

# Project Proposal: Student Performance Analysis

## Project Idea:

This project explores how socioeconomic and learning support factors influence student academic performance. Using the [Student Performance Factors Dataset](#), the analysis will examine how elements such as family income, access to educational resources, study habits, and learning environment affect exam outcomes. The goal is to identify key predictors of academic success and highlight areas where additional student support could improve results.

## Key Research Questions:

1. How do socioeconomic factors (family income, internet access, access to resources) relate to exam scores?
2. Does increased effort (hours studied, attendance, tutoring sessions) consistently lead to higher performance?
3. How do learning challenges, such as distance from home or limited parental involvement, impact achievement?

## Target Audience:

Educational institutions, including schools, colleges, and education support organisations, aiming to improve learning outcomes and close achievement gaps for students from disadvantaged backgrounds.

## Primary Project Objectives

- ETL Pipeline: Design and implement a complete, robust ETL process in Python (Pandas/NumPy).
- Key Insight: Identify the strongest positive and negative correlations with Exam\_Score.

Theme	Relevant Columns	Purpose
<b>Socioeconomic Factors</b>	Family_Income, Internet_Access, Access_to_Resources, Distance_from_Home, Parental_Education_Level	Understand how privilege and access influence performance.
<b>Learning Support</b>	Tutoring_Sessions, Parental_Involvement, Teacher_Quality, School_Type	Measure impact of external and institutional support.
<b>Student Effort / Habits</b>	Hours_Studied, Attendance, Motivation_Level, Sleep_Hours, Physical_Activity	See if higher effort and discipline correlate with better outcomes.
<b>Performance Metrics</b>	Previous_Scores, Exam_Score	Benchmark and target variable for analysis.