Dataset Info – Exams

Dataset Source: <https://www.kaggle.com/datasets/rkiattisak/student-performance-in-mathematics>

This dataset contains information on the performance of high school students in exam, the data was collected from three high schools in the United States. This dataset contains approximately 1,000 records with attributes like:

a. Demographics: Gender, race/ethnicity

b. Socioeconomic: Parental education, lunch type

c. Academic: Preparation course, test scores (Math, Reading, Writing)

**Features (Columns):**

| **Column Name** | **Description** |
| --- | --- |
| gender | Student's gender (male, female) |
| race/ethnicity | Ethnic group classification (group A to group E) |
| parental level of education | Highest education level attained by a parent |
| lunch | Type of lunch received (standard, free/reduced) |
| test preparation course | Completion status of a test prep course (completed, none) |
| math score | Score in mathematics (0–100) |
| reading score | Score in reading comprehension (0–100) |
| writing score | Score in written expression (0–100) |

📂 **File Provided:**

* exams.csv — raw dataset containing 1,000+ student records with demographic and academic performance data.

🧹 **Cleaning Summary:**

* Verified no missing values
* Ensured consistent formatting for categorical variables
* Confirmed numeric scores are within valid range (0–100)
* Ready for exploratory analysis and predictive modeling

📊 Used In:

Capstone Project for the SCTP Associate Data Analyst Program – Python Track

Author: Janet Haw

Project Title: Student Performance in Exams