

Final Project - Step 1 (8 Points)

PSTAT100: Data Science Concepts and Analysis

Due Date

The deadline for this step is Friday, April 25.

Formation of groups:

GROUP NAME

- Matthew Arteaga (matthewarteaga)
- Phuc Lu (pdlu)
- Valerie De La Fuente (valeriedelafuente)
- Hayden Galletta (haydengalletta)
- William Nelson (williamnelson)

1 Tentative Role Distribution Table

Name	Responsibilities
Matthew Arteaga	Develop Research Questions
Phuc Lu	Form Hypothesis
Valerie De La Fuente	Data Preparation (cleaning)
Hayden Galletta	Explatory Data Analysis
William Nelson	Coder

2 Data Information

Data Name/Title: *Student Depression Dataset*

Author/Owner: *Israel Campero Jurado*

Date of Publication: *March 12, 2025*

Publication Venue: *OpenML*

Retrieval Date: *April 22, 2025*

Link: *<https://www.openml.org/search?type=data&status=active&id=46753&sort=runs>*

3 Initial Insights:

This data set is particularly of interest for our group because it will allow us to analyze how a multitude of aspects of a student's life play a role in whether or not they consider themselves depressed. In addition, we can analyze whether or not the increase of certain variables results in a higher rate of depression among groups.