An Inquiry Into Attendance Rates

By Kevin Macias-Matsuura

What and How?





Variables in percentages (0-1)

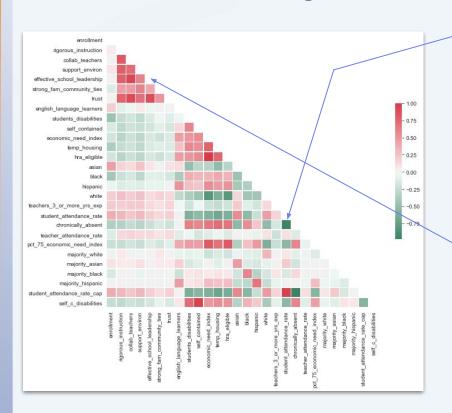
Data: NYC School Quality Report (SQR)

All public middle/elementary (non-charter) schools in NYC

- Charters can choose their evaluation metrics (with approval)

Model / Test The Process Gather / Clean Feature Engineering *p*-value **EDA** Stat - Testing

Statistical Testing



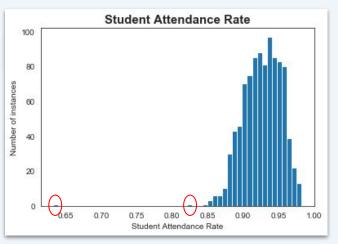
High negative correlation

Chronically Absent Student Attendance Rate

High positive correlation

Came from same section of the questionnaire.

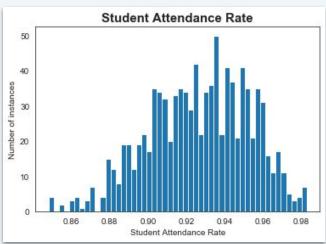




2 outliersNegative skew

0 outliers

Approximately normal distribution



Models

	Features	Measurement	Result
Model #1	27	Train RMSE	0.000908
		Test RMSE	0.015365
		Adj R^2	0.945
Model #2	14	Train RMSE	0.004675
		Test RMSE	0.005376
		Adj R^2	0.966
Model #3	16	Train RMSE	0.004673
		Test RMSE	0.00539
		Adj R^2	0.966

Final Model

Coef Result P-Value Measurement Result Intercept 0.9545 0.000 Train RMSE 0.004676 strong_fam_community_ties 0.0058 0.255 Test RMSE 0.005379 english_language_learners -0.0026 0.202 Adj R^2 0.966 students_disabilities -0.0088 0.006 self_contained 0.0073 0.092 economic_need_index 0.0021 0.280 asian 0.0248 0.000 black 0.0119 0.076 hispanic 0.0156 0.017
strong_fam_community_ties 0.0058 0.255 Test RMSE 0.005379 english_language_learners -0.0026 0.202 Adj R^2 0.966 students_disabilities -0.0088 0.006 self_contained 0.0073 0.092 economic_need_index 0.0021 0.280 asian 0.0248 0.000 black 0.0119 0.076
english_language_learners -0.0026 0.202 Adj R^2 0.966 students_disabilities -0.0088 0.006 self_contained 0.0073 0.092 economic_need_index 0.0021 0.280 asian 0.0248 0.000 black 0.0119 0.076
students_disabilities
self_contained
economic_need_index
asian 0.0248 0.000 0.0119 0.076
0.0119 0.076
hispanic 0.0156 0.017
white 0.0134 0.039
chronically_absent -0.1885 0.000
majority_white -0.0022 0.228
majority_asian 0.0038 0.011
majority_black 0.0016 0.040

13 Features

High Fit / Low Interpretability

Top 3
Chronically Absent
Asian
Hispanic

Takeaways

- Chronic absences are a strong factor in driving attendance down

- Factors one might assume not always statistically impactful

Ex. temporary housing, teacher collaboration, etc

Because money allocated by student attendance, recommend possible implementation of additional questions on the SQR

Ex. How many times a month are you called by the school and/or teachers