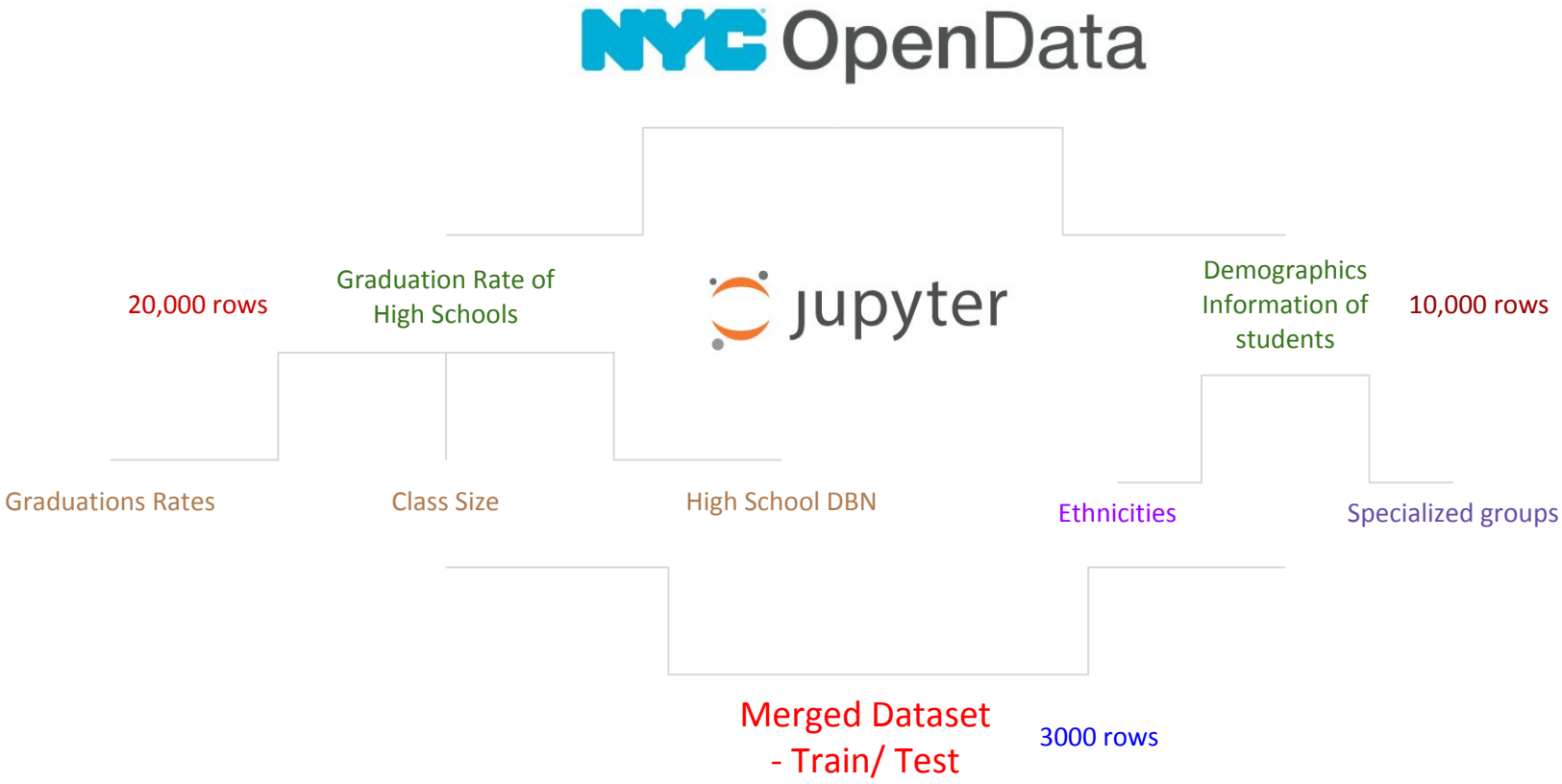
Influence of Demographics on Graduation Rates

Does cultural isolation result in worse performance?

Do schools tailored towards specific groups have better graduation rates than would be expected?

Hypothesis

Students from oppressed groups are less likely to graduate, but that the more students there are of that group at a given school, the more likely those students are to graduate.


NYC OpenData
20,000 rows
Graduation Rate of High Schools
jupyter
Demographics Information of students
10,000 rows
Graduations Rates
Class Size
High School DBN
Ethnicities
Specialized groups
Merged Dataset - Train/ Test
3000 rows

Challenges

Inherent to the data

- **We couldn't access minority graduation rates at specific high schools due to privacy issues.**
- **Students who dropped out or transferred weren't included in cohort size.**

Missing codebook

- **5/6 year cohort columns duplicated data.**
- **Cohort size didn't actually indicate class size.**

Combining datasets

- **We couldn't align 3 year schools between data frames. (due to missing codebook)**
- **Using class size from when the cohort entered 9th grade resulted in data showing graduation size massively larger than cohort size, up to x7000**

Data Cleaning

1. Stripped dataset of whitespaces and converted columns into numeric type.
2. Grouped by Year, DBN and aggregated by max to get one row for each school year.
3. Imputed all the missing student values for different grades & demographics.

Inverse number of students

This set resulted in consistently large P values

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	57.6874	0.571	101.097	0.000	56.568	58.807
ell_percent	1.2601	0.893	1.412	0.158	-0.491	3.011
sped_percent	0.2116	0.742	0.285	0.776	-1.244	1.668
asian_per	-1.2877	1.085	-1.187	0.235	-3.416	0.840
black_per	-1.3439	1.352	-0.994	0.321	-3.997	1.309
hispanic_per	-3.0034	1.368	-2.196	0.028	-5.687	-0.320
invell_num	-0.4765	0.768	-0.621	0.535	-1.983	1.030
invsped_num	0.4115	0.766	0.537	0.591	-1.090	1.913
invasian_num	-0.1824	0.767	-0.238	0.812	-1.688	1.323
invblack_num	-0.1676	0.785	-0.214	0.831	-1.707	1.372
invhispanic_num	-0.8331	0.761	-1.095	0.274	-2.325	0.659
invwhite_num	-1.2078	0.777	-1.554	0.120	-2.733	0.317

Coefficients

After polynomializing and running an RFECV, here are our features and coefficients - >

Our algorithm had an R2 of .42, and had an RMSE of 16, which is 4 percentage points better than predicting the mean.

	feature	coefficient
0	ell_percent	-187.462596
1	sped_percent	-164.993085
2	asian_per	-51.782188
3	black_per	-75.604295
4	ell_percent^2	132.375989
5	ell_percent sped_percent	97.045904
6	ell_percent asian_per	71.845910
7	ell_percent black_per	117.274996
8	ell_percent hispanic_per	23.692446
9	sped_percent^2	357.400553
10	sped_percent asian_per	125.604774
11	sped_percent black_per	-79.417435
12	asian_per^2	64.672046
13	asian_per black_per	90.109775
14	asian_per hispanic_per	-32.851263
15	black_per^2	52.329196
16	black_per hispanic_per	-37.366120

Example

As the percent of the student body that's black increases, the less the graduation rate decreases.

Black %	% and $\%^2$ coef	% coef
20%	-13%	-15%
60%	-27%	-45%
90%	-26%	-68%

Conclusions

1. Specialized students groups such English Language learners and Special Ed.

   students have lower graduation rates.

2. Minority Groups such as Asians and Black students also graduate at a lower rate.

3. However, as the percentage of a group of students increase in a school, their graduation rate decreases at a much lower rate and eventually begins to increase again.

Applications

Suggestion 1	• Graduation requirements might need to be revised for specialized groups.
Suggestion 2	• Further assistance needs to be provided to minority students like the SEEK program of CUNY.
Suggestion 3	• Revise acceptance policies or affirmative action policies.
Suggestion 4	• Provide more support to students who are part of a school's smaller minorities.

Further Research:

- We can use financial information like household income to analyze its effect on graduation rate on top of the demographics.
- How does the geographic location factor into getting into certain schools for minority groups?
- We can use survey information to investigate how students and parents view schools and how it correlates to graduation rate?
- Use population graduation rates of minorities as a balancing feature