

Project Proposal Write-up

Team Introduction

- **Team Name:**

Group 8

- **Team Members:**

Lisa Sverdlova– Product Manager (Sprint 1)

Alvin Lee – Data Engineer (Sprint 1)

Ola Szymula– Data Engineer (Sprint 1)

Liru Chen– Data Storyteller (Sprint 1)

Dataset Choice & Justification

- **Dataset Title:**

Student Performance and Learning Behavior

- **Source Link:**

<https://www.kaggle.com/datasets/adilshamim8/student-performance-and-learning-style>

- **Size/Scope:** [# rows, # columns, file size, data types]

14,003 rows

16 columns

539.61 kB

Boolean, int, real

- **Description:**

The dataset provides a comprehensive view of student performance and learning behavior, integrating academic, demographic, behavioral, and psychological factors.

- **Justification:**

The dataset provides a multitude of practical data that can be used to infer the factors that have the most impact on student performance.

Vision Statement

- **Problem/Question:**

Which factors have the most impact on the academic performance of students?

- **High-Level Goal:**

The goal of the project is to analyze real world data related to the academic performance of students and infer learning patterns and weight coefficients of various factors. The datasets that will be analyzed will include various independent variables, such as learning styles, teaching models, social aspects, personal lifestyle, types of motivation, etc. The final product will include the description of all of the techniques that were used for the analysis, the theoretical background behind them, and practical conclusions and recommendations that can be used in practice for improvement of student academic performance.

- **Potential Users:**

Students themselves, their parents, researchers, the university administrative staff, and counselors can find the results useful.

Repository Setup

- **GitHub Repo Link:**

<https://github.com/sjsu-cs131-f25/Group-8-Education-Data>

- **Repo Name:**

Group8-Education-Data

- **Repo Structure (initial):**

README.md – project overview
data/ – download instructions
notebooks/ – exploratory analysis notebooks (to be added)
src/ – source code (to be added)
.gitignore – standard ignores

Project Management

- **Tool Used:**

Microsoft Excel

- **Board Link:**

<https://github.com/sjsu-cs131-f25/Group-8-Education-Data/blob/main/Project%20Grant%20Chart.xlsx>

- **Sprint #1 Scope:**

<i>1.1.1 Adapt PM Template:</i>	<i>Lisa S.</i>
<i>1.1.2 Scope Week 1 Sprint:</i>	<i>Lisa S.</i>
<i>1.1.3 Define High Level Goals:</i>	<i>Lisa S.</i>
<i>1.1.4 Submit Sprint Tracking Board Link:</i>	<i>Lisa S.</i>
<i>1.2.1 Find Accessible DataSet:</i>	<i>Alvin L. and Ola S.</i>
<i>1.2.2 Provide DataSet Link:</i>	<i>Alvin L. and Ola S.</i>
<i>1.2.3 Set Up GitHub Repository:</i>	<i>Alvin L. and Ola S.</i>
<i>1.3.1 Project Proposal Writeup:</i>	<i>Liru C.</i>