Dataset Description

Dataset Title: Student Stress Monitoring Datasets
Dataset Source: Kaggle ([Student Stress Monitoring Datasets](#))
William Armentrout, David Rayle, Katelyn Ledbetter, Ethan Meadows

Dataset Overview

The dataset provides a look into the causes of stress among college students aged between 18 and 21. The data was derived from a nationwide survey and contains hundreds of anonymous responses. Responses were gathered with a survey using a five point Likert scale ranging from "Not at all" to "Extremely".
The dataset aims to identify the sources and types of stress affecting college students and to explore correlations between emotional states, physical health, academic pressures, and social factors. The dataset also includes a target variable on what type of stress each respondent primarily experiences as eustress, distress, or no stress which makes the dataset suitable for both descriptive and predictive analyses.

Subject Matter and Variables

The dataset spans multiple aspects of student life, grouped into categories:
- Demographics: Age and Gender
- Emotional and Stress Indicators: anxiety, sleep problems, sadness, irritability, and loneliness.
- Physical/Health Indicators: Frequency of headaches, presence of health issues, and weight changes.
- Academic and Enviornment Stressors: Academic workload, peer competition, extracurriculars, attendance, etc.
- Social factors: personal relationships and free time.

Data Organization

The dataset is organized in a tabular format, with a row representing an individual's response and each column representing a variable. This format allows for easier analysis as it works well with various tools in Python.

Limitations

The dataset does rely entirely on self reported responses, which may be influenced by personal perception or willingness to disclose certain information about themselves. The age range restricts generalization beyond traditional college students. The dataset also only uses binary encoding for gender, which excludes non-binary responses. It also does not include socioeconomic factors or cultural differences, which could play a significant role in stress levels.

Why this Dataset

We went with this dataset because of our shared experiences as college students and all the stress that being in college can place on someone. The fact that many of the respondents are in a similar situation to us is something that will help keep our interest in this project as the semester advances.

Citation

Student Stress Monitoring Datasets. (2025). Kaggle. Retrieved from [Student Stress Monitoring Datasets](#)