Student Sleep Patterns

Stat 184 Sec 2

Anna Kasehagen, Gracy Franco Prasanna, Melissa Kim, Sarah Khan

2024-12-13

1 Introduction

This is the citation for the dataset[1]

2 Data Exploration and Research Questions

What to include in this section: Describe the provenance of your data. That is, where did you get the data, who collected the data, for what purpose, who/what make up the cases. Explain how your data meet the FAIR and/or CARE Principles. Describe what attributes you will focus your analysis on (mention if they are part of your data sets or if you created them out of your data sets)

maybe here we can put a portion of the table of the cleaned data? idk just an idea

2.1 Research Question 1

2.2 Research Question 2

After analyzing how sleep duration and physical activity vary by university years, we wanted to see how given two of the attributes for a particular student, how the third would be affected. More specifically, the question we are trying to answer is:

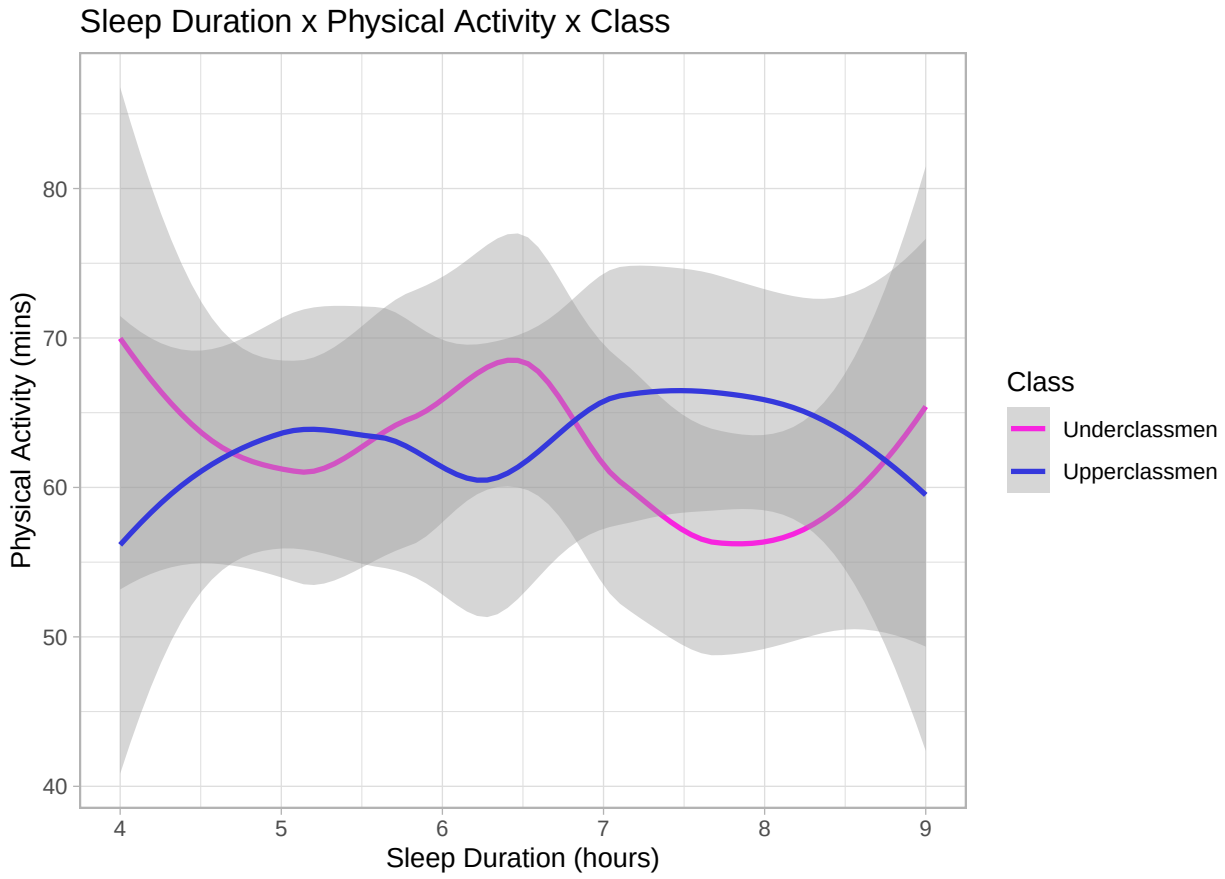
- Given the class group of a student (underclassmen or upperclassmen) and their sleep duration, how do their reported levels of physical activity vary?

To best visualize this relationship, we thought that a line graph would work best, and in that way we would be able to see the trend between the two classes with duration of sleep spanning from 4 to 9 hours. The graph below demonstrates this relationship. Some things to keep in mind are that the x axis, sleep duration, is measured in hours, while the y axis, physical activity levels, are measured in minutes, and underclassmen are represented by the pink line and upperclassmen are

[1] jamal, A. (2024, October 14). Student sleep patterns. Kaggle. https://www.kaggle.com/datasets/arsalanjamal002/student-sleep-patterns

represented by the blue line. The gray bubbles around the lines represent the confidence intervals, but for our analysis we just focused on the average lines themselves.

Figure 1: How Physical Activity Levels vary with Sleep Duration and Class Group



From this graph, there were a few things we noticed and found interesting:

- The first is how the trends for upperclassmen and underclassmen seem to be inverted. That is, how at a given hour for sleep duration, the physical activity of one group goes up while the other goes down, or vice versa.
- Second, if we look at where the lines cross the 8 hour mark (we chose 8 because it is right in the middle of the recommended amount of sleep a person should get every night (7-9)[2]) we can see a big gap between the recorded minutes of physical activity and underclassman might get (just above 55) versus an upperclassman (just above 65).
- The third interesting takeaway from looking at this graph is how we could visually see at what point of sleep duration do upperclassmen and underclassmen hit their maximum physical activity levels. We see that for underclassmen it is at 4 hours whereas for upperclassmen it is at about 7.5/8 hours.

[2]Assess your sleep needs. Sleep Medicine. (n.d.). https://sleep.hms.harvard.edu/education-training/public-education/sleep-and-health-education-program/sleep-health-education-92#:~:text=Although%20there%20is%20some%20genetic,sleep%20as%20long%20as%20possible.

After considering these takeaways, we made some inferences about what might have influenced these sleep patterns and activity levels between classes. The jump between physical activity levels at certain hours and the difference between where the groups reached their maximum levels of sleep duration could be attributed to a difference in time management skills. As students reach their 3rd and 4th years, they typically have figured out a schedule that allows them to balance getting good sleep while maintaining an active lifestyle. This trend could also be attributed to how upperclassmen might have learned the value of maintaining their physical health, while underclassmen might be spreading themselves thin since they have just entered a new environment.

2.3 Research Question 3

2.4 Research Question 4

3 Conclusion

4 References

5 Code Appendix

```
#loading necessary packages
library(tidyr)
library(dplyr)
library(rvest)
library(googlesheets4)
library(ggplot2)
library(esquisse)
# loading data
sleepData <- read_sheet("https://docs.google.com/spreadsheets/d/1BszLI2k3ti0AzKrY6msUY5lqblDG-

#cleaning the data and keeping only what is applicable
CleanedSleepData <- sleepData %>%
  select(4, 5, 9) %>%
  na.omit()

Q2SleepData <- CleanedSleepData %>%
  mutate(
    University_Year = case_match(
      .x = University_Year,
      "1st Year" ~ "Underclassmen",
      "2nd Year" ~ "Underclassmen",
      "3rd Year" ~ "Upperclassmen",
      "4th Year" ~ "Upperclassmen",
      .default = "missing"
    )
```

```
  )
# Make Data Visualization ----
ggplot(
  data = Q2SleepData,
  mapping = aes(
    x = Sleep_Duration,
    y = Physical_Activity,
    colour = University_Year
   )
  ) +
  geom_smooth(se = TRUE) +
  scale_color_manual(
    values = c(Underclassmen = "#F725DF",
               Upperclassmen = "#3538DC")
  ) +
  labs(
    x = "Sleep Duration (hours)",
    y = "Physical Activity (mins)",
    title = "Sleep Duration x Physical Activity x Class",
    color = "Class"
  ) +
  theme_light()
```