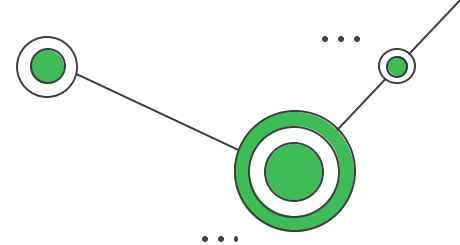


Dev10 Capstone

New York Schools' Financing and
Its Effects on Student
Performance

Meet Our Team!



AZURE ELLER



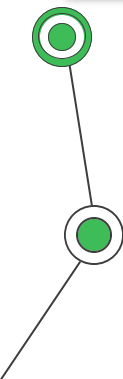
SEUNG MIN KIM



WIDNIE DORILAS



JULIANN GROGLIO



Overview



Project Breakdown

- Explore 2019 education finance and demographic datasets to find what factors contribute to drop out/graduation rates



Data Used

- [US Census Bureau 2019 Education Finance Data](#)
- [NYSED Enrollment Data 2019](#)
- [NYSED Graduation Data 2019](#)



Our Process

- Clean and transform raw data
- Create producer/consumer pipeline
- Analyze and create visualizations/ML model



Main Goal

- Find key features for reducing and predicting dropout rates



Initial Questions



- What are the top 5 districts with the highest and lowest female and male ratios?
- Do districts with higher female to male ratios have higher graduation rates?
- How does a district's homelessness rate affect its graduation rate?
- ★ Does the amount of pupil support expenditures influence Advanced Regents diplomas acquired?
- What is the racial distribution among all school districts?



Initial Questions



- Is there a correlation between property taxes and enrollment?
- ★ What are the top 5 districts with the highest dropout rate?
- What is the racial makeup of the top 5 districts with the highest dropout rate?
- ★ Does an increase in total staff wages increase number of students graduating?
- ★ What is the correlation between the number of economically disadvantaged students and graduation rate?
- **ML:** Can we predict school district dropout rates based on financial allocations and demographic factors?

US Census Bureau 2019 Education Finance Data

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	STATE	IDCENSUS	NAME	CONUM	CSA	CBSA	SCHLEV	NCESID	YRDATA	V33	TOTALREV	TFEDREV	C14	C15	C16	C17	C18
2	01	01500100100000	AUTAUGA COUNTY SCHOOL DISTRICT	01001	N	33860	03	0100240	19	9094	87938	7518	2071	2056	280	0	
3	01	01500200100000	BALDWIN COUNTY SCHOOL DISTRICT	01003	380	19300	03	0100270	19	32267	396834	26739	7240	6905	779	0	
4	01	01500300100000	BARBOUR COUNTY SCHOOL DISTRICT	01005	N	21640	03	0100300	19	755	9877	2175	1180	289	101	0	
5	01	01500300200000	EUFAULA CITY SCHOOL DISTRICT	01005	N	21640	03	0101410	19	5427	42165	4279	1335	961	67	0	
6	01	01500400100000	BIBB COUNTY SCHOOL DISTRICT	01007	142	13820	03	0100360	19	3236	34793	3759	1046	912	130	0	
7	01	01500500100000	BLOUNT COUNTY SCHOOL DISTRICT	01009	142	13820	03	0100420	19	7651	78083	7000	1924	1803	270	0	
8	01	01500500200000	ONEONTA CITY SCHOOL DISTRICT	01009	142	13820	03	0102550	19	1503	15595	1023	196	314	39	0	
9	01	01500600100000	BULLOCK COUNTY SCHOOL DISTRICT	01011	N	N	03	0100480	19	1438	15961	2754	1057	453	209	0	
10	01	01500700100000	BUTLER COUNTY SCHOOL DISTRICT	01013	N	N	03	0100510	19	3015	32230	4810	1606	857	199	0	
11	01	01500800200000	ANNISTON CITY SCHOOL DISTRICT	01015	N	11500	03	0100090	19	1947	24596	4194	1729	637	185	0	
12	01	01500800300000	CALHOUN COUNTY SCHOOL DISTRICT	01015	N	11500	03	0100540	19	8548	91643	8633	2355	2645	345	0	
13	01	01500800400000	JACKSONVILLE CITY SCHOOL DISTRICT	01015	N	11500	03	0101860	19	1622	17069	1766	619	361	59	0	
14	01	01500800500000	PIEDMONT CITY SCHOOL DISTRICT	01015	N	11500	03	0102760	19	1151	12561	1566	434	254	92	0	
15	01	01500880100000	OXFORD CITY SCHOOL DISTRICT	01015	N	11500	03	0102635	19	4083	49006	4296	957	1172	162	0	
16	01	01500900100000	CHAMBERS COUNTY SCHOOL DISTRICT	01017	194	46740	03	0100600	19	3545	38282	5043	1416	943	180	0	
17	01	01500900200000	LANETT CITY SCHOOL DISTRICT	01017	194	46740	03	0101980	19	925	10172	1589	589	263	81	0	
18	01	01501000100000	CHEROKEE COUNTY SCHOOL DISTRICT	01019	N	N	03	0100630	19	3934	46305	4763	1159	1019	167	0	
19	01	01501100100000	CHILTON COUNTY SCHOOL DISTRICT	01021	142	13820	03	0100660	19	7837	73592	8559	2160	1885	299	0	
20	01	01501200100000	CHOCTAW COUNTY SCHOOL DISTRICT	01023	N	N	03	0100690	19	1300	17395	2564	706	466	122	0	
21	01	01501300100000	CLARKE COUNTY SCHOOL DISTRICT	01025	N	N	03	0100720	19	2460	28461	4062	1199	741	182	0	
22	01	01501370100000	THOMASVILLE CITY SCHOOL DISTRICT	01025	N	N	03	0103300	19	1158	12713	1293	268	301	56	0	
23	01	01501400100000	CLAY COUNTY SCHOOL DISTRICT	01027	N	N	03	0100750	19	1868	19529	2099	601	504	75	0	
24	01	01501500100000	CLEBURNE COUNTY SCHOOL DISTRICT	01029	N	N	03	0100780	19	2512	27039	2908	661	787	98	0	
25	01	01501600100000	COFFEE COUNTY SCHOOL DISTRICT	01031	222	21460	03	0100810	19	2388	24027	2371	759	501	116	0	
26	01	01501600200000	ELBA CITY SCHOOL DISTRICT	01031	222	21460	03	0101260	19	622	7927	1207	299	196	38	0	

NYSED 2019 Enrollment Database

ENTITY_CD	ENTITY_NAME	YEAR	NUM_ELL	PER_ELL	NUM_AM_IN	PER_AM_INI	NUM_BLACK	PER_BLACK	NUM_HISP	PER_HISP	NUM_ASIAN	PER_ASIAN	NUM.
000000000001	NYC Public Sch	2019	136480	15	10328	1	205266	22	377831	41	165994	18	
000000000002	Large Cities	2019	16546	16	538	1	42273	42	31736	31	6668	7	
000000000003	High Need/Res	2019	33092	16	574	0	47933	23	83955	41	7700	4	
000000000004	High Need/Res	2019	2062	1	2365	2	4706	3	10781	8	1227	1	
000000000005	Average Need	2019	27677	4	2770	0	46283	6	110378	15	27470	4	
000000000006	Low Need Dist	2019	12091	3	530	0	14120	4	46628	13	41132	11	
000000000007	Charter School	2019	10102	7	1499	1	77919	53	52595	36	4065	3	
000001000000	ALBANY Count	2019	2032	5	68	0	7914	20	3663	9	3640	9	
000002000000	ALLEGANY Cou	2019	4	0	22	0	65	1	94	2	46	1	
000003000000	BROOME Coun	2019	531	2	69	0	2386	9	1747	7	859	3	
000004000000	CATTARAUGUS	2019	21	0	781	6	180	1	342	3	115	1	
000005000000	CAYUGA Count	2019	70	1	14	0	277	3	383	4	57	1	
000006000000	CHAUTAUQUA	2019	684	4	192	1	363	2	2749	15	102	1	
000007000000	CHEMUNG Cou	2019	45	0	17	0	848	8	353	3	185	2	
000008000000	CHENANGO Co	2019	26	0	32	0	112	2	181	3	60	1	
000009000000	CLINTON Coun	2019	27	0	27	0	239	2	145	1	96	1	
000010000000	COLUMBIA Cou	2019	326	5	7	0	427	6	688	10	255	4	
000011000000	CORTLAND Col	2019	39	1	1	0	100	2	162	3	39	1	
000012000000	DELAWARE Col	2019	44	1	12	0	67	1	267	5	42	1	
000013000000	DUTCHESS Cou	2019	1371	4	29	0	4615	12	7563	20	1253	3	
000014000000	ERIE County	2019	8128	7	998	1	24039	20	11782	10	6672	5	
000015000000	ESSEX County	2019	6	0	6	0	29	1	30	1	18	1	
000016000000	FRANKLIN Cou	2019	7	0	935	13	32	0	74	1	25	0	

NYSED 2019 Graduation Rate Database

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	report_scl	aggregatic	aggregatic	aggregatic	aggregatic	entity_in	lea_beds	lea_name	nrc_code	nrc_desc	county_cc	county_na	nyc_ind	boces_coc	boces_na	members	members	members	subgroup	subgroup
2	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	1	All Studer
3	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	2	Female
4	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	3	Male
5	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	4	American
6	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	5	Black
7	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	6	Hispanic
8	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	7	Asian/Pac
9	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	8	White
10	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	9	Multiracia
11	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	10	General E
12	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	11	Students v
13	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	12	Not Englis
14	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	13	English La
15	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	14	Formerly I
16	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	15	Economic
17	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	16	Not Econo
18	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	17	Migrant
19	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	18	Not Migra
20	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	20	Homeless
21	2019-20	0	Statewide	0	All Districts and Charters											6	152	2014 Total	21	Not Home

ETL Process



Extraction

- Downloaded CSVs from each site
- Uploaded CSVs into storage container




Transformation

- Dropped NAN rows
- Created consistent naming convention
- Modified datatypes
- Dropped unneeded columns



Load

- Create DDL schema on SQL server
- Load data frames through pipeline into SQL database



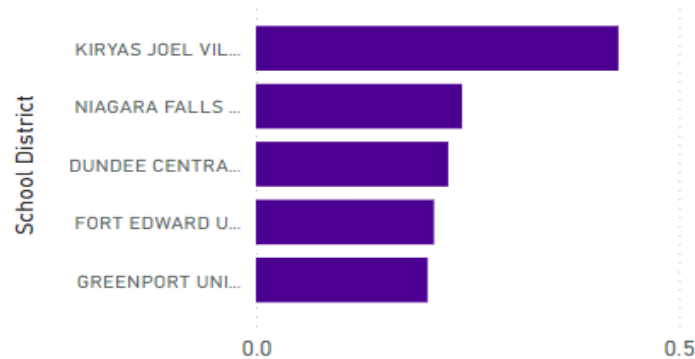
4 Main Questions

What is the racial makeup of the top 5 districts with the highest dropout rate?

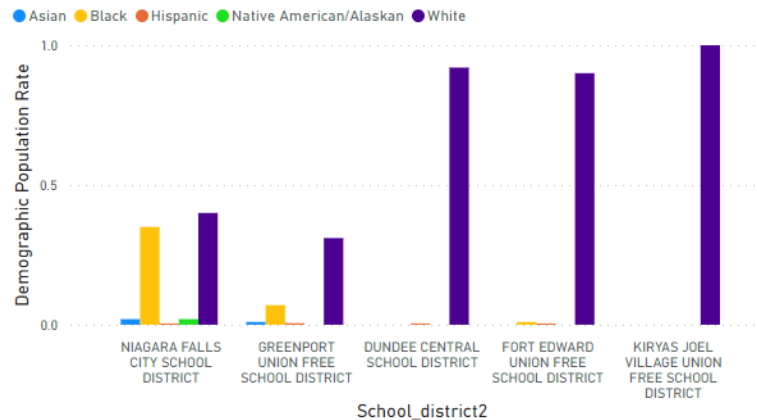
Findings

- No consistent pattern in racial distribution between school districts

School Districts with Highest Dropout Rate



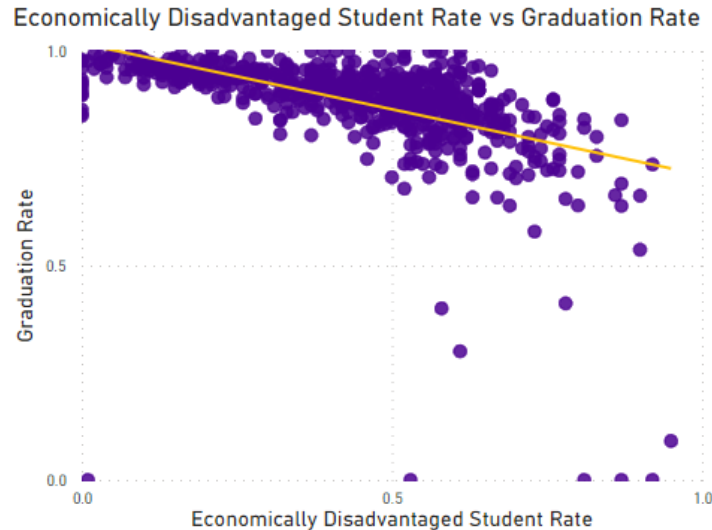
Racial Distribution for Districts with Highest Dropout Rate



What is the correlation between the number of economically disadvantaged students and graduation rate?

Findings

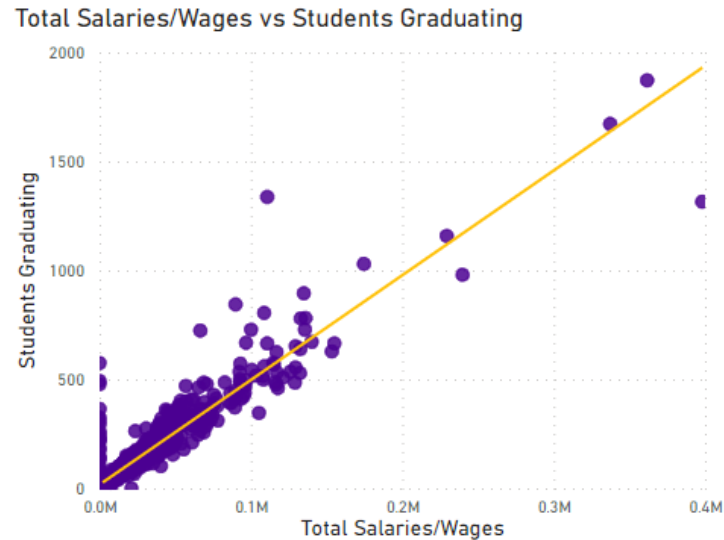
- Negative relationship
- The more economically disadvantaged students, the lower the graduation rate
- R-Squared value of = -0.622



Does an increase in total staff wages increase number of students graduating?

Findings

- Positive relationship
- The higher the wages are, the higher number of students graduate
- R-Squared Value = 0.949

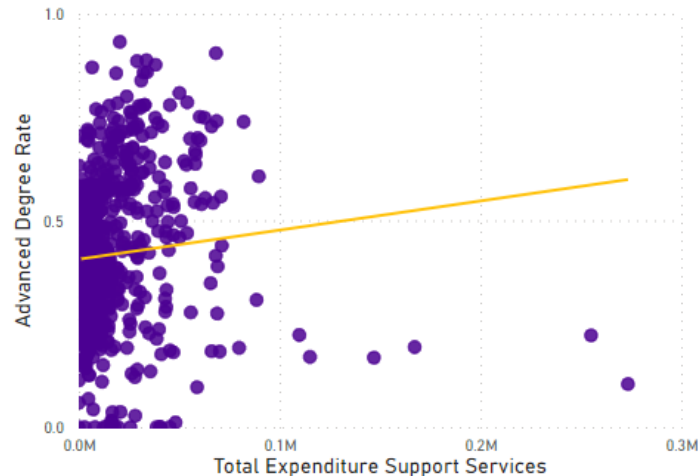


Does the amount of pupil support expenditures influence Advanced Regents diplomas acquired?

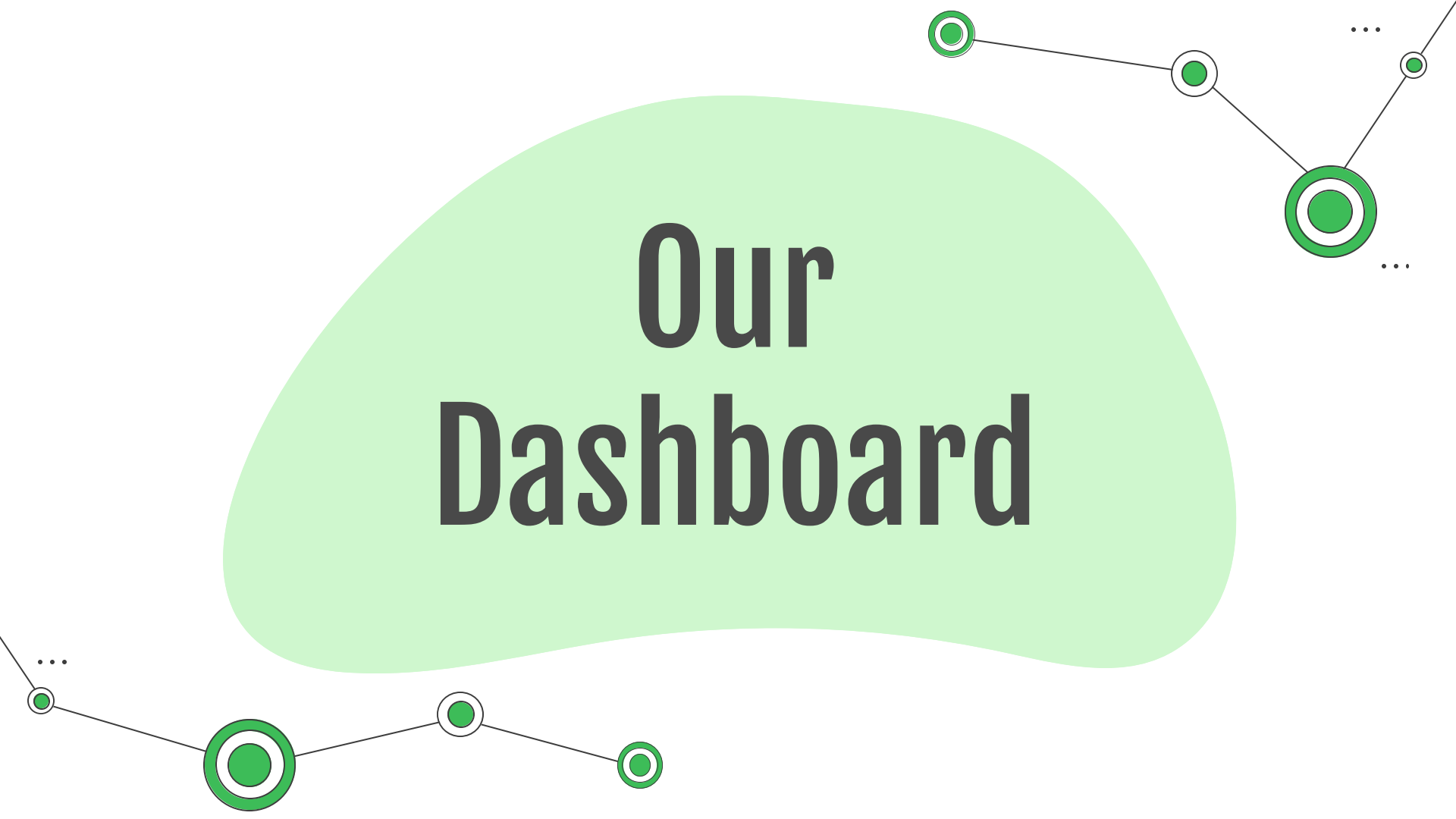
Findings

- Positive relationship
- The more spent on pupil support services, the more advanced degrees obtained

Total Expenditure Support Services vs Advanced Degree Rate



Our Dashboard



Dev10 Capstone: New York Schools' Financing and its Effects on Student Performance

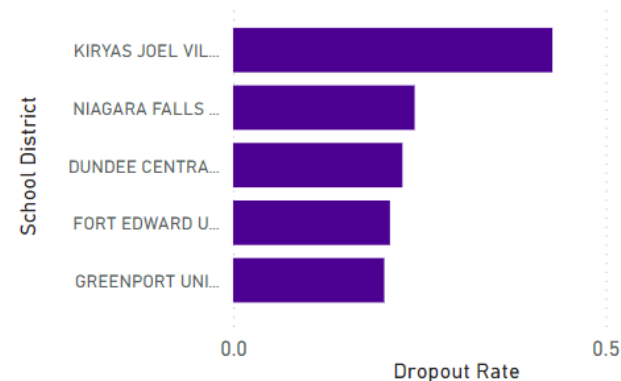
Datasets:

[2019 US Census Bureau School Finance Data](#),
[NYSED Data Site \(2019-20 Enrollment Database\)](#),
[NYSED Data Site \(2019-20 Graduation Rate Database\)](#)

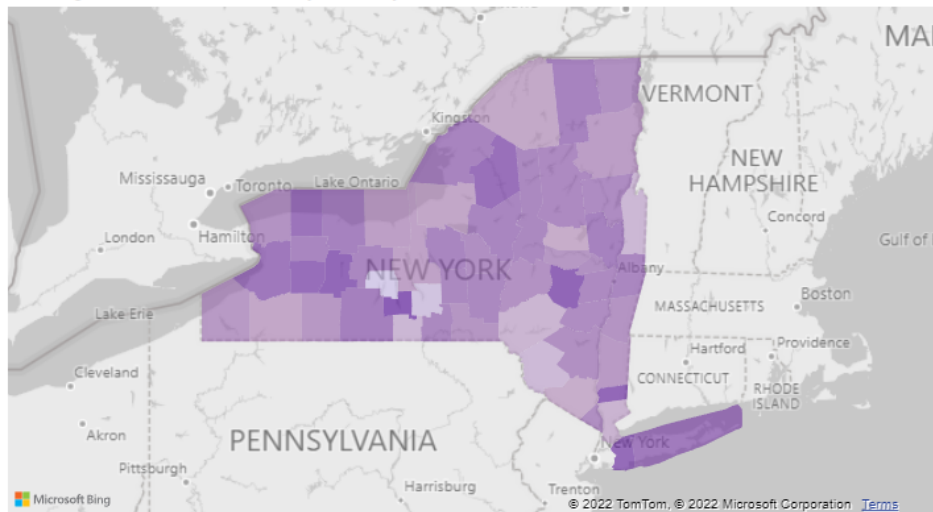
Members: Seung Min Kim, Azure Eller, Juliann Groglio, Widnie Dorilas

Description: This dashboard refers to data pulled from the New York State Education Department and the US Census Bureau regarding financial allocation, demographic, and performance data excluding NYC School districts. This report looked at graduation rates and enrollment demographics which include: the types of diplomas earned, homelessness rates, dropout rates, gender, and racial demographics. We will predict district dropout rates based on revenues, expenditures, racial breakdowns, and gender makeup for each district. Based on our research, we will give school districts a path towards decreasing dropout rates.

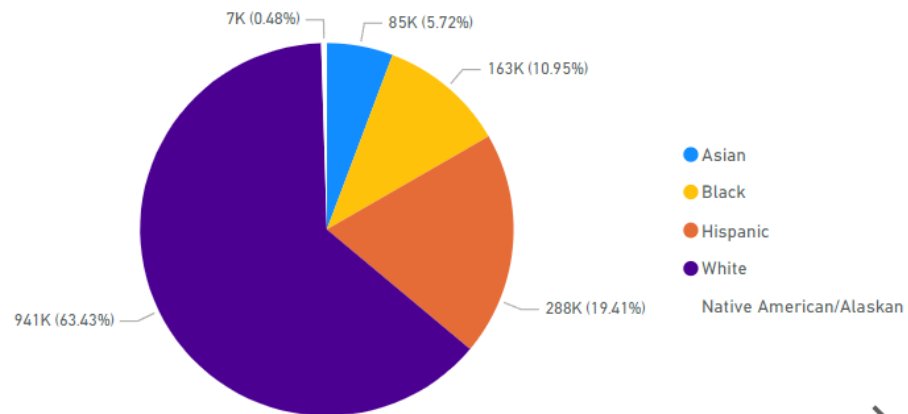
School Districts with Highest Dropout Rate



Average Graduation Rate by County



Racial Makeup of NYS School Districts



Finances

Description: Relationships between dropout rates, graduation rates, and Key financial indicators.

Top 5 School Districts Highest Revenue

ERIE
2859750
Total Revenue

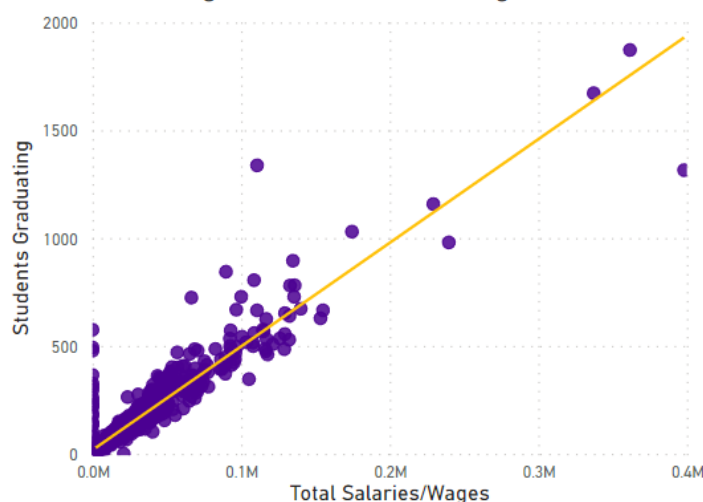
MONROE
2640003
Total Revenue

NASSAU
5525750
Total Revenue

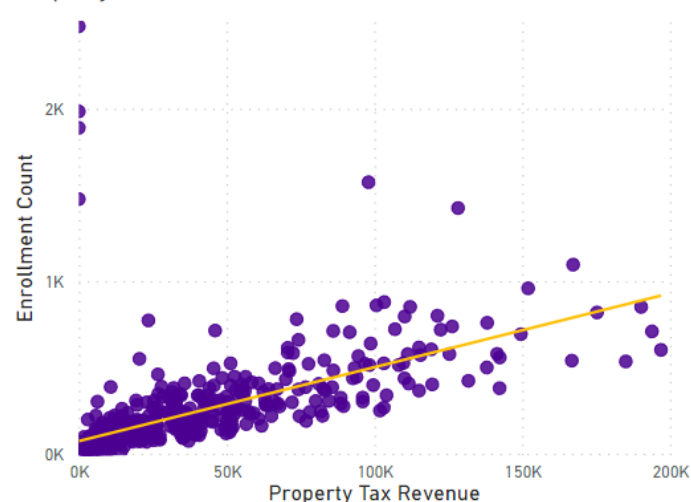
SUFFOLK
6489188
Total Revenue

WESTCHESTER
3776125
Total Revenue

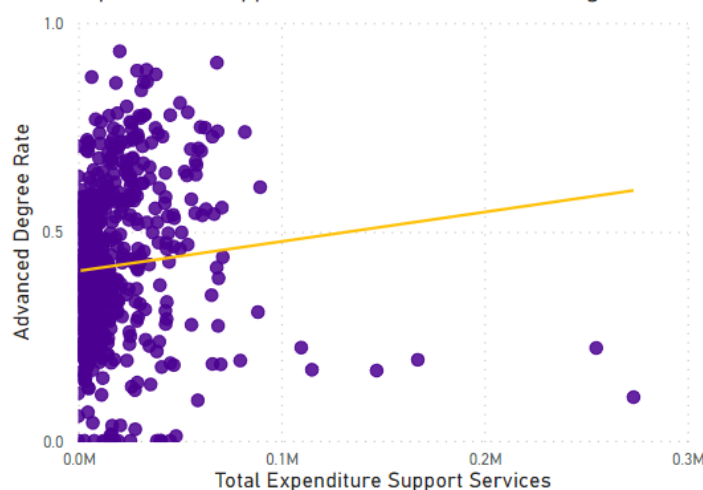
Total Salaries/Wages vs Students Graduating



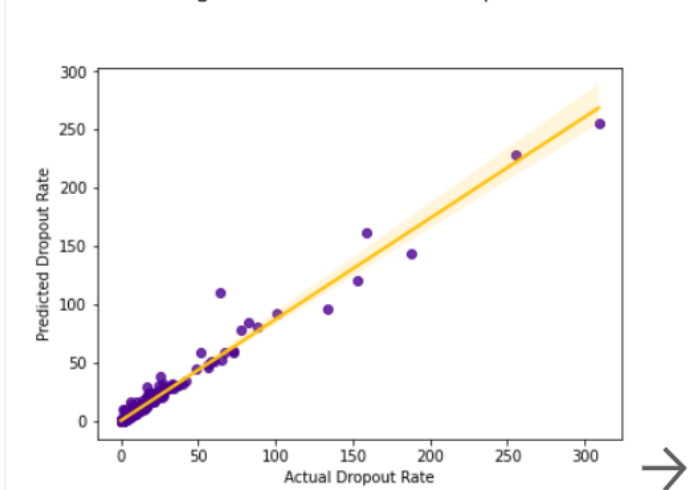
Property Tax Revenue vs Enrollment Count



Total Expenditure Support Services vs Advanced Degree Rate



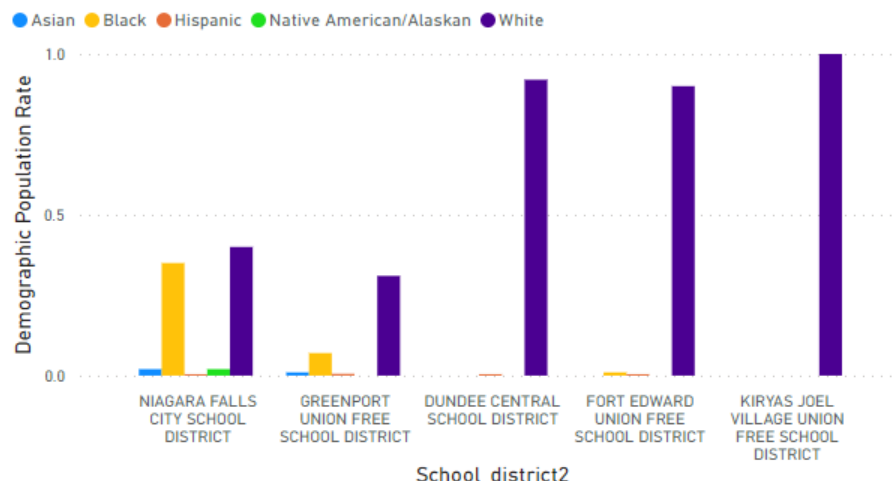
Machine Learning: Predicted vs Actual Dropout Count



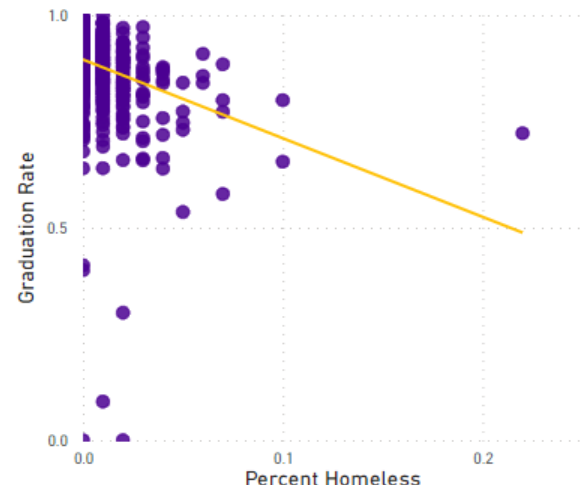
Demographics

Description: Relationships between dropout rates, graduation rates, and demographic data.

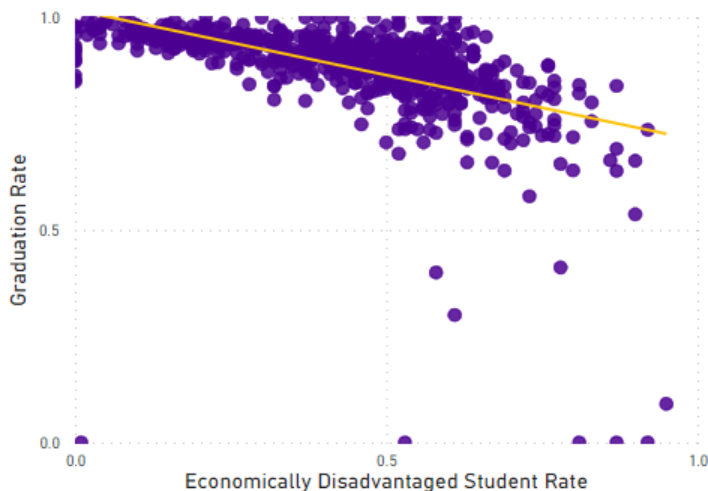
Racial Distribution for Districts with Highest Dropout Rate



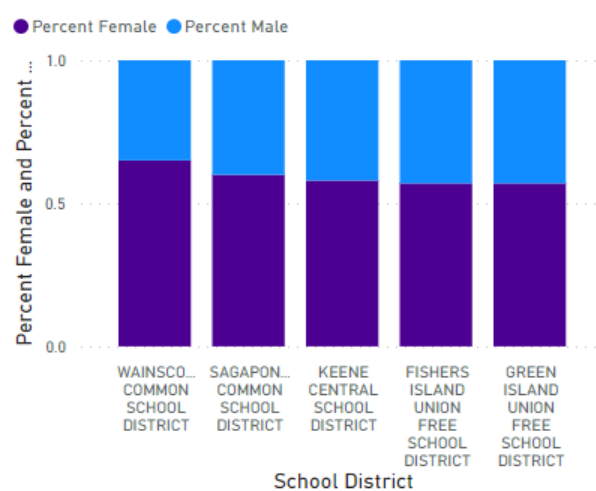
Percent Homeless vs Graduation Rate



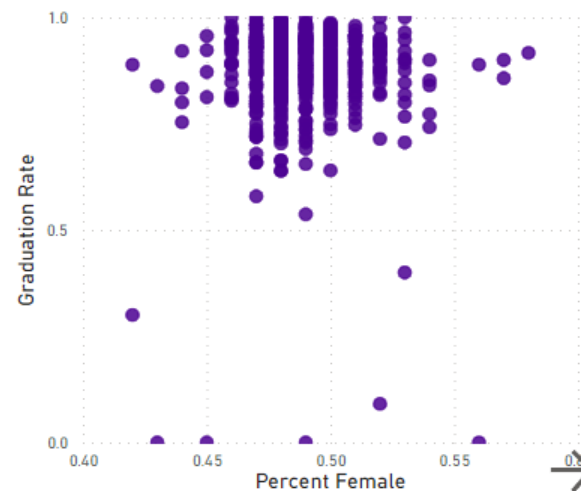
Economically Disadvantaged Student Rate vs Graduation Rate

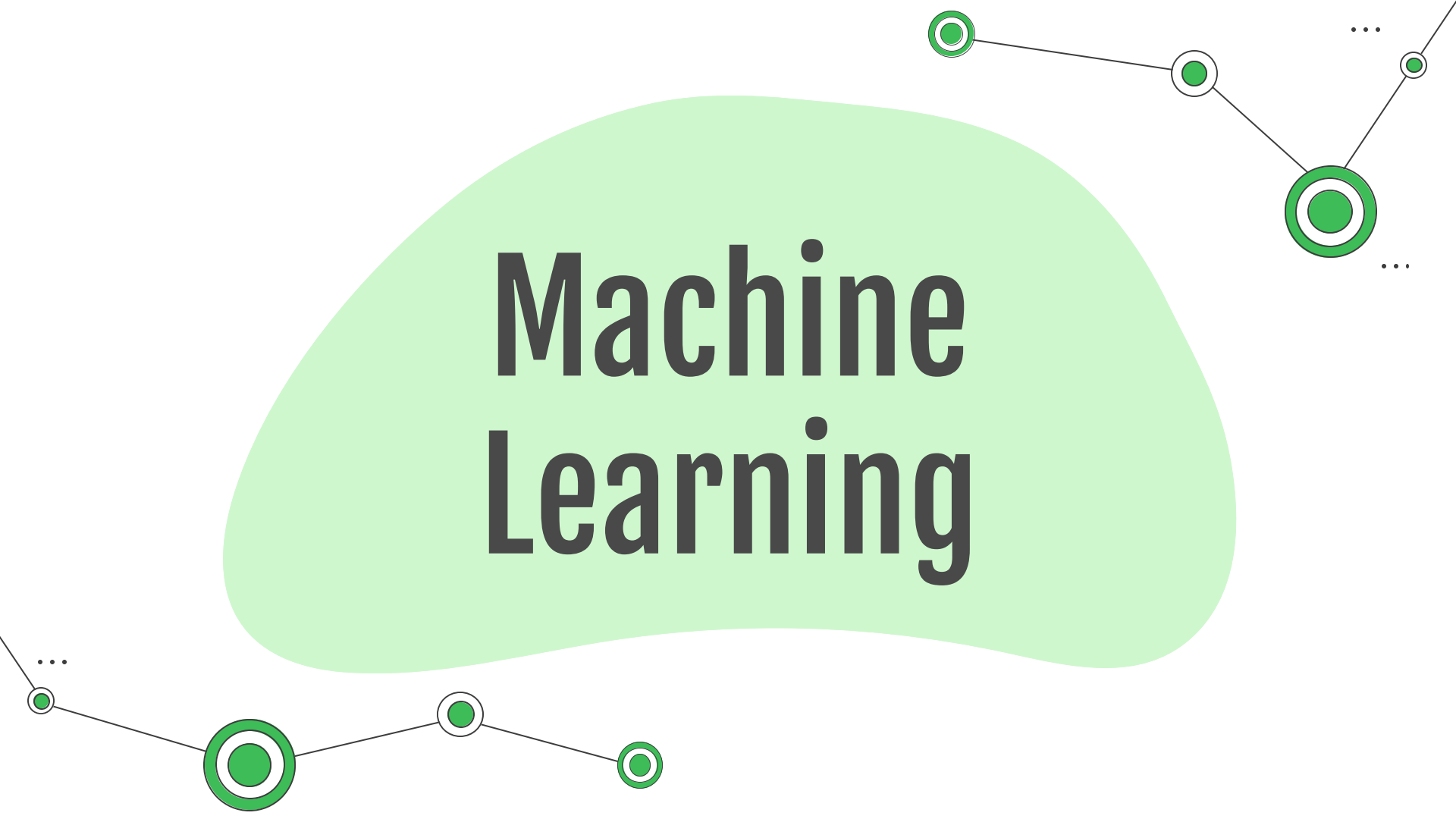


Percent Female and Percent Male by School District



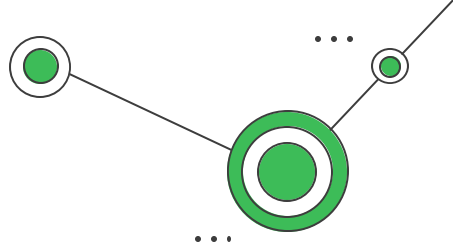
Percent Female vs Graduation Rate



The image features a central light green rounded rectangle containing the text "Machine Learning". Surrounding this central element are several green circles of varying sizes, some of which are connected by thin black lines. These circles are arranged in a network-like pattern, with some having concentric rings. Ellipses (...) are used to indicate that the network continues beyond the visible elements. The overall design is clean and modern, using a green and grey color palette.

Machine Learning

Machine Learning

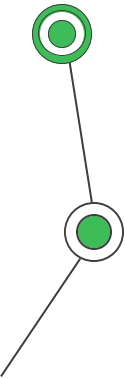


What we want to predict:

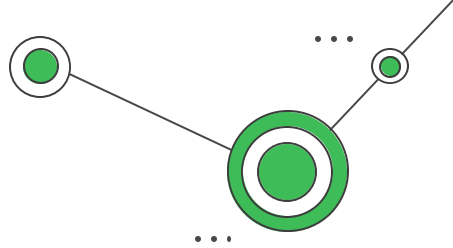
School district dropout rates based on financial allocations and demographic factors

How we will do it:

Using the Random Forest Machine Learning Method, we will be able to create a model and tune it to its best capability for predicting dropout rates based on the data we have



Tuning



- Performed a grid search on the Random Forest hyperparameters: bootstrap, max_depth, min_samples_split, min_samples, max_features, and n_estimators
- Found that max_depth= 10, max_features= sqrt, min_samples_split= 2, min_samples_leaf= 1 and n_estimators= 90 were the best adjustments for a best-performing model

max_depth

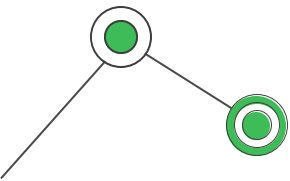
10

max_features

sqrt

n_estimators

90

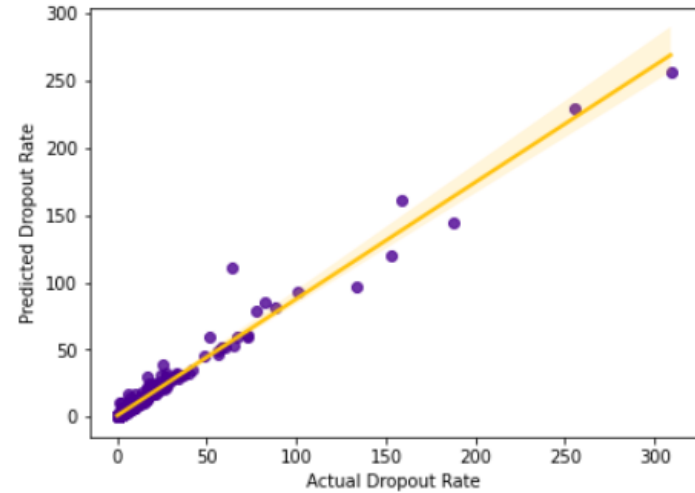


Results

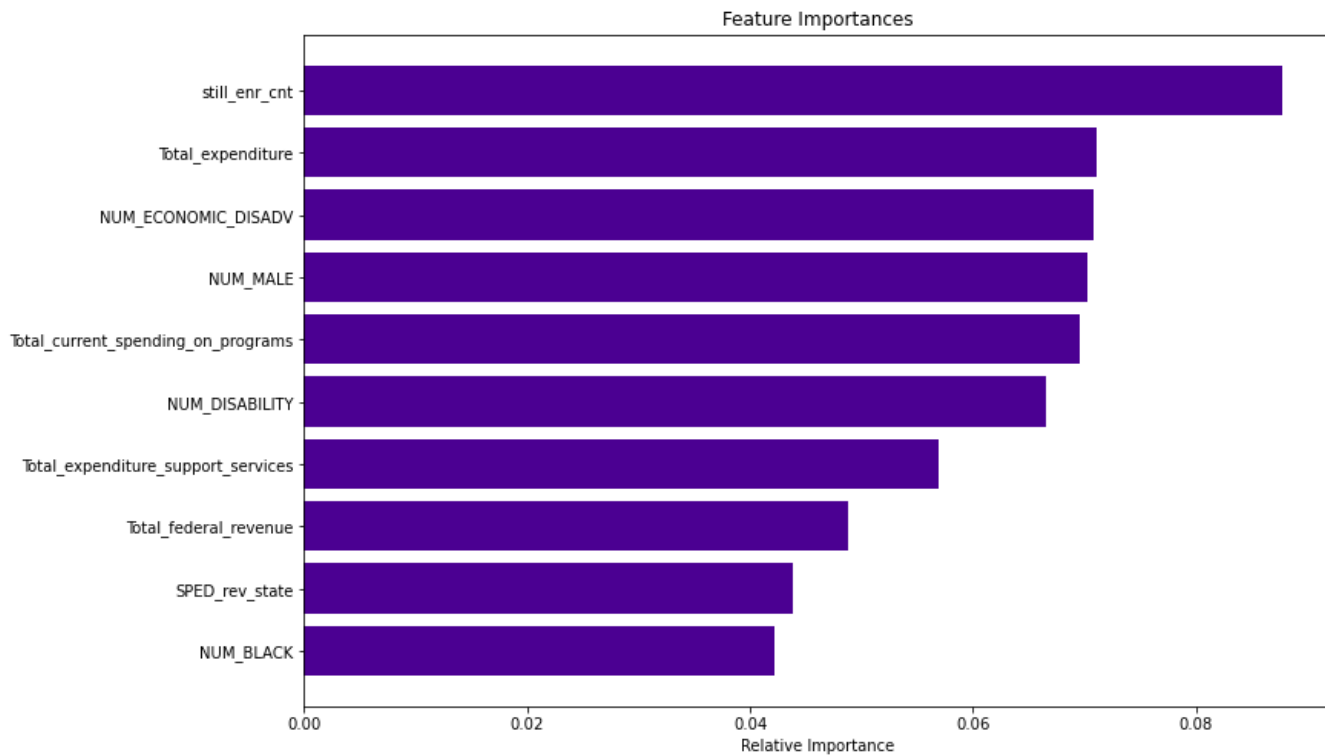
After tuning our model, we were able to improve our Coefficient of Determination of Testing from...

0.7139 to **0.9628**

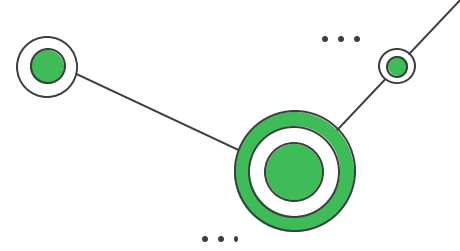
Predicted vs Actual Dropout Count



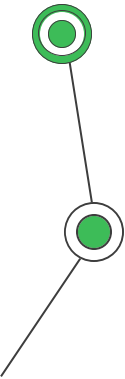
Features of Importance



Our Recommendations



1. Help students and families find affordable housing
2. More research into how students are economically disadvantaged, and how to best help those students
3. Have an equal male to female ratio of students
4. Spend a greater percentage of budget on staff wages
5. Increase pupil support expenditures



Thanks!

Do you have any questions?

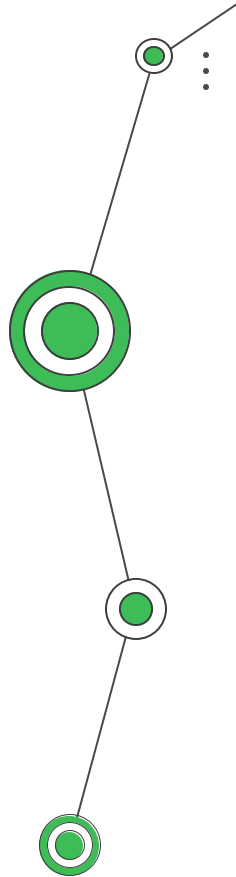
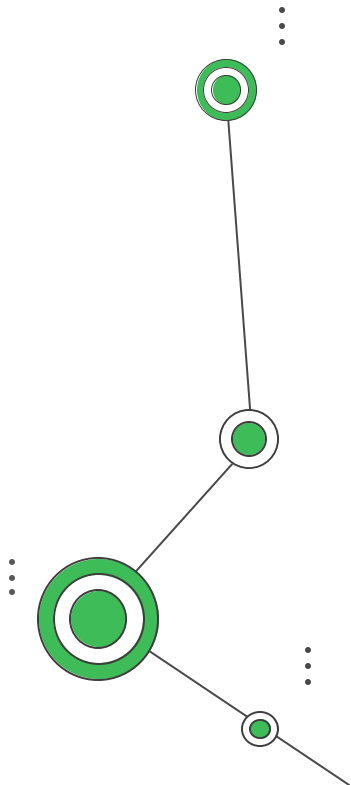


jgroglio@dev-10.com

wdorilas@dev-10.com

aeller@dev-10.com

s-kim@dev-10.com



Resources

- [2019 Education Finance Census Data \('All Data Items' CSV\)](#)
- [NYSED Data Site \(2019-20 Enrollment Database\)](#)
- [NYSED Data Site \(2019-20 Graduation Rate Database\)](#)

