Author: Tyler Stephens

What have we done since the original project proposal? We first set up two separate Feature Selection python programs dealing with the separate District level ISD data and another one dealing with each individual Highschool in the State of Texas.

With the District level we wanted to see how the different expenditure and revenue of each school district correlated with the various minority students and to see if there was a correlation with the school’s size and student population.

With the Individual Highschool level we wanted to see how the various races of the students correlated with the number of students receiving free lunch with could indicate that school is a poverty-stricken school. Secondly, we wanted to see if the school’s location and Student size had any correlation with average number of students receiving free lunches.

After setting up both and looking at the correlation chart for the district level Feature section we realized that there was too much bias in the data and not much actual information that was useful for determine whether said district was in the poverty threshold, so we decided to drop the district and continue on squarely with the Individual Texas Highschool data.

Once we moved on, we determined that the best indicators through the data selection would be the populations of Hispanic students, Caucasian students, and African American students with the other indicator being Free Lunches per Total Students which is the percentage of free lunches given out per the total number of students in a given school.

That’s not to say that these indicators may not change because as we get access to more data in the years closer to today we may be able to find better indicators in the datasets that are better for predicting whether or not a school is title 1 eligible.