SAT scores in New York City

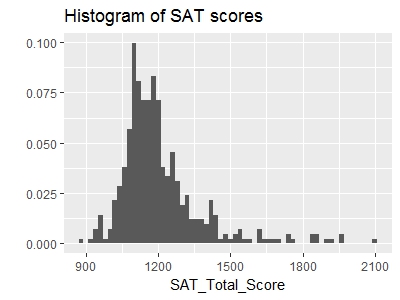
This analysis of SAT scores of high schools in New York City serves exclusively as a descriptive analysis.

Based on a data set with the SAT scores and additional data sets about demographic information, information about each high school as well as a survey the following questions are going to be discussed1:

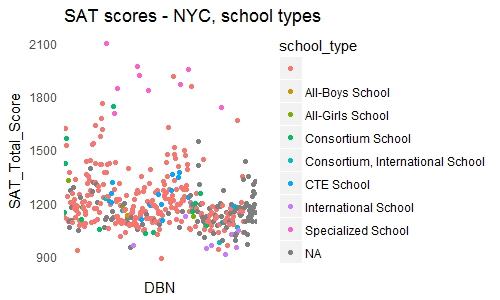
* What variables have an influence (in terms of correlation) on the total SAT scores?
* Is demographic information important?
* Are the SAT fair? Concerning the allegations about the SAT being unfair to certain racial groups or biased by gender?

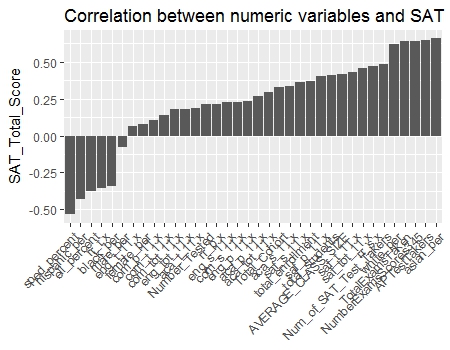
This analysis of SAT scores serves exclusively as a descriptive analysis. Everything that is computed with the data is used to gain some informative insights from and for the current data only. The analysis should not be used as a predictive. Conclusions about future data can only be carefully suggested.

Taking a look at the total SAT scores one can see that most schools reach an average score between 1100 and 1250 points.



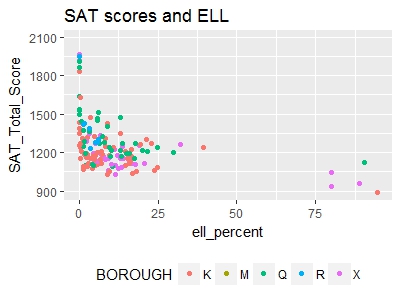
Let's take a look at where the differences in the average scores might come from. Looking at the scores again but colored according to the school type there is an interesting phenomenon going on. The schools with the highest SAT scores are so called Specialized Schools. This does not surprise since those schools were established specially to serve the needs of academically and artistically gifted children.



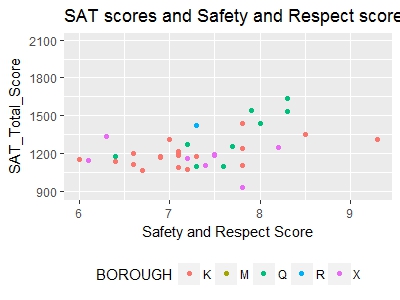


There are several points and insights that might need to be explored.

* ell\_percent and correlate strongly negatively with SAT scores.
* The survey responses do not correlate highly with SAT scores, except the Student response rate and the Safety and Respects scores.
* The percentage of males correlates negatively with SAT scores, whereas the percentage of females correlates positively with SAT scores.
* There seems to be a strong racial inequality in SAT scores.
* The number of AP Test Takers correlates highly positively with SAT scores.



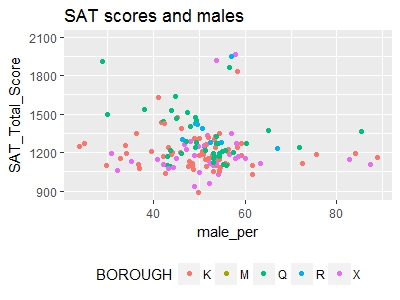
The percentages of students that are eligible for ELL or special-education services correlate highly negatively with SAT scores. There may be an effect on the scores when there are ELL or sped students or not rather than a direct relationship between the percentage and the scores.

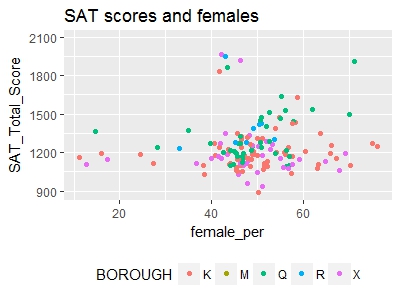


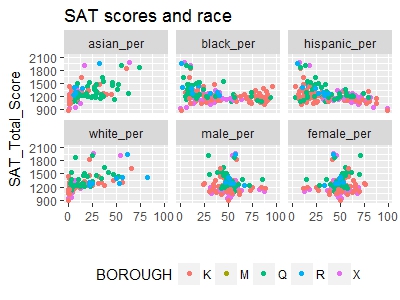
The correlation between SAT scores and the Safety and Respects scores might be explainable. If a school provides a safe and respectful learning environment, students might learn better. Especially if teachers feel safe and respected their patience and ability to teach individually might improve, which also reflect on their students.

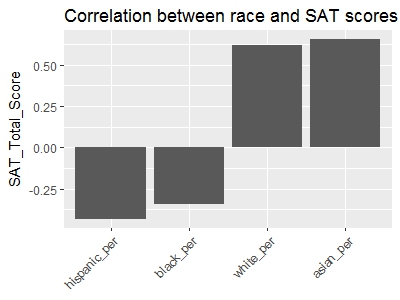
**Exploring race and gender differences in SAT scores**

One of the other angles to investigate involves gender and race and SAT scores. Looking at the percentages of males and females in each school the opposite effect also seen in the correlation can again be found in scatterplots.

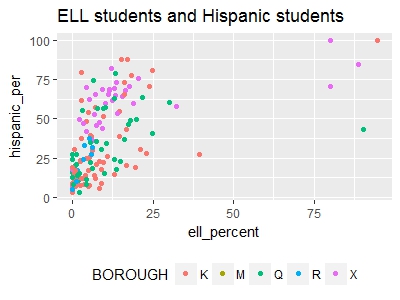


One possible explanation for the positive correlation of SAT scores and percentages of females could be that schools with an extremely high percentage of females tend to be elite schools that focus on the performing arts. This may be a reason for positive correlation between higher female percentages and SAT scores, and the inverse correlation between higher male percentages and lower SAT scores.



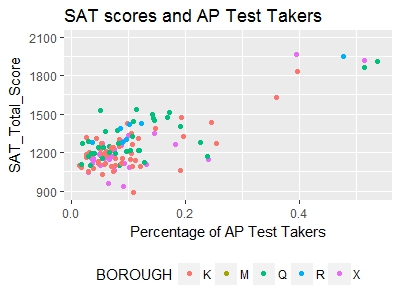


It seems that higher percentages of hispanic and black students correlate negatively with SAT scores whereas higher percentages of white and asian students correlate positively with SAT scores. For hispanic students this may be because schools with a higher percentage of hispanic students tend to also have a higher percentage of ELLs, because they might be more recent immigrants. Looking at the percentage of ELLs and hispanic students enhances this additional correlation.



**AP Scores**

The quite highly positive correlation between the number of students which took an AP test and the SAT scores is worth being further investigated. When looking at the scatterplot above, one can actually see the positive correlation. However there seems to be a cluster in the right top area which may be a group of observations that are biased.



All high schools which have a percentage of AP test takers higher than 30% are listed below. Those are all schools with an admission test so highly selective schools. It seems reasonable that especially those schools would have a high proportion of AP test takers.

Since several states use Advanced Placement data for accountability purposes and to rank high schools a critical analysis is more or less inevitable2. The greatest advantages of AP programs lie in the academic achievements for those students who actually took the test.

**What can NOT be concluded?**

This analysis of SAT scores has shown some interesting insights which are worth to be further investigated. However, there are limits to it (like to any other data analysis). For example, regarding gender and race percentages, the highly negative correlation between hispanic students and SAT scores does not mean, that hispanic students do worse than other students on the SAT in general. It may however indicate together with the highly negative correlation between SAT scores and ELL students that an earlier improvement of English as the teaching and learning language is necessary. Since the SAT scores are the average score for each school a lower value does not mean that sending one's own child to this school will cause a bad SAT score later on.