

This dataset is created by Professor Gregory Samanez-Larkin at Duke University, which was used for a final exam on a statistics class (source: <https://osf.io/5egfq/>)

### **Overview: Physical activity, sleep, and academic performance**

The dataset was collected by Fitbits on activity (number of steps taken by a student) and sleep as well as some data related to academic performance in a class and in general. You can use the dataset to examine how overall amount of steps and sleep are related to academic outcomes, or evaluate the effects of sleep variability (the opposite of sleep consistency) on academic outcomes. Studies have shown that sleep variability is more strongly related to grades than overall sleep. Of course, you can also ask any other interesting questions from this dataset.

### **Description of column names**

The data file stepsleepacademics.csv includes the following variables

student: A unique ID code for each student participant in the class (note: this has been changed to something random so that no individual student can be identified).

stepsave: Average number of daily steps taken by the student across the semester.

stepslmh: This variable groups students into low (1), medium (2), and high (3) levels of average daily steps across the semester.

stepssd: Standard deviation across days of daily steps taken by the student across the semester as a measure of step variability.

stepssdlmh: This variable groups students into low (1), medium (2), and high (3) levels of step variability across the semester.

sleepave: Average number of daily minutes asleep for the student across the semester.

sleeplmh: This variable groups students into low (1), medium (2), and high (3) levels of average daily sleep across the semester.

sleepsd: Standard deviation across days of daily sleep for the student across the semester as a measure of sleep variability.

sleepsdlmh: This variable groups students into low (1), medium (2), and high (3) levels of sleep variability across the semester.

female: 1 if female, 0 if male.

psetz: Average grades across all assignments (psets) in PSY201. Grades were standardized by z-scoring to mask student identity. Based on the 9 assignments completed individually; the assignment completed in groups during class is excluded from this score.

understandse: At the beginning of lecture one day students self-reported how well they understood the concept of standard error using a 1-7 scale.

corrguesserror: During lecture one day students played [guessthecorrelation.com](http://guessthecorrelation.com) and reported their average errors (deviation between actual correlation depicted and student guess).

hwcondiff: During lecture students self-reported how difficult it is to concentrate while doing homework assignments in general (not specific to this class) using a 1-7 scale.

lecondiff: During lecture students self-reported how difficult it is to concentrate during lectures in general (not specific to this class) using a 1-7 scale.

lecturetired: At the beginning of lecture one day students self-reported their current level of tiredness using a 1-7 scale.

preexamstress: During lecture students self-reported their average level of stress before exams (not specific to this class) using a 1-7 scale.