

# Your project title

Wale: Liane, Amy, Eshan, Will

2024-10-31

Your written report goes here!

## ! Important

Before you submit, make sure your code chunks are turned off with `echo: false` and there are no warnings or messages with `warning: false` and `message: false` in the YAML.

## Exploratory Data Analysis

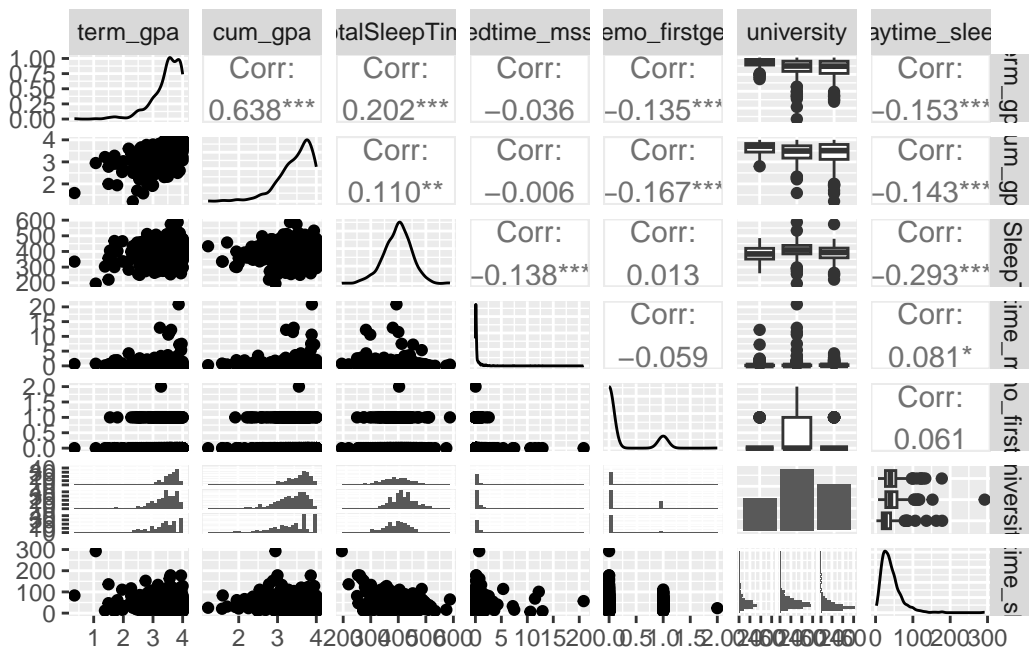
Description of the data set and key variables.

The data was originally collected in 2019, with the participants being first-year students at the following three universities: Carnegie Mellon University (CMU), a STEM-focused private university, The University of Washington (UW), a large public university, and Notre Dame University (ND), a private Catholic university. To collect data on sleep, each participating student was given a Fitbit device to track their sleep and physical activity for a month in the spring term, and grade and demographic data was provided by university registrars.

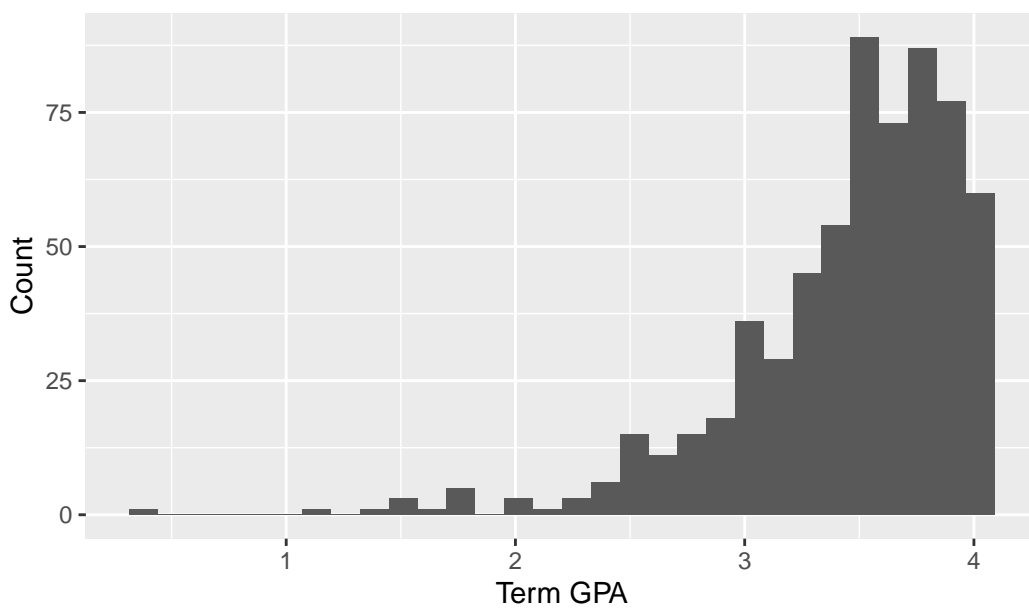
There are 634 observations, representing the 634 participants in this study. Race is a binary variable separated into underrepresented students and non-underrepresented students with 0 being underrepresented and 1 being non-underrepresented. Students are considered underrepresented if either parent is Black, Hispanic or Latino, Native American, or Pacific, and students are deemed non-underrepresented if both parents have White or Asian ancestry. The gender of the subject is also binary with 0 being male and 1 being female. First-generation status is binary with 0 being non-first gen and 1 being first-gen. The mean successive squared difference of bedtime measures the bedtime variability, specifically the average of the squared difference of bedtime on consecutive nights.

## Clean Data

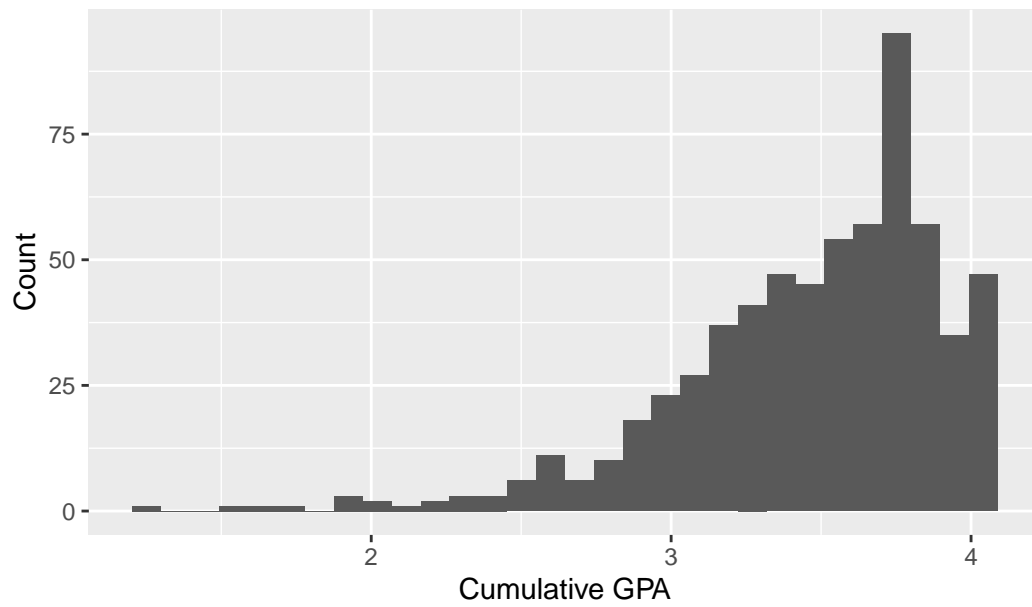
Univariate EDA of The Response & Key Predictor Variables



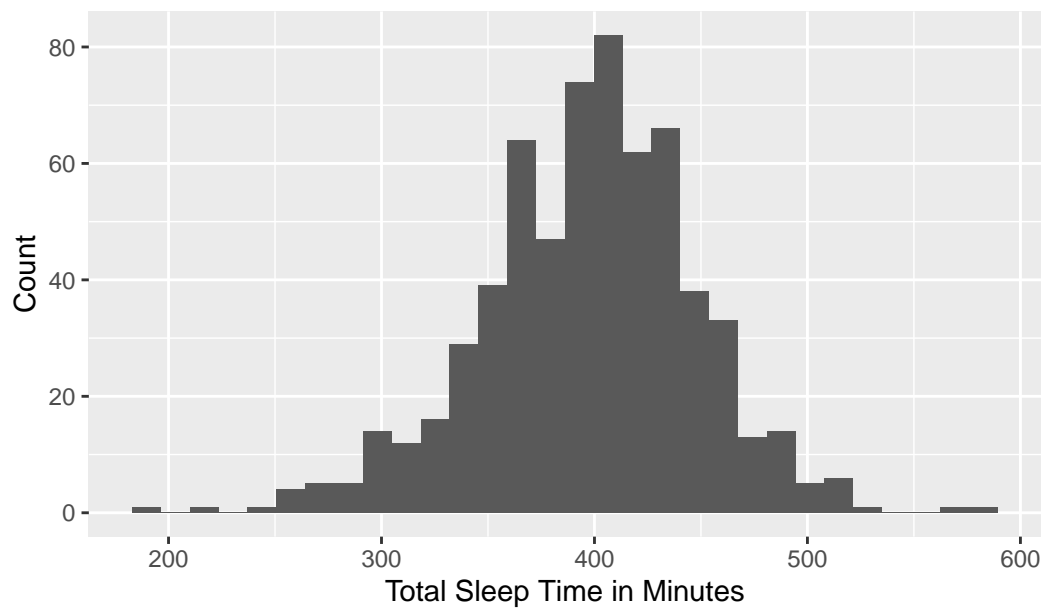
Distribution of the Term GPA



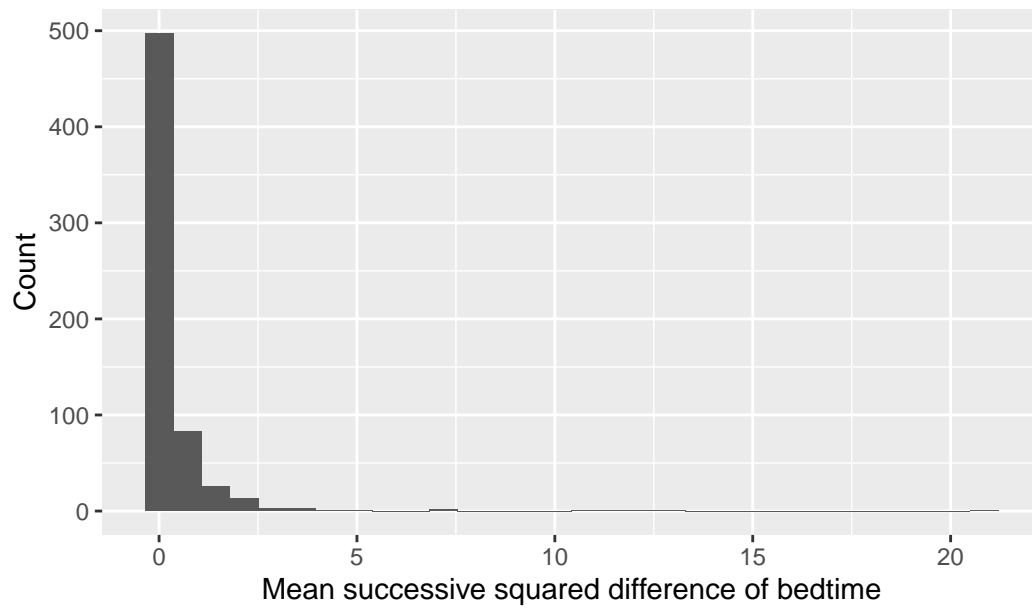
Distribution of the Cumulative GPA



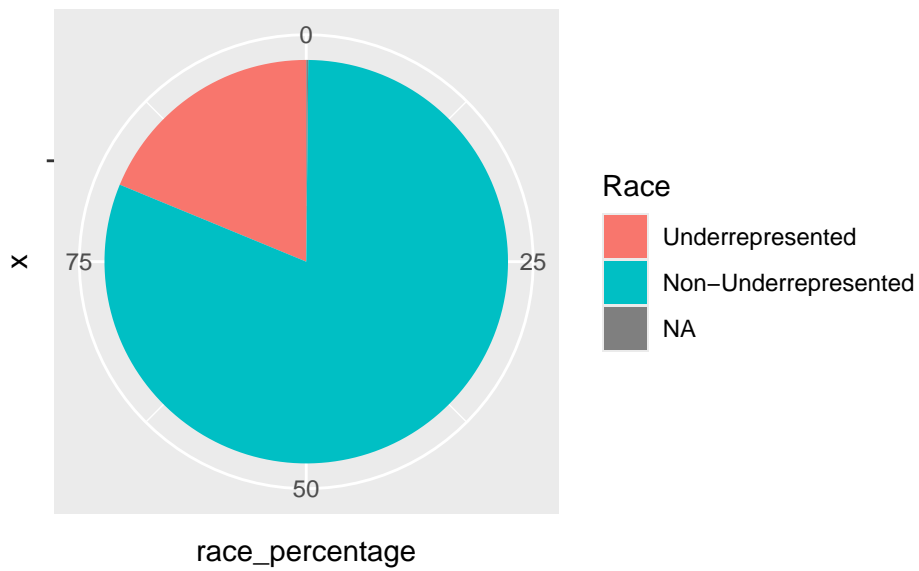
Distribution of the Total Sleep Time



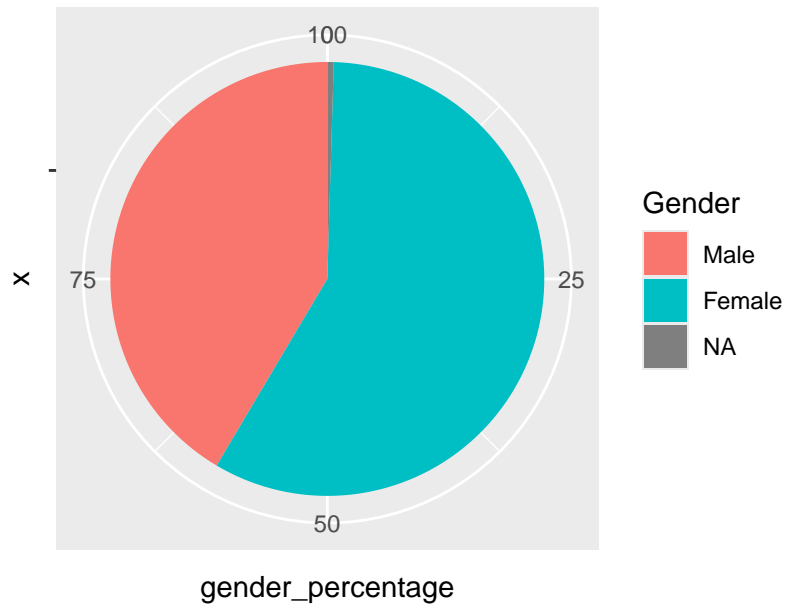
Distribution of the Bedtime Variability



Distribution of Underrepresented Vs. Non-Underrepresented Students



Distribution of Gender



Bivariate EDA of The Response & Key Predictor Variables

