

Executive Summary

Through an analysis of the academic year 2020-2021 data from the New York State Educational Data site, we created a model using teacher experience, teacher credentials, chronic absentee rate, and per-pupil funding to predict individual school success, as measured by a composite score on selected elementary and secondary assessment exams.

Our project serves state and local legislators who determine public and charter school policy, school building, and district leaders, who are primarily responsible for policy implementation, as well as parents and teachers.

We created both linear and non-linear models to study the relationship among all variables. Based on these initial models, we find that the selected variables cannot accurately predict school success. The r-squared value of approximately .4 for our non-linear models indicates that our selected features predict up to 40% of the data variance. Of the features studied, our model shows that chronic absenteeism most accurately predicts school success, with a negative correlation.

Our analysis assists legislators in constructing effective policy by isolating some variables contributing to school success. Given the breadth of data available on the NYSED data site, our model should be further developed to determine how other features interact to predict other measures of school success, such as graduation rates and college enrollment.

The next step is to refine our work toward a more nuanced analysis. One such refinement might predict whether any of the NYSED-reported school features predict how different New York schools serve demographics. Another analysis might determine whether selected features might predict different outcomes for charter schools than for other public schools.