

Stakeholders:

Tara Rajagopalan (trajagop@calpoly.edu)

Daniel Alvarez (dalvar70@calpoly.edu)

Dataset Link: <https://www.kaggle.com/datasets/mahmoudelhemaly/students-grading-dataset>

Our Organization:

We are a private academy trying to build a successful educational curriculum for middle and high school students for grades 6 through 12. We would like to see how several factors affect students' academic performances in order to best optimize our academic program.

Dataset Name: Student Performance and Behavior Data Set on Kaggle

Dataset Description:

This data set is about college students' academic performance. The data set provides information about each student in three main categories, including demographic background, family income, and grades in coursework and exams. The information presented in these main categories are used to assess student performance and ultimately to understand students academic needs.

List of questions needed for data analysis / machine learning:

- What is the average number of hours students study per week?
- Given a student, what is the average number of hours of sleep they receive per night?
- What proportion of students come from a home with no internet access?
- How does a student's participation score differ between income levels (low, medium, high)?
- Does parent education level impact students' average score on quizzes and assignments?
- Does the student's stress level have a huge impact on how well they perform on final exams?
- Does participation in extracurricular activities impact the amount of hours that a student studies per week?

Purpose for Asking Questions:

A student's daily lifestyle factors such as amount of sleep, amount of study hours, and participation in extracurricular activities can make a positive or negative impact on academic performance.

We want to find a correlation between a student's home life and their academic performance. Internet access at home allows students to conveniently access academic resources for studying. Oftentimes, students who come from high income, well educated families have easier access to tutoring resources. We would like to see if the data presents this information to us.