Analysis of Student Data to Determine Significance of SAT/ACT as Predictor of College GPA

Lisa Siefker
March 28, 2025

Introduction

Colleges across the U.S. are considering updating pandemic era test-optional admission policies. However, experts continue to debate the usefulness and equity of considering standardized test scores in college admission decisions.

Experts who support mandatory test requirements argue that test scores provide an objective measure of academic readiness and are a useful predictor of college success. Opponents of standardized tests have traditionally argued that test requirements disadvantage low-income and minority students[1]. Specifically, low-income and minority students may not have access to expensive, high-quality test preparation resources. They may also have less time to devote to test preparation since they may have jobs or care for family members. In these situations, high school GPA may be able to more accurately predict college GPA.

In this study, I examine student data to help answer the question of whether high school GPA is just as effective at predicting college GPA as standardized test scores. Specifically, I separately analyze whether high school GPA and SAT/ACT scores are significant predictors of college GPA across white, Asian, and minority populations.

The research goal is explanatory because it aims to test a specific hypothesis about the relationship between SAT/ACT scores and college GPA.

Methods

Data

This study used visualization and statistical analysis to examine data from 1,528 students at a selective four-year U.S. college. For each student, the following relevant information was collected:

- Parental education levels for two parents
- Student race (white, Asian, or minority)
- High school weighted GPA, ranging from 0 to 5
- SAT/ACT score (highest score, where ACT is weighted to equivalent SAT)
- College program (Arts & Letters, Business & Economics, Math & Science)

[1] Leonhardt, D. (2024, January 7). The Misguided War on the SAT. *New York Times*.

- College GPA, ranging from 0 to 4

This data has a natural time order because SAT/ACT scores represent student data measured before college admission, and college GPA measures student data after college admission. It is not clear from the dataset when each piece of data was collected, but I will assume the data is retrospective. In other words, the data represents past SAT/ACT scores and college GPAs.

Statistical Analysis

The statistical analysis will use scatterplots and linear regression to test the null hypothesis that SAT/ACT score has no linear relationship to college GPA against the alternative hypothesis that there is a linear relationship between SAT/ACT score (independent variable) and college GPA (dependent variable). Similarly, I will use scatterplots and linear regression to test the null hypothesis that high school GPA has no linear relationship to college GPA against the alternative hypothesis that there is a linear relationship between high school GPA (independent variable) and college GPA (dependent variable).

Next, I will fit a multiple regression model to simultaneously evaluate SAT/ACT score and high school GPA as predictors of college GPA.

Finally, I will examine the output of all three models to compare performance and check model significance using ANOVA to determine whether adding SAT/ACT as a predictor significantly improves the model that predicts based on high school GPA alone. I will create a scatter plot of both relationships to visually examine the trends.

Other Variables

Other variables in the study that could potentially predict college GPA are parental education level, and type of college attended.

Parental education level is a potential confounding variable because it may indicate higher socioeconomic status and access to resources, which could distort the true relationship between SAT/ACT scores and college GPA. This variable should be considered in future analyses since the effect of socioeconomic advantages may have an impact on both SAT/ACT scores and college GPA.

Type of college attended could potentially be a moderating variable because academic admission requirements vary by college, and different college programs have different levels of difficulty. For those reasons, college type could potentially impact test scores required for admission and college GPA. This variable should be considered in future analyses to determine whether SAT/ACT scores are equally predictive across different colleges.

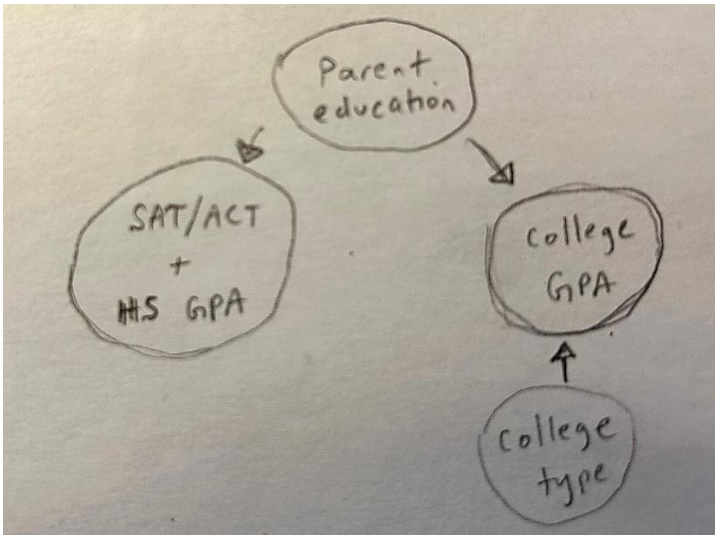

Fig. 1 illustrates the relationship of SAT/ACT and high school GPA covariants, the potential confounding variable of parent education, and the potentially moderating variable of college type.

Results

Simple Linear Regression Models

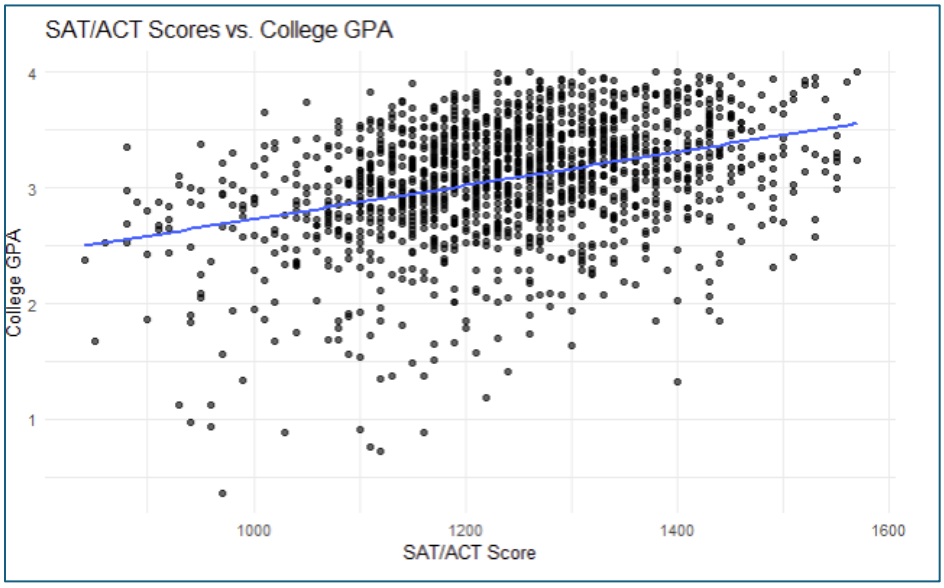

Fig. 2 is a scatterplot of SAT/ACT score v. college GPA. The regression line shows a strong positive association between SAT/ACT score and college GPA. As SAT/ACT score increases, college GPA also increases.

The simple linear regression model of the SAT/ACT score v. college GPA relationship shown in Fig. 2 has a slope of about 1.456, indicating a strong positive linear relationship. The effect is statistically significant at the 5% significance level, so I will reject the null hypothesis that SAT/ACT has no linear relationship to college GPA.

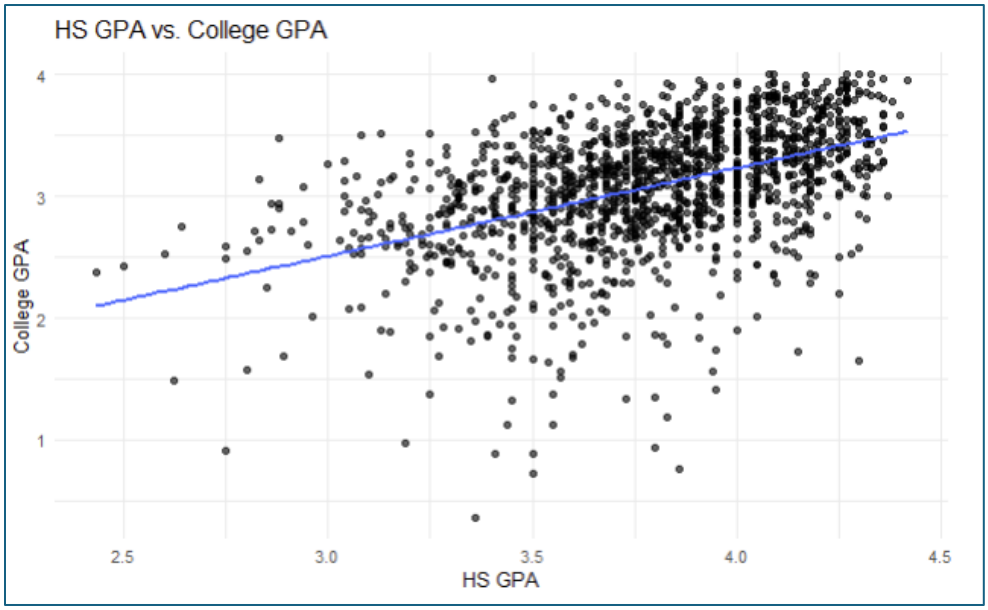

Fig. 3 is a scatterplot of high school GPA v. college GPA. The regression line shows a strong positive association between high school GPA and college GPA. As high school GPA increases, college GPA also increases.

The simple linear regression model of high school GPA v. college GPA relationship shown in Fig. 3 has a slope of about 0.724, indicating a positive linear relationship. The effect is statistically significant at the 5% significance level, so I will reject the null hypothesis that high school GPA has no linear relationship to college GPA.

Comparing the two models shows a stronger positive relationship between ACT/SAT and college GPA (1.456) than high school GPA and college GPA (0.724). The adjusted R squared values indicate that SAT/ACT accounts for only 13% of variance in college GPA while high school GPA accounts for 21.4% of variance in college GPA. This indicates a better fit for the high school GPA model, but both values are low, which indicates that other factors contribute to variation in college GPA.

Multiple Regression Model

The multiple regression model that simultaneously includes SAT/ACT and high school GPA as predictors shows that both predictors are statistically significant. The adjusted R squared value indicates that the multiple regression model explains only about 26% of variance in college GPA. The adjusted R squared value is very close to the R squared value, suggesting that both SAT/ACT and high school GPA contribute meaningfully to the model.

Comparing the multiple regression model to the SAT/ACT only model (Fig.2) indicates that SAT/ACT alone is not a strong predictor of college GPA. Adding high school GPA improves the model significantly, explaining more variance in college GPA. ANOVA also confirmed that adding high school GPA improves the model beyond using high school GPA alone.

Discussion

This study confirmed that both SAT/ACT scores and high school GPA are useful predictors of college GPA. However, standardized test scores and high school GPA only explain about 26% of the

variance in college GPA, suggesting that additional factors impact college GPA. Future research and analysis should focus on other potential confounding and moderating variables, such as parental education, type of college degree, drug and alcohol consumption, and study habits.