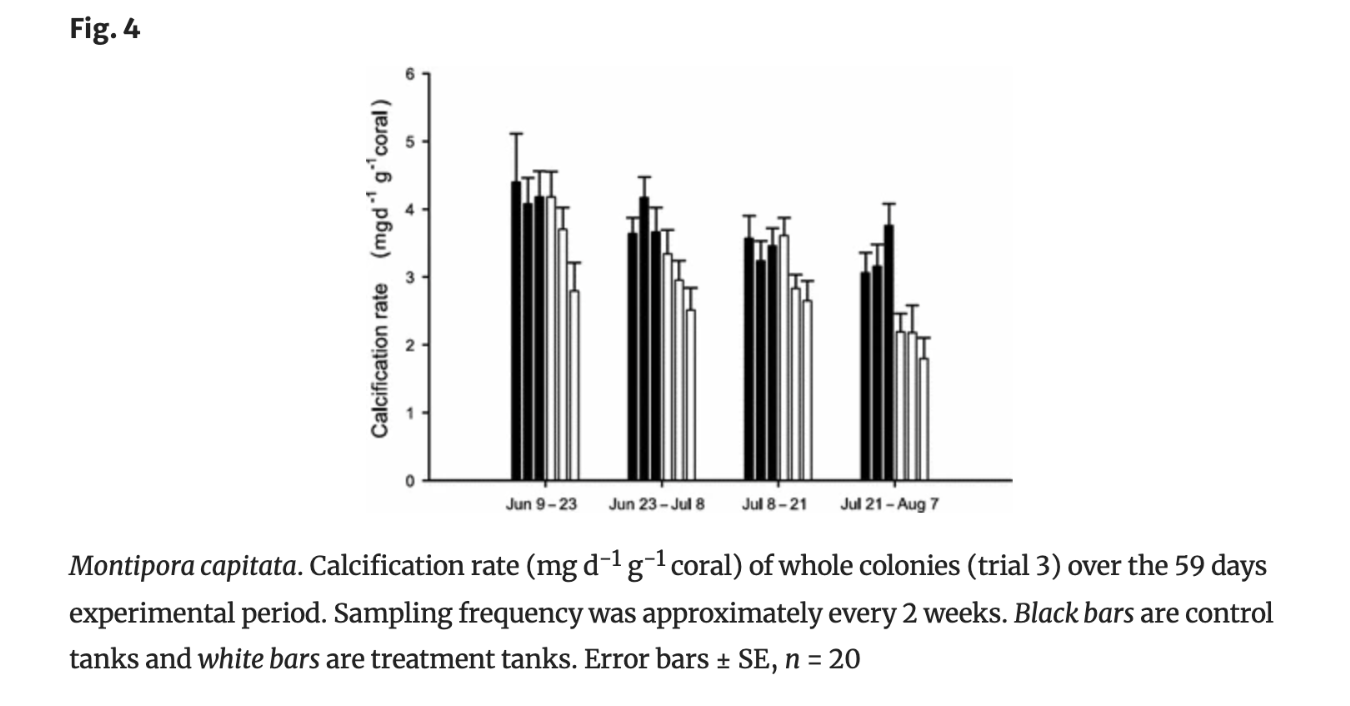
### D.

My piece shows my different shoes and my walking pace while I was wearing them. I was inspired by Stefanie Posavec and Giorgia Lupi’s Dear Data project. My work is a digital drawing on my ipad. I did this by drawing each shoe and then making an axis for walking pace (min) at the bottom and jittering each point (shoe) vertically.

#### Problem 3.

### A.

This research paper aimed to answer how ocean acidification impacts coral growth, which was done by performing a one way ANCOVA test that showed the reduction in calcification rates of coral in the acidification treatment was significant in comparison to the control group.



### B.

The authors were able to visually represent their statistics very clearly by showing the mean calcification rate for each colony, as represented by each bar on the graph. Additionally, the error bars provide a visual representation of standard error for the data. Furthermore, the x and y axis of the bar chart are clear and understandable with calcification rate on the x-axis and the dates for each observation period on the y-axis.

### C.

The bar chart has a good data:ink ratio because there is no excess visual clutter. For example, the chart does not have any grid lines and has very simple numbers listed on the y-axis. The authors did a good job making an aesthetically pleasing figure because only the essential information is displayed, making the figure easy to understand and interpret quickly.

### D.

One recommendation that I would make for this figure is to move the figure number from the top of the figure to the caption so that everything is grouped together in one chunk. Visually, I think that this would make more sense and be more clear. In addition, the researchers could also add a short descriptive title at the top of the figure so that the viewer immediately knows what data is being displayed. Furthermore, I would recommend changing the colors to something slightly more intuitive. Personally, I would assume that the white bars are the control because they seem “blank,” however it is actually the opposite in this figure. I think it would make most sense to change the treatment color to a light red to represent an acidic environment and the control color to be more blue to represent the natural ocean pH since these colors are widely accepted and intuitive for a pH scale.