SAT School Participation and Performance: 2012-2013 Data Analysis

Shudi Zhao, Zeyu Guan, Genghua Chen

December 2, 2019

Contents

Introduction

The SAT(Scholastic Aptitude Test) is a standardized test meant to show schools how prepared you are for college by measuring key skills like reading comprehension, computational ability, and clarity of expression and the SAT test also is an entrance exam used by most colleges and universities to make admissions decisions. Its achievements are an essential reference for high school worldwide students to apply for US university admissions and scholarships. Because so many students take the test, it also provides schools with data about how you compare to your peers nationwide.

As you know, you are required to take the SAT or ACT if you're applying to colleges or universities in the United States, since most require you to submit test scores with your application. Depending on where you want to apply, your SAT score can account for more than 50% of the admission decisions, so a strong standardized test score is vital. The SAT is similar to the national college entrance exam in China, but the difference is the SAT can be taken many times, China can only take it once. So we took the SAT test more than once.

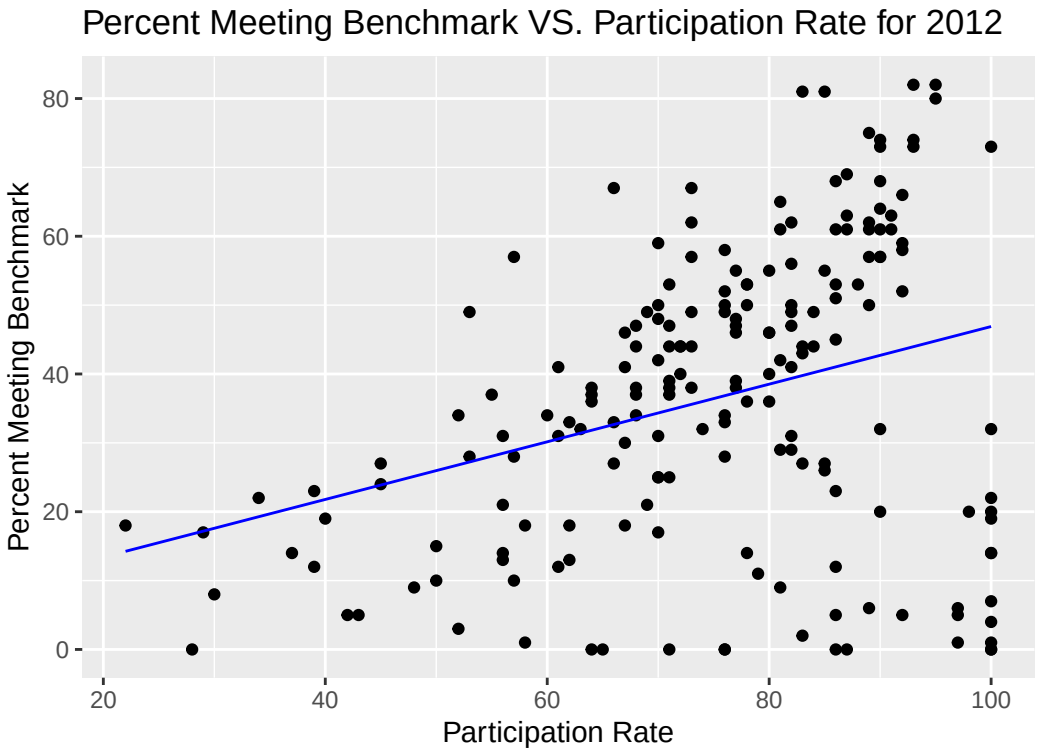
This dataset contains data by school on student SAT for the graduating classes of 2012 and 2013. Also, this dataset was randomly selected 197 high schools and provides 12 variables, including district, school, test takers from 2012, test takers from 2013, test takers change rate from 2012 to 2013, estimate participation rate from 2012, estimate participation rate from 2013, estimate participation change rate from 2012 to 2013, percentage for meeting benchmark from 2012, percentage for meeting benchmark from 2013. They estimated SAT Participation Rate starting from 2012 to 2013, which equals the number of SAT test-taking seniors in the school divided by the number of seniors enrolled in the school as a percentage. Benchmark performance is a measure of a student's SAT score standard. The state can use the SAT Benchmark to measure college and career readiness of groups of students. The College Board conducted extensive research to develop a fair and appropriate indicator of expected college performance. But the Benchmark Performance only reflects the results of only those who participate in the SAT. Because participation rates vary from school to school, the CSDE provides both the estimated Participation rates together with the Benchmark Performance in order to promote fair and valid comparisons across schools. This report provides information to universities or colleges in student selection and the education system on how to influence or improve academics.

But since our dataset is limited, we found another dataset to explain our question.. This dataset contains data by CORGIS Dataset Project. Across the nation SAT score recorded starting from 2005 to 2015. In this data is collected 53 states in United State and provide 99 variables. The data is very detailed, reflecting the examination of each subject. Such as, the average GPA of all students in this state during this year in English or the average GPA of all students in this state during this year in Math. But we only use 6 variables in this dataset: "Total.Math", "Total.Verbal", "Academic Subjects.English.Average GPA", "Academic Subjects.English.Average year", "Academic Subjects.Mathematics.Average GPA", "Academic Subjects.Mathematics.Average year". Both dataset are very tidy, so we don't have to do too much data cleaning.

For this project, after all the discussion, we came up with one question: "What affects the SAT percentage meeting benchmark performance and SAT score?" Our assumption is that the test takers and the participation rate will affect the percent meeting benchmark. We will consider the two cases which is the data for 2012 and the data for 2013. Then, the average GPA for English and Math classes will affect the students SAT score. The number of years English and Math study will also affect the students SAT score.

What affects the SAT percentage meeting benchmark performance and SAT score?

First, we assume that test takers and participation rate will effect the percentage meeting bechmark performance for each schools. We did two plots for participation rate vs. percent meeting benchmark for 2012 and 2013. We also did another two plots for test takers vs. percent meeting benchmark for 2012 and 2013. Then, we did four models for each of those plots, since we saw some positive trend in those plots. Next, we put our models and plots together to see how good is our model. The plot below is one of our plot which is participation rate vs. percent meeting benchmark for 2012.



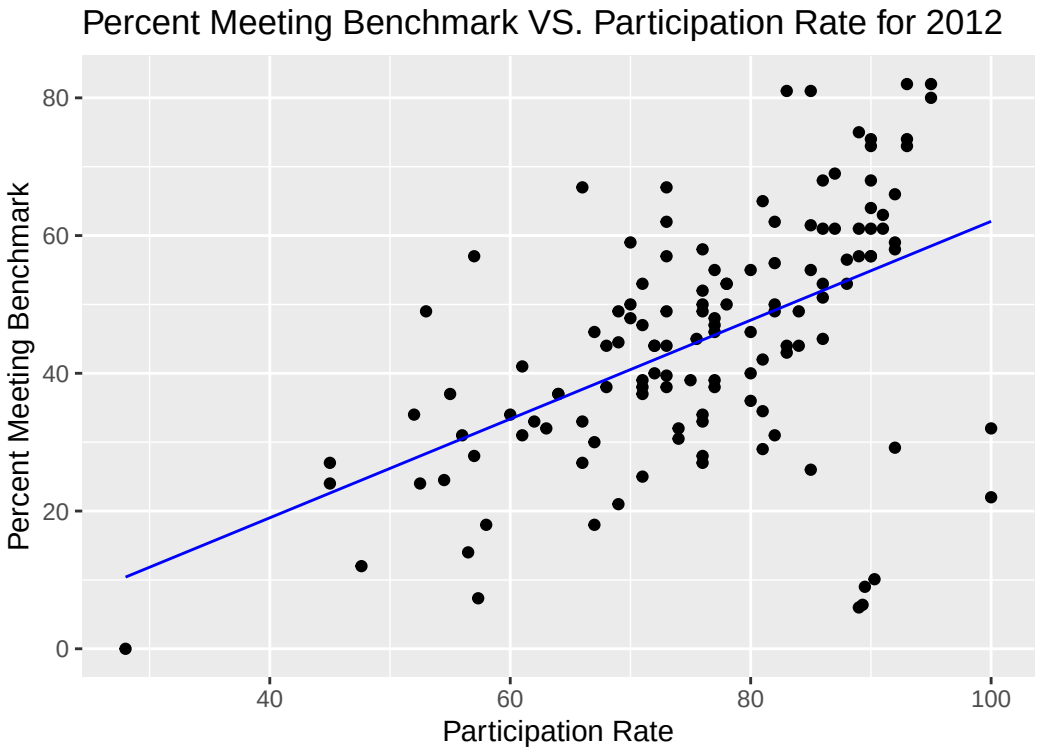
We didn't put all the models here, because we found another approach. We found 197 rows in the table which means this dataset contains 197 high schools, but there are 134 distinct districts in this dataset. There are

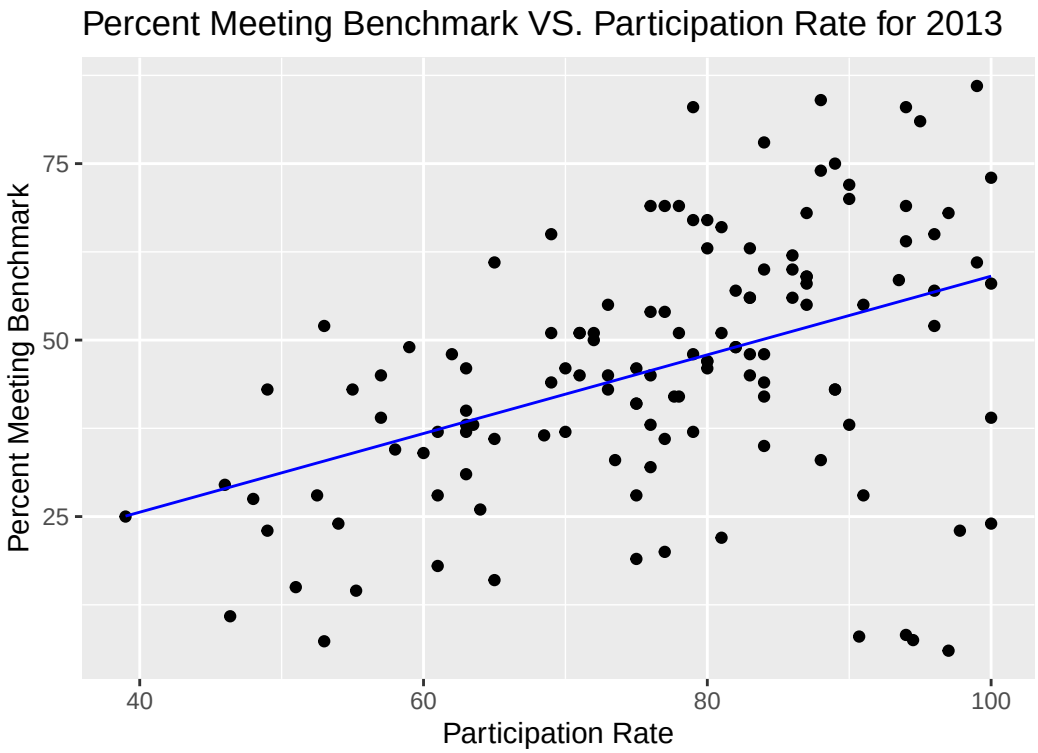
113 districts that only contains one high school and 21 districts contains more than one high schools. Then, we decided to see what affects the SAT percentage meeting benchmark performance for each districts, so we have to group by district and do some modifications. We repeat what we did in the previous plots, and we also obtained four models. First, We did two models for participation rate vs. percent meeting benchmark for 2012 and 2013. Then, We did another two models for test taker vs. percent meeting benchmark for 2012 and 2013.

In total, we did eight models, so we want to know which models have the best performance. We compared the r squared for models by school with the models by district. For instance, the r squared for model one (Percent Meeting Benchmark VS Participation Rate for 2012 by schools) is about 0.09996407 and model three (Percent Meeting Benchmark VS Participation Rate for 2012 by districts) is about 0.2686636. The r square for model three is greater than model one. Then, we also found the performance for the other models by district is also a lot better than the models by schools, so we dicided to display our results by using this new approch: group by district.

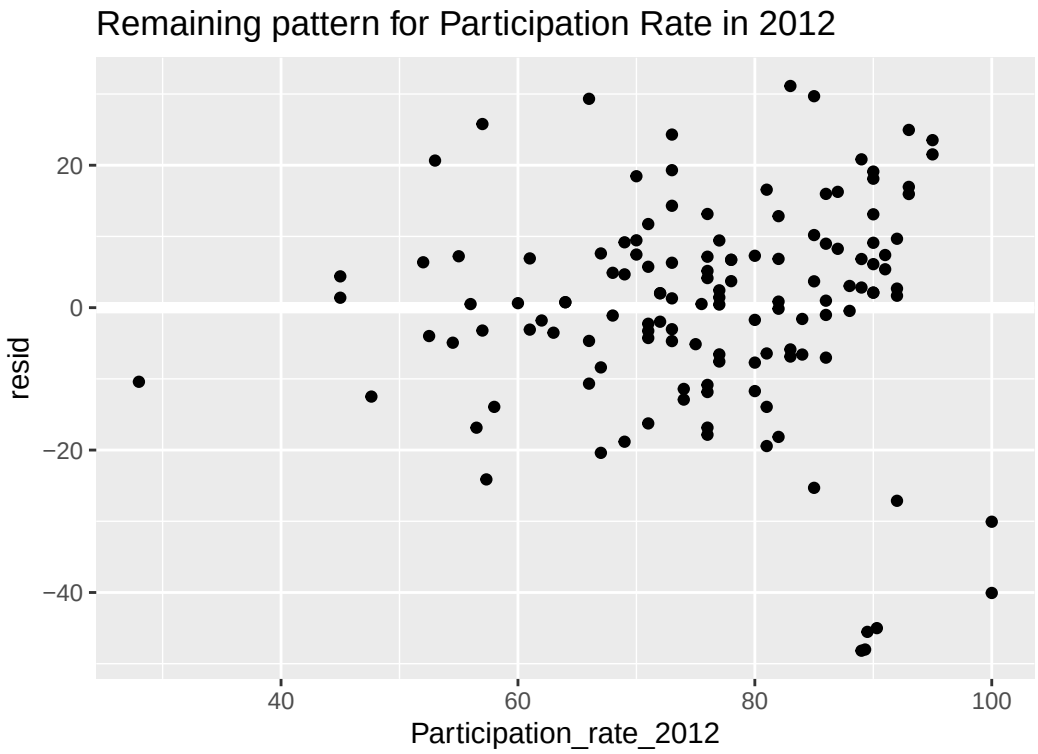
Participation Rate

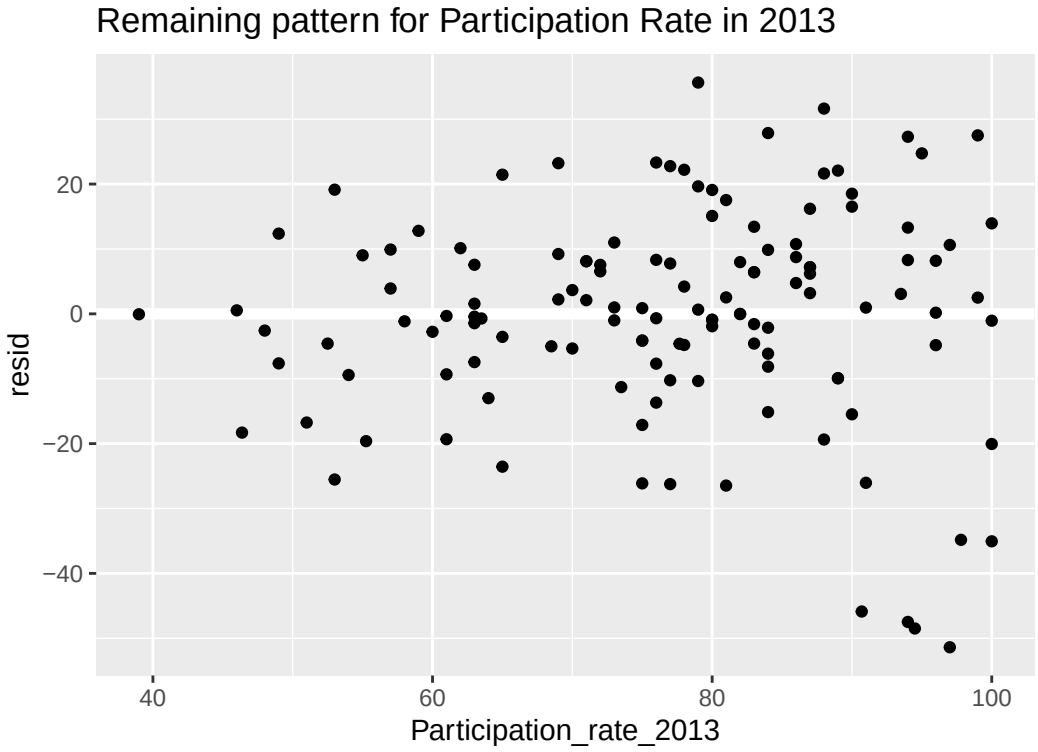
The plots below is for the participation rate vs the percent meeting benchmark for 2012 and 2013 by district. We can see there are some positive trend in this two plots, the percent meeting benchmark increased as the participation rate increased. However, the r squared for both models are small. The r squared for the first model is about 0.2686636 and the r squared for the second model is about 0.1884118. This means the predict power for this two models is very low, but we still can see that the participation rate is slightly effected the percent meeting benchmark.





The two scatter plots below is the residuals for participation rate for 2012 and 2013. In the two plots below, it looks like random noise, suggesting that our model has done a great job of capturing the patterns in this dataset. To prove my assumption, I calculated the average of the residuals. I got 1.045401e-14 for the participation rate in 2012 and -2.301163e-14 for the participation rate in 2013. They are both closed to 0, so my two models captured the most patterns in this dataset.

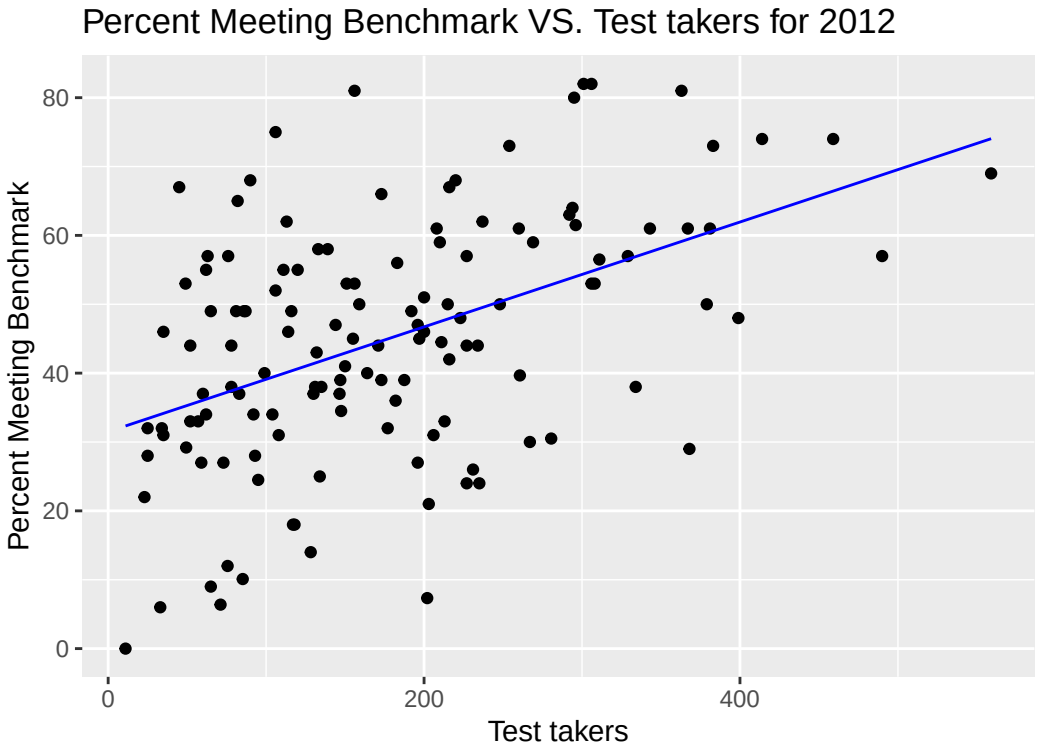


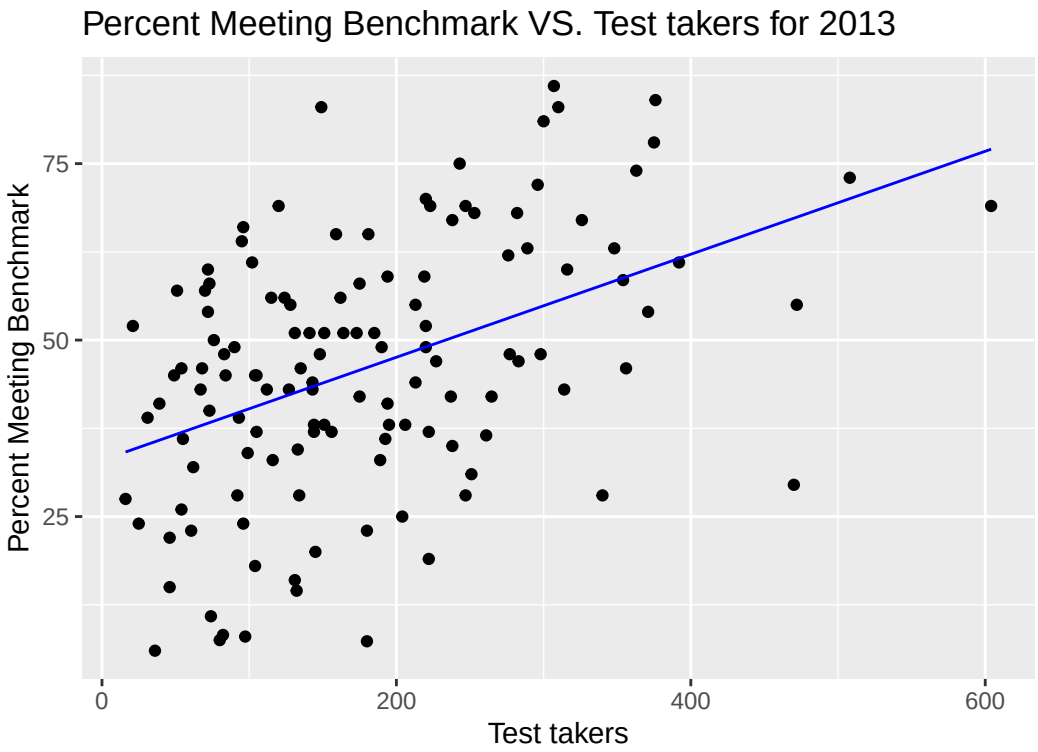


Actually, we can make our models more precisely. We can cut our binwidth to 5 and find the mean for both percent meeting menchmark and participation rate. Then, we can create another model with those modified points to see what does it looks like. However, we didn't figured it out how to plot it. It's a pity that we give up, since we don't have enough time to figure it out.

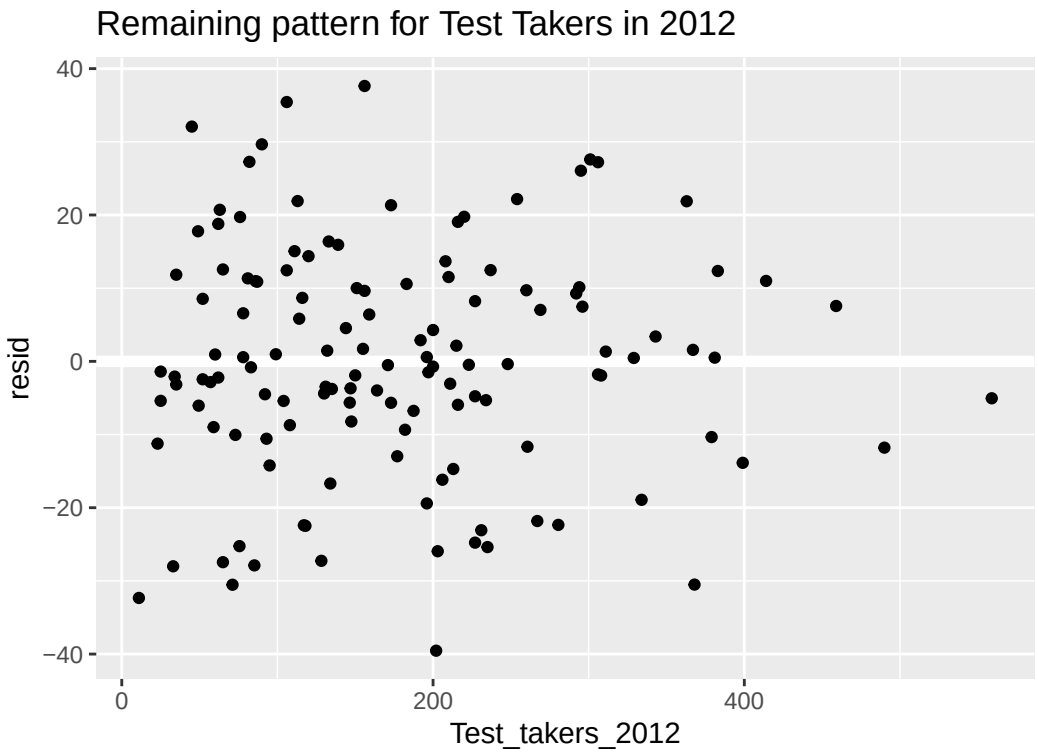
Test takers

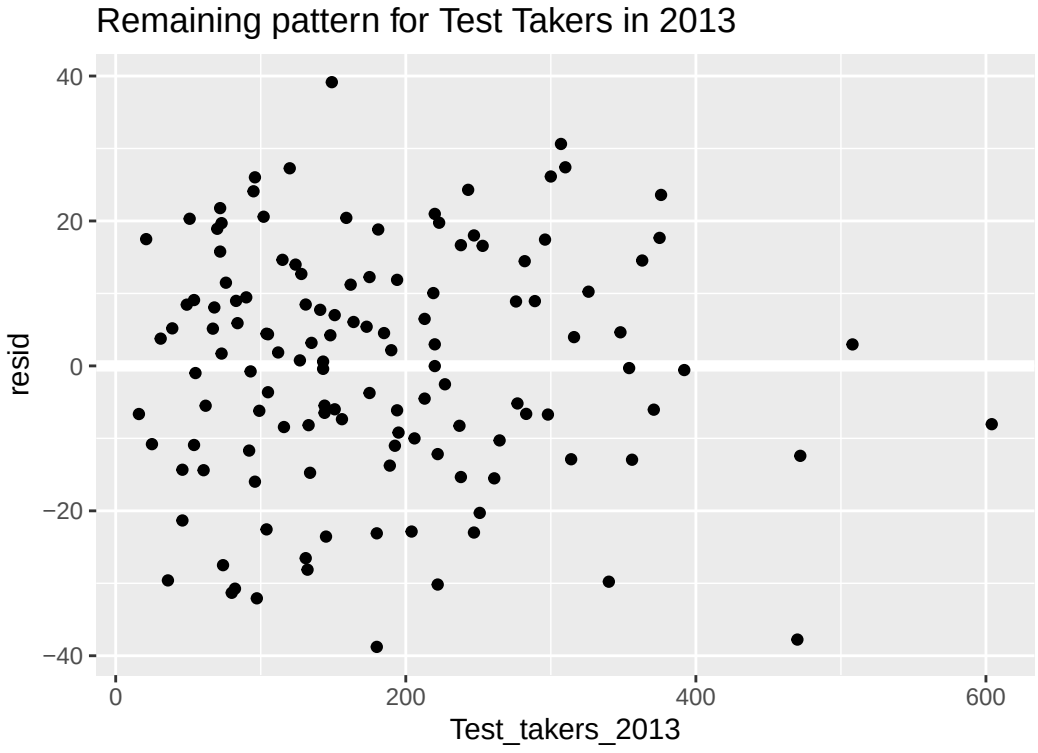
As we have mentioned before, another two models for the test takers vs. percent meeting benchmark for 2012 and 2013 are going to be provided. The reason is simple, first, the variables of the dataset are limited, so there are only two factors in the dataset that can directly affect the percent of the meeting benchmark. Second, to think in normal logic, the data of test_takers might be more realistic than participation, since I personally think the number of test_takers looks more intuitive than the percent of the rate of students who willing to participate in the SAT. Based on the two reasons above, we determined it's necessary to analyze the relationship between test takers and the percent meeting benchmarks in 2012 and 2013, respectively. For convenience, the two plots below are two full data of test_takers VS. percent meeting benchmark in 2012 and 2013. From the two plots below, we can see that there is a positive trend in both plots.





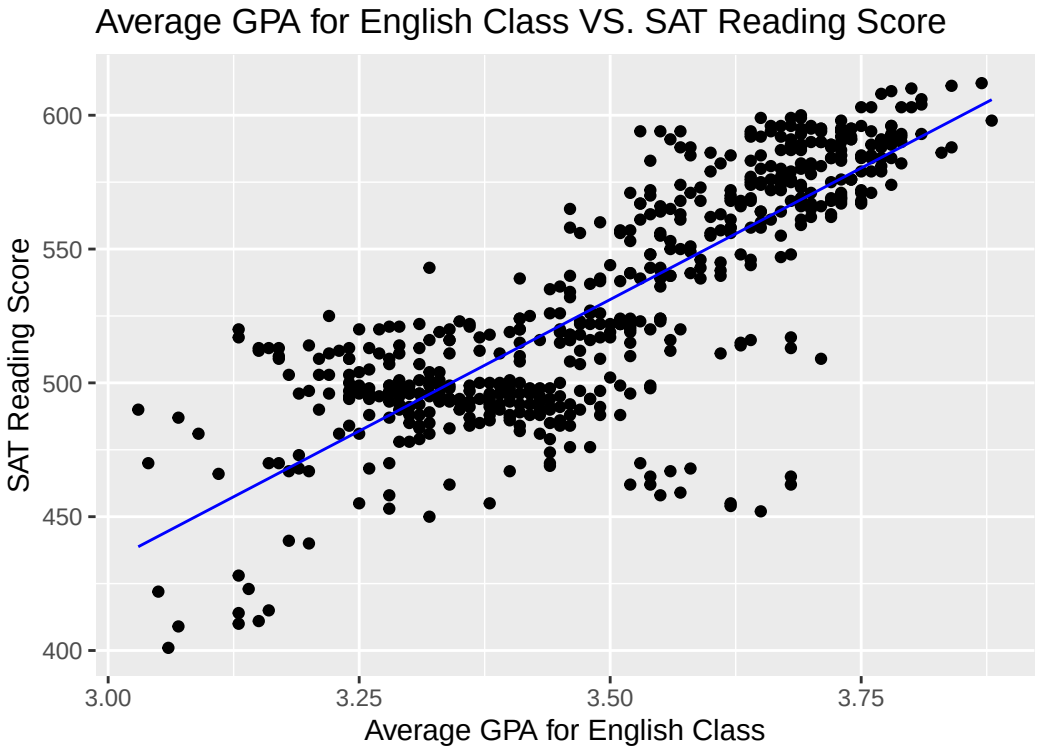
Next, We calculated the r squared for this two models to check the performance for these models. The r squared for the test takers vs. percent meeting benchmark in 2012 is approximately 0.2218803 and the r squared for the test takers vs. percent meeting benchmark in 2013 is approximately 0.1997816. From the r squared, our models didn't do a good job on predicting the percent meeting benchmarks. However, that's not our goal. We just want to know whether the test takers will effect the percent meeting benchmark or not. Then, I also looked at the residuals for this two models (See the two plots below). The residuals for this two models have many random noise, so I calculated their average of the residuals. I got -2.172655e-14 for the participation rate in 2012 and 3.676743e-14 for the participation rate in 2013. Since they are both very closed to 0, so our models captured most of the patterns in this dataset. The result is that test takers does affect percent meeting Benchmark, But again, test takers has no decisive effect on Percent_Meeting_Benchmark.


Remaining pattern for Test Takers in 2012
resid
Test_takers_2012
40
20
0
−20
−40
0
200
400

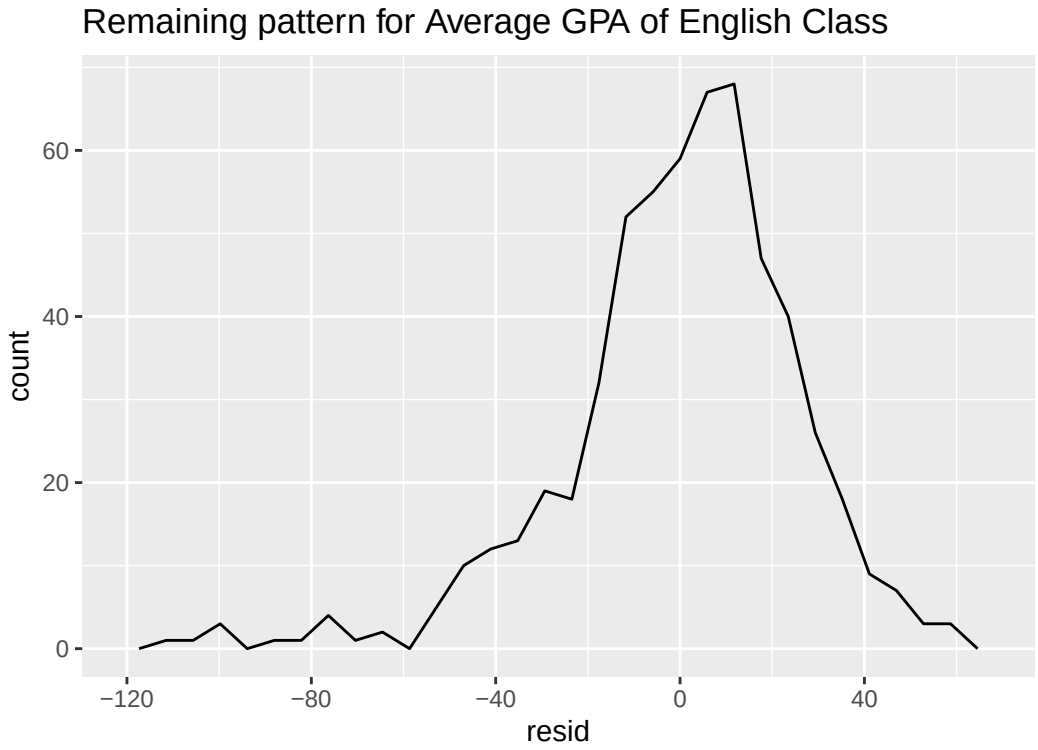

Remaining pattern for Test Takers in 2013
resid
Test_takers_2013
40
20
0
−20
−40
0
200
400
600

The average GPA

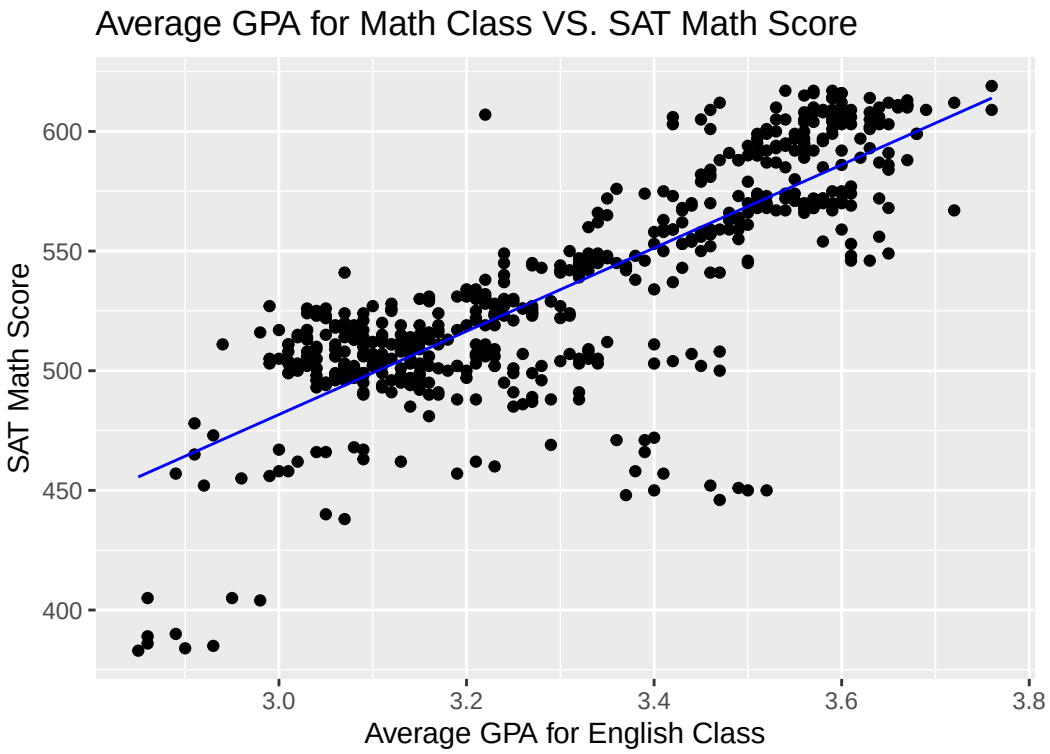
"After a storm comes a calm." Because of the previous limited dataset, the following new rich dataset is going to show us some textbook-like plots that also boosted our team's confidence and excitment. The following model can prove that the average GPA of all students in those states during this year in English played a decisive role on students SAT reading scores. As the average GPA for English increased the SAT reading score is alse increased. We also calculated the r squared for this model, is approximately 0.6771001. It means approximately 67.71001% of the observed variation can be explained by the model's inputs, so the performance of this model is pretty good.



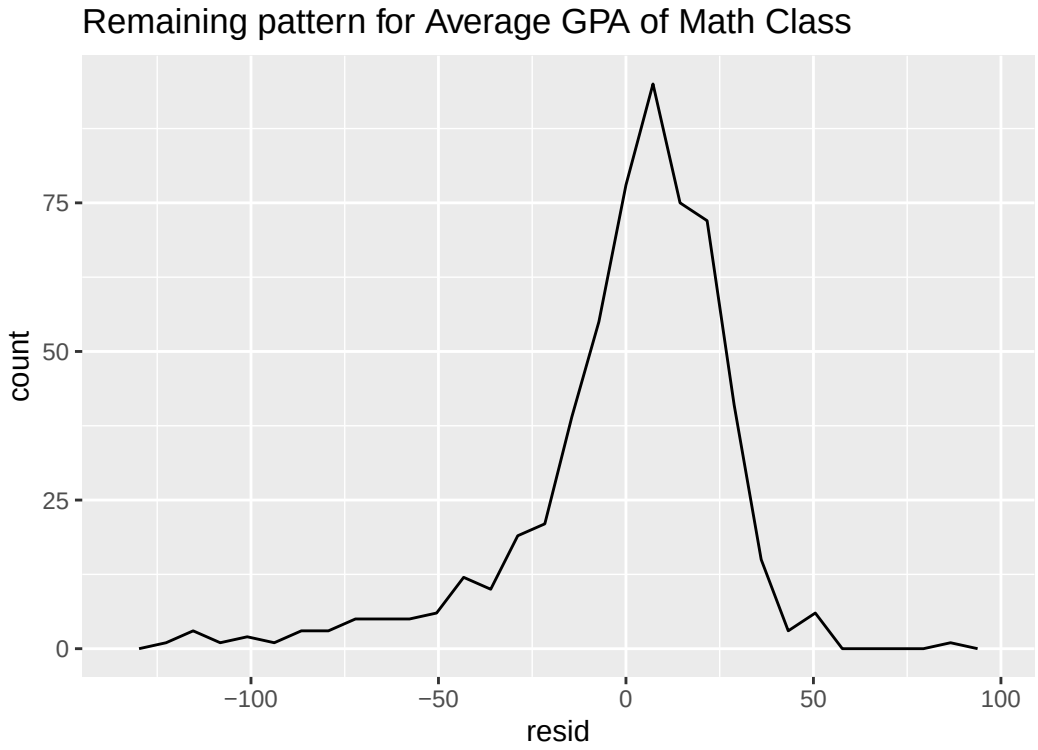
We also calculated the residuals for this model to see if we captured the partterns in this dataset. From the plot below, we can see that the graph it looks like is symmetric to 0, so I calculated the average of the residuals. I got an average of 5.057034e-13 for the residual, which is closed to 0. Thus, it means that our model has captured the most patterns in this dataset. As a result, the average GPA of all students in those states during this year in English have a higher chance to improve students SAT reading scores.



At the same time, in order to dispel people's guess that this might be just a single case, not representative, we choose another example on math as a comparison, and confirmed that the average GPA to the average SAT score really has a decisive effect. In the plot below, we can see that there is also a positive trend. As the average GPA for math increased the average SAT math score is also increasing. Then, we also calculated the r.squared, which is approximately 0.6577988. It means approximately 65.77988% of the observed variation can be explained by the model's inputs, which means our model did a great job.

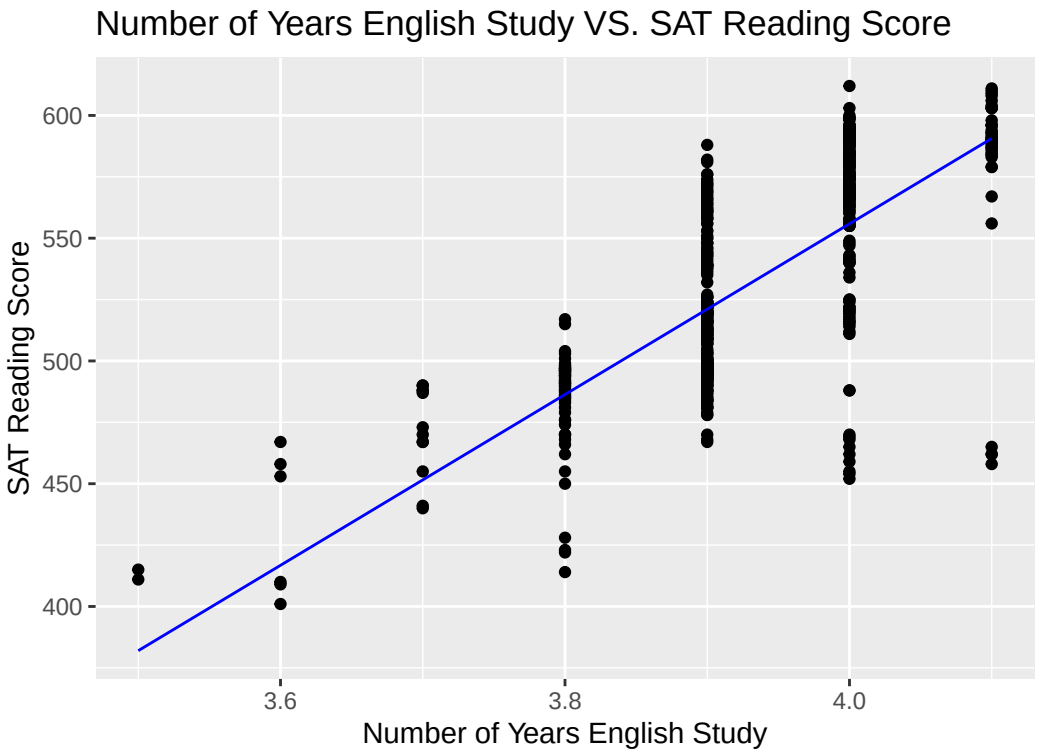


Then, We also calculated the residuals for this model to see if we captured the partterns in this dataset. From the plot below, it looks like the frequency is symmetric to 0, so I calculated the average of the residuals. I got an average of 4.569416e-14 for the residual, which is closed to 0. Thus, it means that our model has captured the most patterns in this dataset. We can conclude that the average GPA of all students in those states during this year in math have a higher chance to improve students SAT math scores.

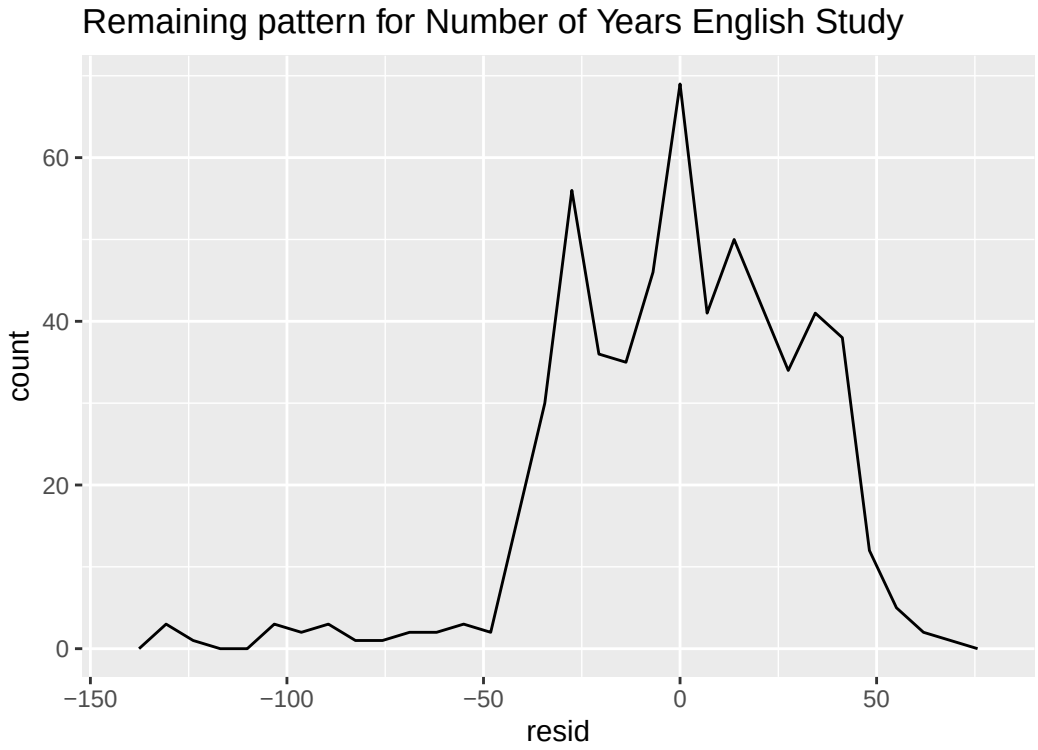


The years of study

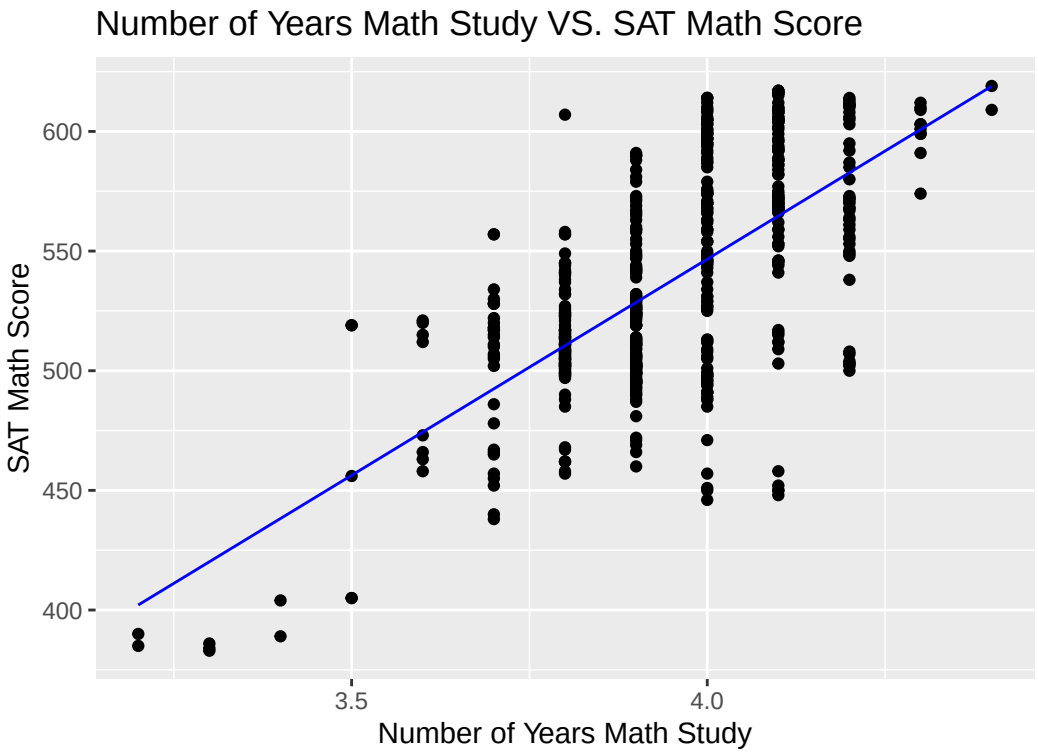
Another factor that seriiously affect SAT score is the average number of years that a student has studied English when they took the exam. The following plot is the average years of English study vs. the SAT reading scores. We can see the positive trend in this plot, the average SAT reading score increased as the average years of English study increased. Then, we calculated the r squared for this model to see the performance of our model, which is approximately 0.5322423. It means approximately 53.22423% of the observed variation can be explained by the model's inputs. In this case, our model performance is just ok.



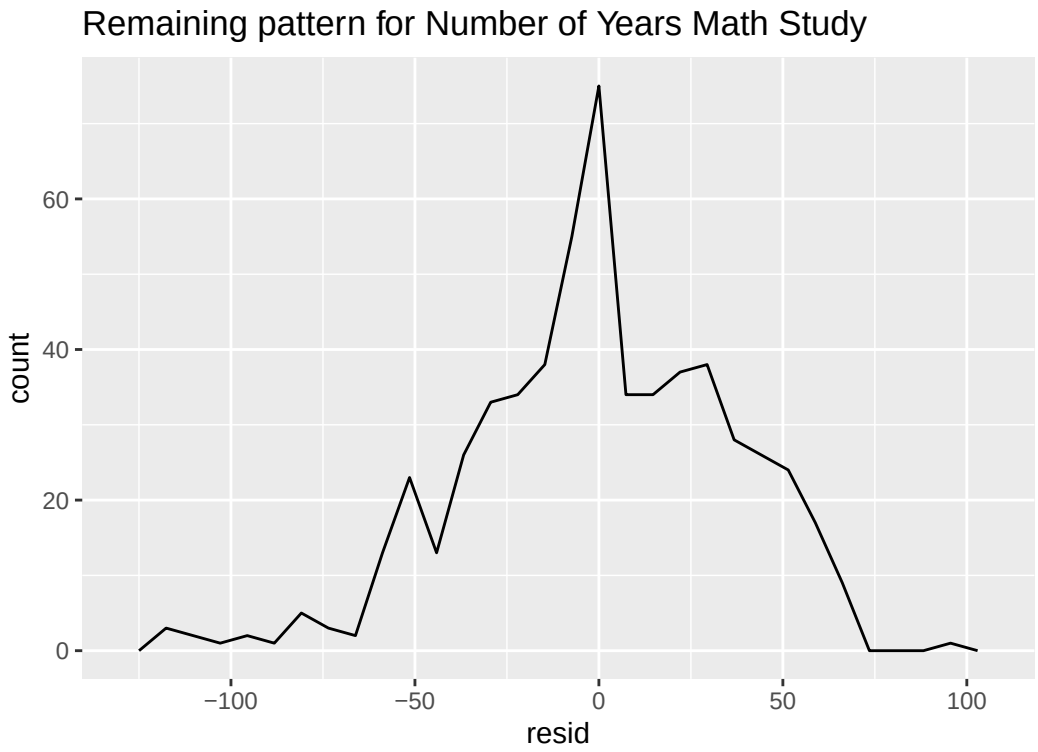
Next, We want to calculate the residuals for this model to see if we captured the partterns in this dataset. From the plot below, it looks like the frequency is symmetric to 0, so I calculated the average of the residuals. I got an average of -4.445211e-12 for the residuals, which is very closed to 0. Thus, it means that our model has captured the most patterns in this dataset. As a result, the average number of years that a student has studied English when they took the exam have a higher chance to improve students SAT reading scores.



Then, we want to know whether the average years of math study will increase the average SAT math scores or not. The plot below showed us the result. We can see there is also a positive trend between this two variables. The average SAT math scores increased as the average years of math study increased. Then we calculated the r squared for this model to see how good our model is. The r squared for this model is approximately 0.4332555. This indicates that approximately 43.32555% of the observed variation can be explained by the model's inputs, so our model performance is not that good.



We also calculated the residuals for this model to see if we captured the partterns in this dataset. From the plot below, it looks like the frequency is symmetric to 0, so I calculated the average of the residuals. I got an average of 1.958536e-13 for the residuals, which is very closed to 0. Thus, it means that our model has captured the most patterns in this dataset. As a result, the average number of years that a student has studied math when they took the exam have a higher chance to improve students SAT math scores.



Conclusion

The first result we found is that participation rate can indeed affect percent meeting benchmark in both 2012 and 2013, respectively. Then, as we expected, another important factor in the dataset, which is test takers can also slightly affect the percent meeting benchmark in both years. For the new dataset, we found that the average GPA for English took a decisive effect on the SAT reading scores and, to increase conviction, we use the same principles to analyze and found the average GPA for math can also take a decisive effect on SAT Math scores. Then, the last result we found is that the more time you spend on studying, the better SAT score you will most likely to get, whether it is reading or math.

In conclusion, education is an important topic in the United states. The SAT is important to colleges because they use it to judge your academic preparedness for college. However, it's also important to colleges because the SAT can reflect a student's real level of math and English, allowing college to better understand a student's real knowlege level. Also, if you're admitted, your SAT score will be incorporated into their yearly SAT statistics. We can also better know which aspects of students are strong, which are weak, can better to give students some necessary help.

Contribution

This report is evenlly contributed by our group members. We used audio calls when we were doing it, so everybody were doing it at the same time. First, we came up with one question together. Then, we started to explore. We finished all those models together, then we started to wrap it up. We collaborate each other, Genghua finished the introduction and the conclusion. Shudi finished the analysis for the two effects, test taker and participation rate. Zeyu finished the analysis for another two effects, average GPA and number of years study. At the end, we checked each others work and improve it to get this final version. Before we submitted, each our team member read this report over one more time to make sure that we didn't make any mistake.