Bay City Recommendations for Improved STAAR scores

Introduction

I) Problem Statement:

It is a primary goal of every Texas school district to maximize student outcomes as measured by the STAAR. Though Bay City already strives for excellence in educator professional development and supplementary interventional programs, my goal was to examine other indirect factors that may influence STAAR performance.

II) Background:

Bay City ISD is located in Matagorda county. In 2018 the district had 5 schools, 3,562 students, 36% of it's students at the "meets" level for STAAR (compared to a 48% state average), and a district accountability rating of "C".

III) Goal:

This project aims to provide supplementary district level recommendations to improve STAAR meets performance without changing existing academic infrastructure (such as CHAMPS, Read Grant, and Texas Lesson Study). Data was collected from TEA snapshot records from all Texas districts to see which features most closely correlate to higher STAAR performance. Of those features, 4 of them are "actionable", in that direct action by the district can alter them for better or worse. These include :

-Attendance Rate
-Staff % Teachers
-Teacher % With 5 or Fewer Years of Experience
-Teacher Turnover Rate

Datasets

I) TEA District Snapshot 2018:

The Texas Education Agency publishes a yearly snapshot with details about each school district (both charter and non-charter). The 2018 snapshot contains 137 features including identifying information for each district, student and staff demographic information, financial information (including tax rates, fund balances, and expenditures), and testing information broken down by student demographics. These were downloaded in a .csv format directly from the TEA website and then renamed (for example 'DDA00A001219R' was renamed to 'staar all subjects meet') to be more user friendly.

II) District Snapshot 2019 & 2020:

The data from these snapshots were not downloaded directly as a .csv file. Instead the relevant features from the target school district were input directly from the TEA district detail search. The 2019 data was used to verify the accuracy of the model, and the 2020 data was used to forecast this year's results. Along with the data from the district detail search the 2020 data was joined with alterations that included projected improvements to the features listed above. These included optimistic features wherein all actionable changes were perfect (100% attendance rate, 5% teacher turnover rate, 10% of teachers with 5 or fewer years of experience, ect), as well as scenarios with those features altered by reasonable amounts.

III) Unscaledsnap:

After exploratory data analysis the most relevant of the 137 features in the original dataset were cleaned and kept. These included:

- Students % Hispanic
- Students % White
- Students % Asian
- Students % Economically Disadvantaged
- Students % Gifted and Talented Education
- Attendance Rate (2017-18)
- Annual Dropout Rate Gr. 9-12 (2017-18)
- Staff % Teachers

- Staff % Minority:
- Teacher % With 5 or Fewer Years of Experience:
- Teacher Turnover Rate:', '78. Operating Revenue % Federal
- Expenditures % Basic Education Services
- Expenditures % State Compensatory Education:',
- Racial Dif Stu/Tea'
- District Size
- Community Type (Rural, Urban, Suburban, ect)

Data Cleaning and Wrangling

I) Imputing missing values:

TEA records missing values with a period. These were replaced with NaN values for analysis. Most of the missing data was due to student privacy concerns. TEA will not release specific test scores for demographics if the number of students in that demographic group is too small. As a result this meant SAT and ACT results had the most missing data. Since those are too highly correlated to student academic performance to help predict STAAR scores (It should be obvious that a district with higher SAT scores will also have higher STAAR scores), I dropped all standardized testing data altogether. I also dropped longitudinal graduation rates since it would be more difficult to measure long-term graduation rates as an “actionable” feature. For the rest of the dataset I used the mean across all districts to fill in missing values.

II) Correcting erroneous values:

There were also a couple of specific instances where I took the liberty to correct what seemed like erroneous information which I will list here. Richland Collegiate Highschool was recorded as having 615 students per teacher while Imagine International Academy was recorded as 0. These were both replaced with the average ratio.

III) Generating a racial difference feature:

Finally, some research has suggested that minority representation in media has an impact on childhood self-esteem[1]. With this in mind I wanted to know if staff representation had an impact on student performance. To measure this I created a “Racial Dif Stu/Tea” column by averaging the absolute difference between the percentage staff of each race and the corresponding student racial percentage.

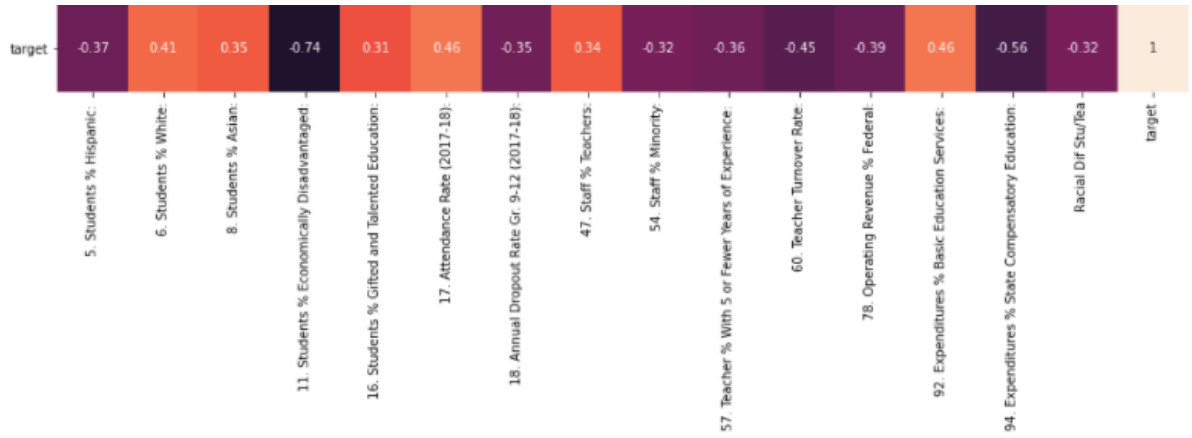
---

[1] Okoro CA, Hollis ND, Cyrus AC, Griffin-Blake S. Prevalence of Disabilities and Health Care Access by Disability Status and Type Among Adults — United States, 2016. MMWR Morb Mortal Wkly Rep 2018;67:882–887. DOI: http://dx.doi.org/10.15585/mmwr.mm6732a3

Exploratory Data Analysis and Initial Findings

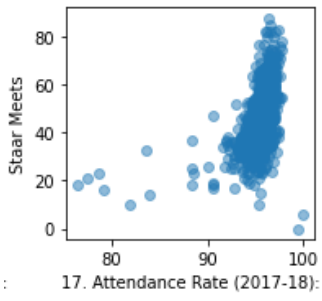
I) Correlations to the target:

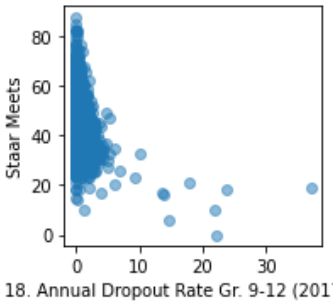
All the numeric features were compared to the target feature using a Pearson score. The following heat map illustrates the features with the highest correlation to the target.

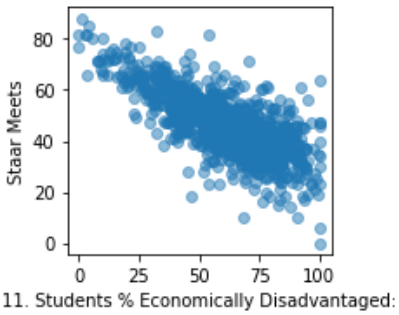


II) Notable feature characteristics:

A few other feature attributes became apparent during EDA. Notably attendance rates did not seem to impact student performance until around 95%. Similarly high drop out rates were associated with poor performance, but after a certain threshold (about 5%) it widened to encompass a range of scores. Other than that, most features were only moderately correlated to the target with the exception of Students % Economically Disadvantaged which was the strongest predictor of STAAR performance.

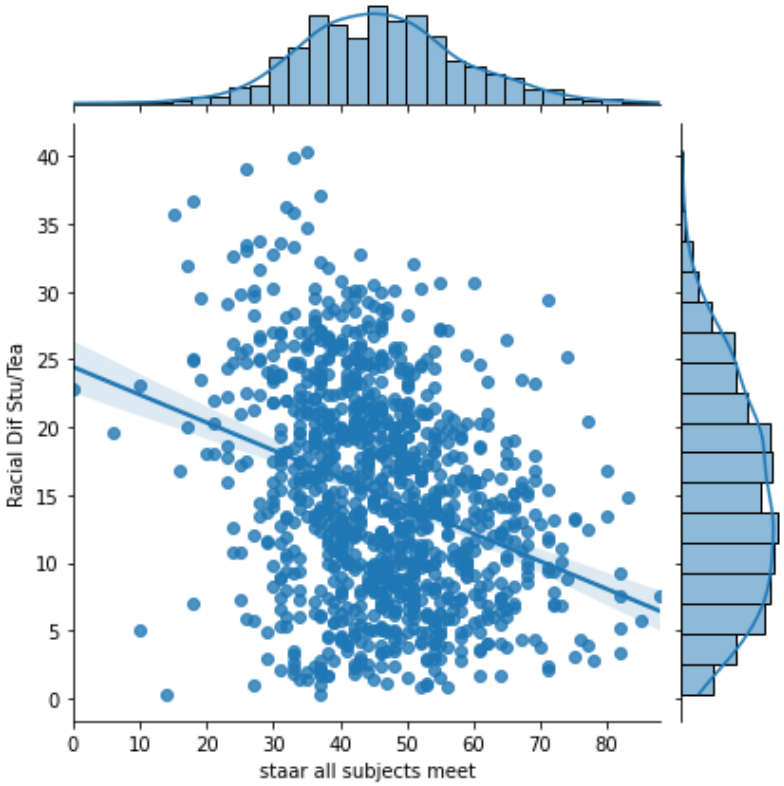




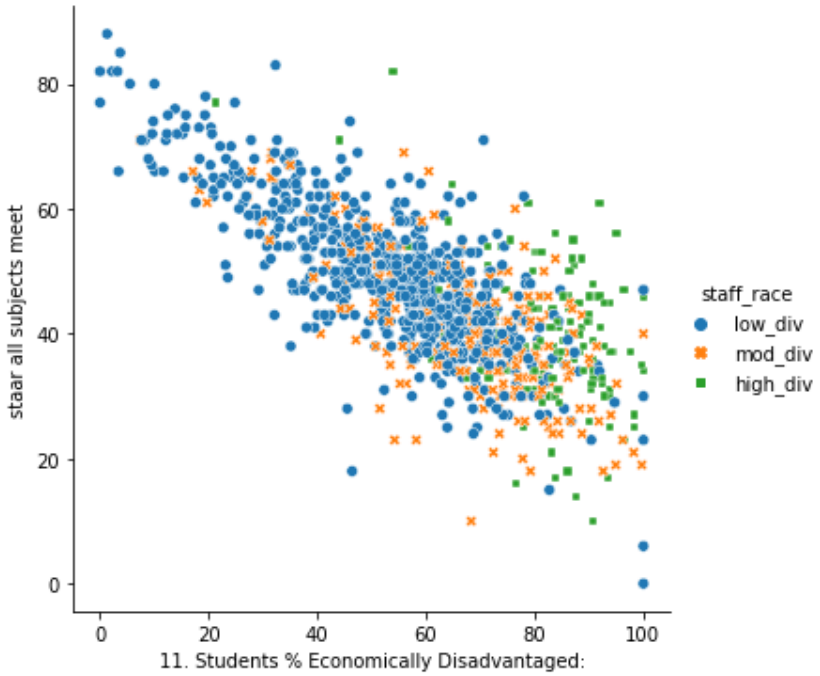


III) Staff Diversity vs District Performance

The fact that the difference in racial makeup between staff and students negatively correlated to student performance seemed to lend credit to the idea proposed above that staff minority representation was important. The less of a difference there was between the racial makeup of the students and staff the better students seemed to do (in a general sense).

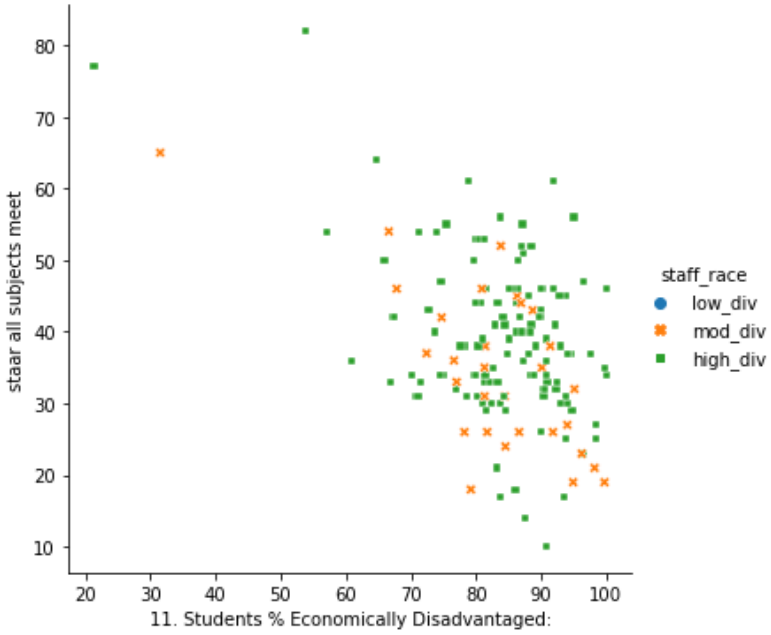


Since there was also a negative correlation between staff % minority and staar performance I wanted to make sure that the above correlation wasn't simply due to schools with a homogenous racial makeup. I suspected that the reason for the negative correlation to staff % minority may have been because higher minority populations were correlated with lower income districts and it might still have been the case that lower income districts with a more diverse staff outperformed lower income districts lacking that diversity. This idea seemed to be illustrated when teachers were grouped by the level of staff minority representation (grouped by 0-33%, 33.1-66%, and 66.1-100%)



However, even factoring for wealth by filtering out districts with less than 66% of the student population labeled as “economically disadvantaged” there was still a negative correlation associated with minority staff, even if it was much smaller than the dataset as a whole.

Eventually it became apparent that there was no clear correlation between staff racial diversity and STAAR performance that I could see. Even filtered for just schools with a high minority population, the diversity of the staff had no clear effect. Staff diversity could no longer be considered an “actionable” feature, though it would still serve the model for predictive purposes.



Results and Recommendations

I) Model Selection:

Several models were experimented with in this project and the results of each of them can be reviewed in the model selection notebook. The model with the most satisfactory results was the gradient boosting model. The basic essence of this model is that it uses a weighted combination of many weaker models to produce a powerful overall prediction. In this case the model was able to account for 82% of the variance from the mean in the training data. When used to predict on the testing data (a random selection of 20% of the overall dataset), this model had a root mean square deviation of 7.26. Essentially this is the percentage point by which the model might be “off” on it’s prediction.

To verify the accuracy of the model and give an informal assessment of Bay City ISD’s performance in relation to its demographics I then retrained the model based on all of the 2018 data and used it to predict the STAAR performance for 2019. The model predicted that 34% of students would attain a score of “meets” or higher on the STAAR against the actual 2019 score of 37%. This was exciting because not only did it imply a high level of accuracy for mode predictions, but it implied that Bay City actually slightly outperformed its predicted performance based on demographics alone.

II) 2020 Prediction and recommendations:

As of the completion of this project, 2020 STAAR results have not yet publicly been released. Based on the model used above, Bay City ISD is predicted to have had 33% of it’s students attain the “meets” STAAR score. However, with certain factors adjusted the model predictions improve significantly. For the purposes of this project I forecasted two difference scenarios for Bay City:

	Attendance Rate	Drop-out Rate	Staff % teachers	Teachers with <5 years exp	Teacher turnover rate	Predicted STAAR meets
Optimum Changes	100% (+5%)	0% (-1.5%)	50% (+6.4%)	10% (-21.8%)	5% (-22.9%)	53.6%(+20.6%)
Moderate Changes	97% (+2%)	1% (-0.5%)	46% (+3.4%)	25%(-6.8%)	17%(-10.9%)	43%(+10%)

Note that exploratory data analysis revealed that attendance and dropout rates did not appear to impact student performance until they met a certain extreme threshold. This means

that retaining experienced educators seems to be the most important factor contributing to student success. Bear in mind also that causation was not a given for these results. It is possible that these features correlate to some unmeasured factor to explain student performance (Much like how the number of sports cars owned could be used to predict a person's wealth, but buying a sports car would not make you a wealthier person). That being said, I believe it is plausible to expect experienced teachers to lead to higher student achievement.

In all the data would suggest that if the district were to cut it's turnover rate by just 11%, which would in turn have a positive impact on the number of experienced teachers in the district, the percent of students to meet the staar test would rise by between 2.7% and 17.3%.

Future Work

Managing teacher turnover is no small task and I'm not aware of any strategies the district is already deploying in this area. I do know that Bay City ISD conducts a small survey for each out-going educator asking them why they are leaving, among other questions. If the results of these surveys were combined into a dataset along with other information about each teacher (examples include: Salary, commute or home address, years teaching, content area, ect) additional data analysis might provide insight into how to improve retention.

Additionally if similar records were available about past hires, machine learning might be useful in determining whether a potential candidate teacher had a high likelihood of leaving the district within a certain number of years.