An Analysis of Student Performance

By Bogdan Tesileanu

Abstract

Students are faced with a barrage of test throughout their schooling. The outcome of those tests are based on the student's performance throughout the year. However, different students perform differently based on their individual circumstances at home. The ones studied in this analysis tend to be mostly socioeconomic. When looking at a large group of students, it was found that those with better parental education, those without welfare aid from the government, and those who have taken an out-of-pocket preparation course performed much better on exams in math, reading, and writing. It is important to note that a student's ability to learn is not only based on their achievement in school, but also what they are able to practice and learn at home.

Introduction

Learning institutions are everywhere, and can be as diverse as their students in terms of practices and support. For this reason, understanding the situations and motivations of each student allows the school and teachers of that school to best prepare students for the future ahead. As a result, schools must keep in mind the differences between each student and their own economic situation at home.

Background

Education is of utmost importance in most countries, and thus has been widely studied. While some of the research happened across the seas, it is all relevant in terms of what boosts test

performance and what doesn't. Author Jennifer Barry highlights the importance of a student's socioeconomic status, finding that it was the biggest factor out of all that were studied. A lower status implies more stress and conflict at home, as well as less access to resources. The lunch status, whether free or reduced, is also indication of this status, and thus should be noted as a negatively correlated factor to test scores. Moreover, author Orhan Kara reported that school absences correlated negatively at about 22% with grades. Once students get behind, they are stuck in a vicious cycle of making up work while more work is accumulating. Another study by Michumbu highlighted the importance of parent involvement in the student's life. Doing things such as carpooling, clubs, chaperoning, and field trips allowed parents to be more involved in their child's life, and create a bond that would encourage success at school. Similarly, parent income was also positively correlated in this case. A higher income yields better resources, and more resources means the student is able to gather and absorb all the information they need more efficiently. More recently, Realyvásquez-Vargas decided to do the same investigations under the COVID-19 pandemic. He focused on factors that affected student learning at home. These included lighting, noise, and temperature. These factors were chosen since they differ from a regular school-setting classroom to one's bedroom. It was found that lighting and temperature had the biggest impact on concentration and focus, followed by noise. In a classroom, such variables are held constant, while at home (especially with poorer families), many of these conditions may be less than ideal. Last but not least, Kawtar Tani conducted a study that, like Kara's, focused on student absenteeism. According to the study, students in higher level classes had better attendance record than those in lower, more basic classes, and that increase meant better grades. The study also found that an increased number of dependents at home was correlated to an increase in grades. This effect could be attributed to an increased sense of responsibility at home, and a desire to succeed to help the family or significant other.

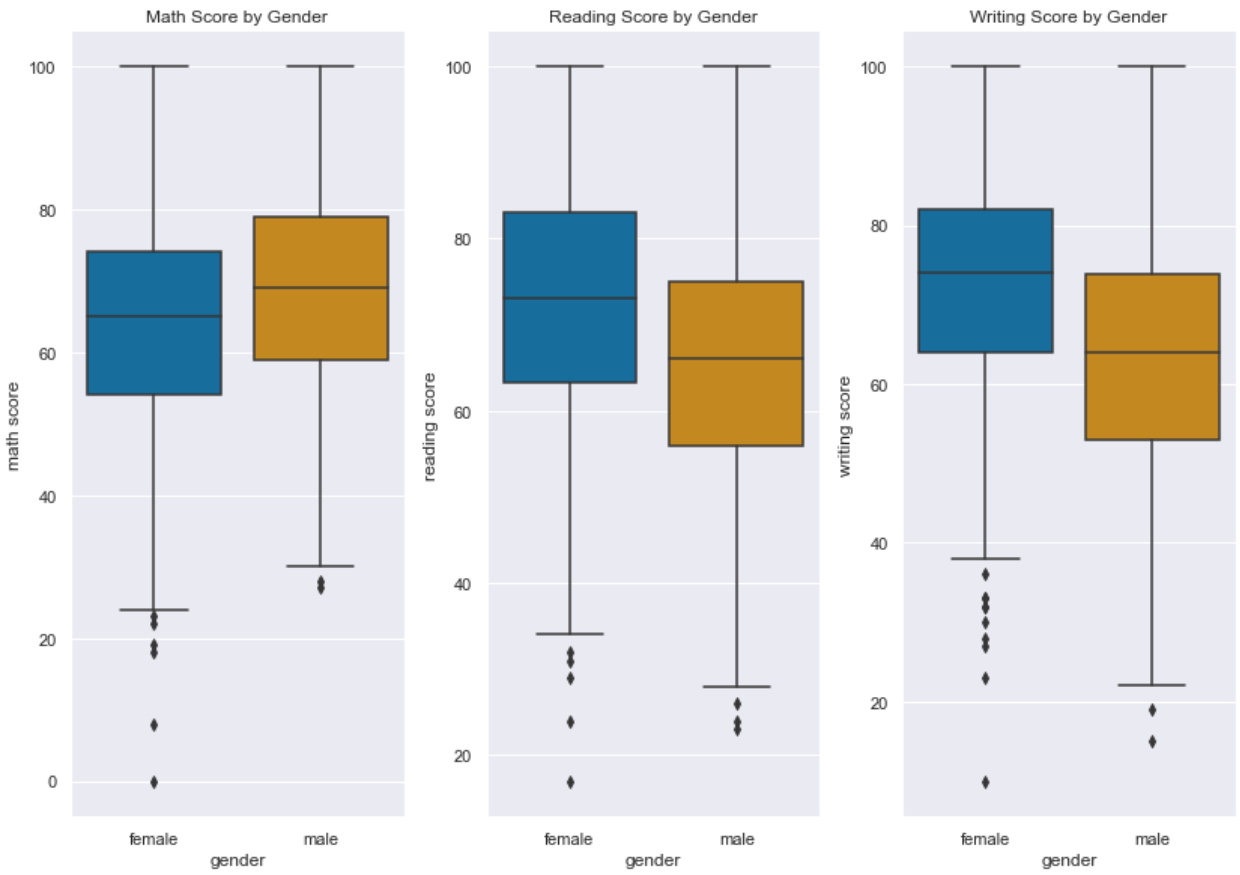
Approach

In this project, we selected the "Student Performance" dataset. The dataset has 8 variables, 5 of which are categorical, which we will use as features, and 3 of which are numerical, which will be our target variables. The dataset describes 5 student attributes: gender, race, parental level of education, type of lunch aid, and preparation course taken. The 3 target variables are the students' math, reading, and writing scores on their exams. A data summary table is below:

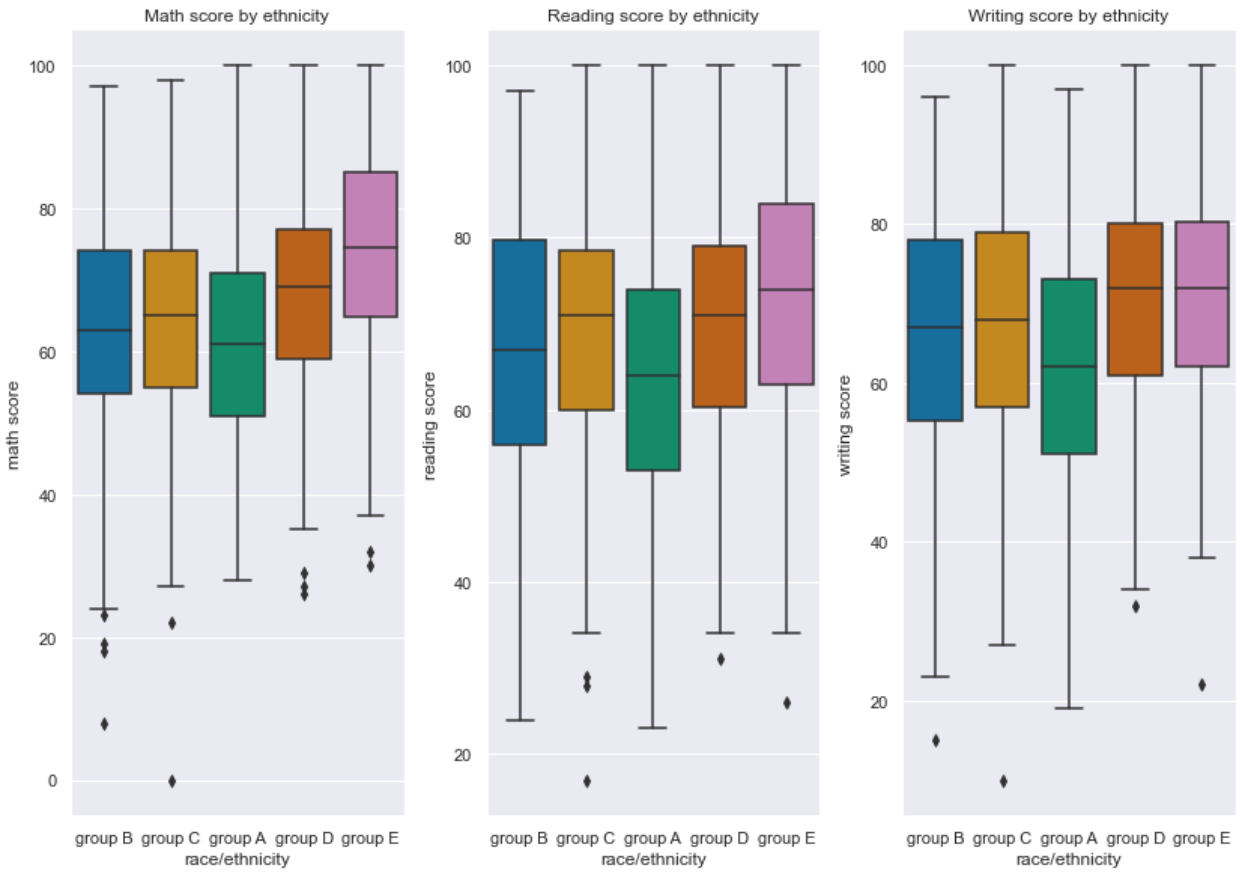
Feature	**Description**	**Breakdown**
Gender	Gender of the student	Female: 518 Male: 482
Race/ethnicity	Race or ethnicity of the student (not disclosed)	Group C: 319 Group D: 262 Group B: 190 Group E: 140 Group A: 89
Parental level of education	Highest education level attained by the student's parents	Some college: 226 Associate Degree: 222 High School: 196 Some high school: 179 Bachelor's Degree: 118 Master's Degree: 59
Lunch	Whether or not the student is given free or reduced lunch	Standard: 645 Free/reduced:355
Test preparation course	Whether or not the student has taken a test preparation course	None: 642 Completed: 355
Math score	Final score on the math exam	Mean: 66.09 Std: 15.16 25% QT: 57 Median: 66 75% QT: 77
Reading score	Final score on the reading exam	Mean: 69.17 Std: 14.60 25% QT: 59 Median: 70 75% QT:79
Writing score	Final score on the writing exam	Mean: 68.05 Std: 15.20 25% QT: 57.75 Median: 69 75% QT:79

The summary data table is fairly informative. We notice that there is close to even amount of male and female students in the study. Next, we see that there are 5 races/ethnicities listed. The data source did not have a data dictionary listing the different groups, so this data is assumed to confidential. Regardless we see that the most common race had 319 students while the least common had 89. Next, the parental level of education field describes the highest education level achieved by either parent. The choices range from “some high school”, meaning the parent did not finish high school, all the way to “master’s degree”. The fewest amount is attributed to Master’s degrees, while the majority of students’ parents held some sort of college degree. It is noticeable that almost 400 either did not finish high school or only graduated high school. Next, the lunch feature describes a student’s lunch aid: whether their lunch is free/reduced due to their economic status or not. In this dataset, about a third of the students receive free/reduced lunch. The “test preparation course” describes whether the student took a non-school related preparation course for the exam. This includes help from 3rd party companies or tutors. In this case, only about a third of the students took a course, while the rest did not. Finally, the last 3 categories describe the test scores in math, reading, and writing final exams. The mean of all exams was in the high 60s, with a similar median.

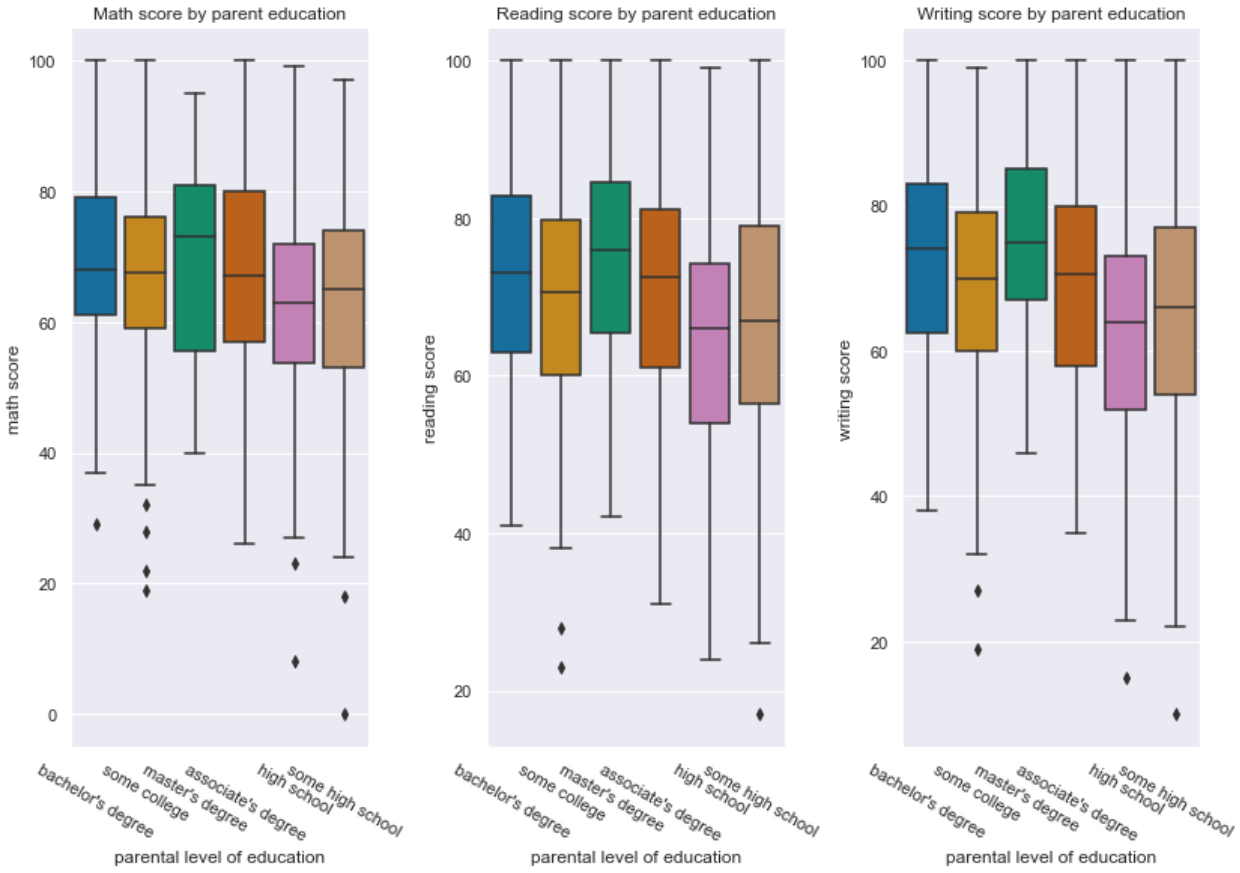
Results and Discussion



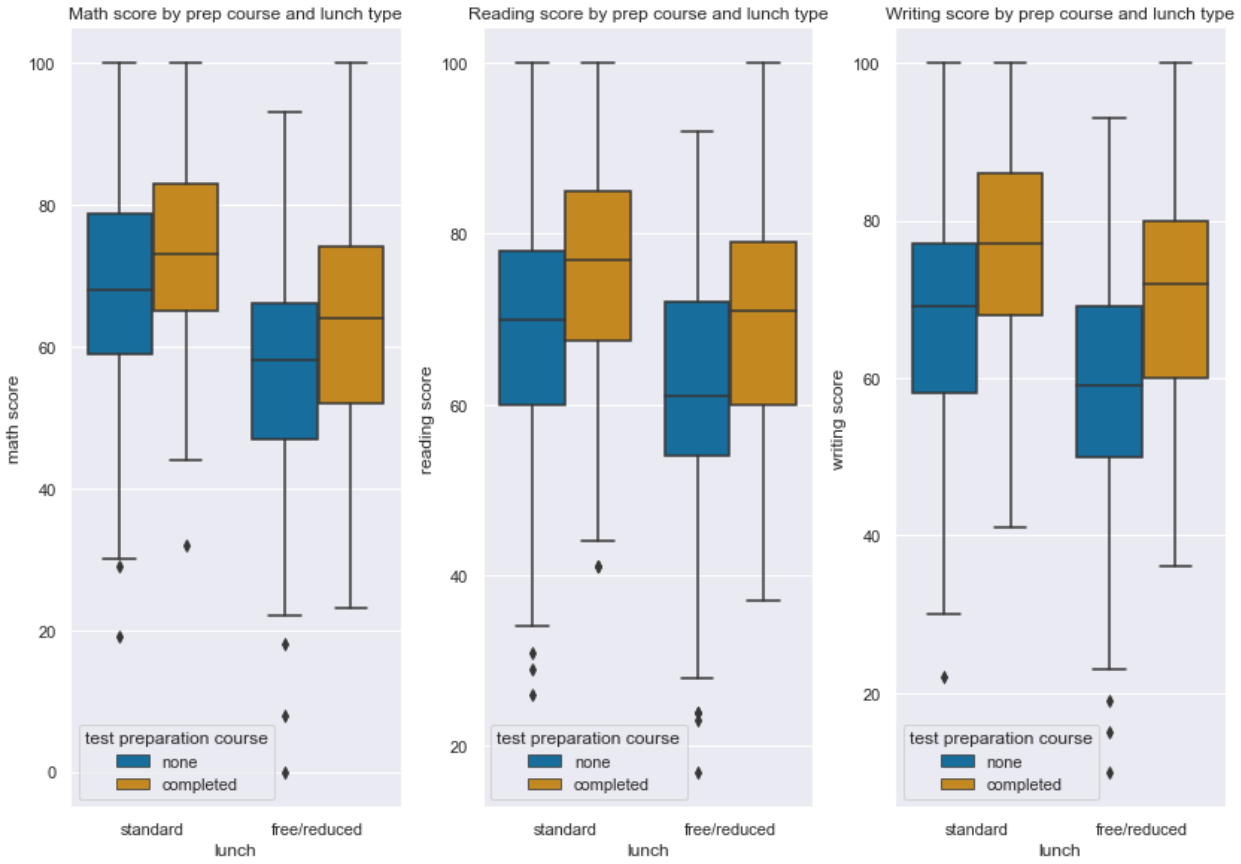
We first analyzed the correlation between gender and the different subjects. We observe a proclivity for male students to have higher math scores, while female students have a higher reading and writing score. The median and first quartile for each graph supports this hypothesis, even though there are very few outliers on the contrary. It also makes sense that those who score well on reading tests also score well on writing tests, as the two are tightly intertwined.



Next, we look at the correlations between race and test scores. We notice that among the three different test, group E usually scores the highest and group A usually scores the lowest. Group A also has the smallest amount of students represented in the study, which leads to a question of underrepresentation or prejudice towards group A, whichever it may be. It is also worthy of mentioning that group E has a noticeable lead in math and reading scores, but are almost equal to group C and D and writing.



In regard to parental level of education, we expect to see a positive correlating with the test score. First, the green “Master’s degree” box is highest in all subject, including its median and third quartile. On the other hand, “high school” and “some high school” are the bottom two categories in all the tests. This is unfortunate but may be cause by the parents’ expectations or responsibilities at home. While some parents will favor education, some will favor hard work and helping around the home, if that is what they grew up with.



Lastly, lunch type and preparation course were analyzed simultaneously since both depend on economic status. Lower economic families are less likely to have standard lunch and are less likely to pay for a prep course. Thus, we see in all the subjects that a preparation course (yellow) was very helpful in increasing scores by at least 5 points. Likewise, we see that the boxplots on the left side of each subplot are dramatically higher than those on the right side, indicating a better scoring student if that child did not have free/reduced lunch. Unfortunately, it can be implied that a lower socioeconomic status will negatively impact test scores, either due to more responsibility at home or simple lack of resources.

Conclusion

In summary, a general trend has revealed itself that a student's socioeconomic status will directly impact their test grades, regardless of subject. The parental level of education will have an impact on the parent's values and expectations of students, who will inevitably follow along. The range of test scores between the highest education level and "high school" was as high as 9 points. Likewise, taking a preparation course for each test outside of school was as impactful as having a standard lunch. A standard lunch means less money and resources at home for the student to use to prepare, as does a lack of preparation course. Interestingly, each one had about the same effect, so all else being equal, a student without a prep course that had a standard lunch scored about the same as a student with a prep course who had free/reduced lunch. Regardless of the causation for these differences in test scores, schools should attempt to minimize the economic effects on student learning through different enrichment programs.

References

Barry, Jennifer. “THE EFFECT OF SOCIO-ECONOMIC STATUS ON ACADEMIC ACHIEVEMENT.” Https://Soar.wichita.edu/, Dec. 2006, soar.wichita.edu/bitstream/handle/10057/616/t06115.pdf?sequence=3&isAllowed=y.

Kara, Orhan, et al. “Factors Affecting Students Grades In Principles Of Economics.” American Journal of Business Education (AJBE), vol. 2, no. 7, 2009, pp. 25–34., doi:10.19030/ajbe.v2i7.4581.

Michubu, Mweti John. “Socio-Economic Factors Influencing Student’S Academic ...” Http://Erepository.uonbi.ac.ke/,erepository.uonbi.ac.ke/bitstream/handle/11295/56391/Mweti_Socioeconomic%20factors%20influencing%20student%E2%80%99s%20academic%20performance.pdf?sequence=3.

Realyvásquez-Vargas, Arturo, et al. “The Impact of Environmental Factors on Academic Performance of University Students Taking Online Classes during the COVID-19 Pandemic in Mexico.” Sustainability, vol. 12, no. 21, 2020, p. 9194., doi:10.3390/su12219194.

Tani, Kawtar, et al. “Evaluation of Factors Affecting Students' Performance in Tertiary Education.” Journal of Pedagogical Research, vol. 3, no. 2, 2019, pp. 1–10., doi:10.33902/jpr.2019252504.