homework-3-submission

Yonatan Grossman

2025-05-28

Here’s the link

https://github.com/Yonatan-Grossman/ENVS-193DS\_homework-03.git

library(tidyverse) #read in packages

── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
✔ dplyr 1.1.4 ✔ readr 2.1.5  
✔ forcats 1.0.0 ✔ stringr 1.5.1  
✔ ggplot2 3.5.2 ✔ tibble 3.2.1  
✔ lubridate 1.9.4 ✔ tidyr 1.3.1  
✔ purrr 1.0.4   
── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
✖ dplyr::filter() masks stats::filter()  
✖ dplyr::lag() masks stats::lag()  
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(here)

here() starts at C:/Users/16177/193 DS/ES 193DS/git/ENVS-193DS\_homework\_3/ENVS-193DS\_homework-03

library(flextable)

Attaching package: 'flextable'  
  
The following object is masked from 'package:purrr':  
  
 compose

library(janitor)

Attaching package: 'janitor'  
  
The following objects are masked from 'package:stats':  
  
 chisq.test, fisher.test

library(dplyr)

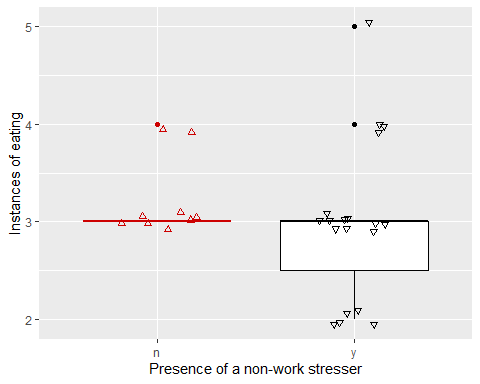
## Problem 1

# a.

I could make side-by-side box and whisker plots comparing the number of times I ate on days where there is and isn’t a non-work stresser. I find that when I’m stressed I tend to work more and spend less time eating so I think the mediam number of times I ate when stressed will be lower, and I think the range will also be lower since the more stressed I am the overall fewer opportunities I give myself to eat.

# b.

personal\_data <- read.csv("C:/Users/16177/193 DS/ES 193DS/git/ENVS-193DS\_homework\_3/ENVS-193DS\_homework-03/data/Personal data - Sheet1.csv") #reading in the personal data  
  
data\_clean <- personal\_data |> #makes a new object called data\_clean which is just the data from personal\_data  
 clean\_names() #cleans the data  
  
  
ggplot(data\_clean,  
 aes(x = presence\_of\_a\_non\_work\_stressor\_y\_n,  
 y = instances\_of\_eating,  
 color = presence\_of\_a\_non\_work\_stressor\_y\_n,  
 shape = presence\_of\_a\_non\_work\_stressor\_y\_n)) +  
 geom\_boxplot(show.legend = FALSE) + #boxplot  
 geom\_jitter(height = 0.1, width = 0.2, show.legend = FALSE) + #jitterplot  
 scale\_color\_manual(values = c("y" = "black", "n" = "red3")) + #custom colors  
 scale\_shape\_manual(values = c("y" = 6, "n" = 2)) + #custom shapes  
 labs (x = "Presence of a non-work stresser", y = "Instances of eating") #axis labels



# c.

**Figure 1: Number of times I eat on days I have a non-work stresser and days I don’t:**  
Comparison of number of times I eat on days I’m stressed for a reason unrelated to work (*n* = 19) and days where I’m not (*n* = 9) The boxplots represent median values, IQR, and spread of observations of number of times I ate food those days. Colors represent the location (black = stressor present, red = no stressor). The triangles represent individual observations.

# d.

personal\_summary <- data\_clean |>   
 group\_by(presence\_of\_a\_non\_work\_stressor\_y\_n) |>   
 summarise(  
 median = median(instances\_of\_eating), #calculating the following  
 n = length(instances\_of\_eating),  
 IQR = IQR(instances\_of\_eating),  
 min = min(instances\_of\_eating),  
 max = max(instances\_of\_eating),  
 .groups = "drop"  
 ) |>   
 rename("non-work stressor?" = presence\_of\_a\_non\_work\_stressor\_y\_n) |> #renaming the titles   
 flextable() |> #make a flextable   
 set\_table\_properties(layout = "autofit")  
personal\_summary #output the flextable

| non-work stressor? | median | n | IQR | min | max |
| --- | --- | --- | --- | --- | --- |
| n | 3 | 9 | 0.0 | 3 | 4 |
| y | 3 | 19 | 0.5 | 2 | 5 |