# NEW YORK STATE PUBLIC SCHOOLS GRADUATION RATES: PREDICTING SUCCESS

Stefano Bastianelli, Brooke Brehm, Derek Caramella, Miguel Novo Villar, & Kenzie Potter

# AGENDA

- Introduce Vision & Goals
- Exhibit Data Model
- Display Data Visualizations
- Discuss Data Preprocessing
- Describe Model Performance
- Propose Future Enhancements

# INTRODUCTION

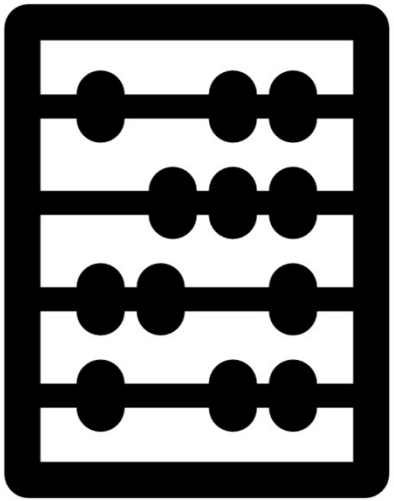
- Qualified teachers directly impact educational outcomes [2]
- Teacher experience, income, & standardized tests shaped the educational landscape & affected learning conditions
- Identifying patterns can help policy makers update requirements and allocate resources more efficiently

# GOALS

Is there a change in marginal graduate rate gain from incremental government funding?

Does the quality of teachers & principals impact graduation rate?

How can educational resources be better allocated?



# DATA MODEL


GraduationRate
Id int
Institution_ID int
Entity_ID int
Subgroup_ID int
Cohort_ID int
Grad_Rate int
InexperiencedTeachersPrincipals
Id int
Institution_ID int
Entity_ID int
Year datetime
NumTeach int
NumTeachInexp int
NumPrinc int
NumPrincInexp int
ExpendituresPerPupil
Id int
Institution_ID int
Entity_ID int
Year datetime
PupilCountTot int
PerFederalExp int
PerStateLocalExp int
PerFedStateLocalExp int
AccountabilityStatus
Id int
Institution_ID int
Entity_ID int
Year datetime
OverallStatus_ID int
InstutionGrouping
Id int
Institution_ID int
Entity_ID int
Group_ID int
Lkp_Entity
ID int
Code varchar(10)
Description varchar(300)
TeachersOutofCertification
Id int
Institution_ID int
Entity_ID int
Year datetime
NumTeachOc int
NumTeachOutCert int
Lkp_Cohort
ID int
Code varchar(10)
Description varchar(300)
Lkp_Institution
ID int
Code varchar(10)
Description varchar(300)
Lkp_OverallStatus
ID int
Code varchar(10)
Description varchar(300)
Lkp_Subgroup
ID int
Code varchar(10)
Description varchar(300)
Lkp_Group
ID int
Code varchar(10)
Description varchar(300)
dbdiagram.io

# DATA MODEL

## FACT TABLES

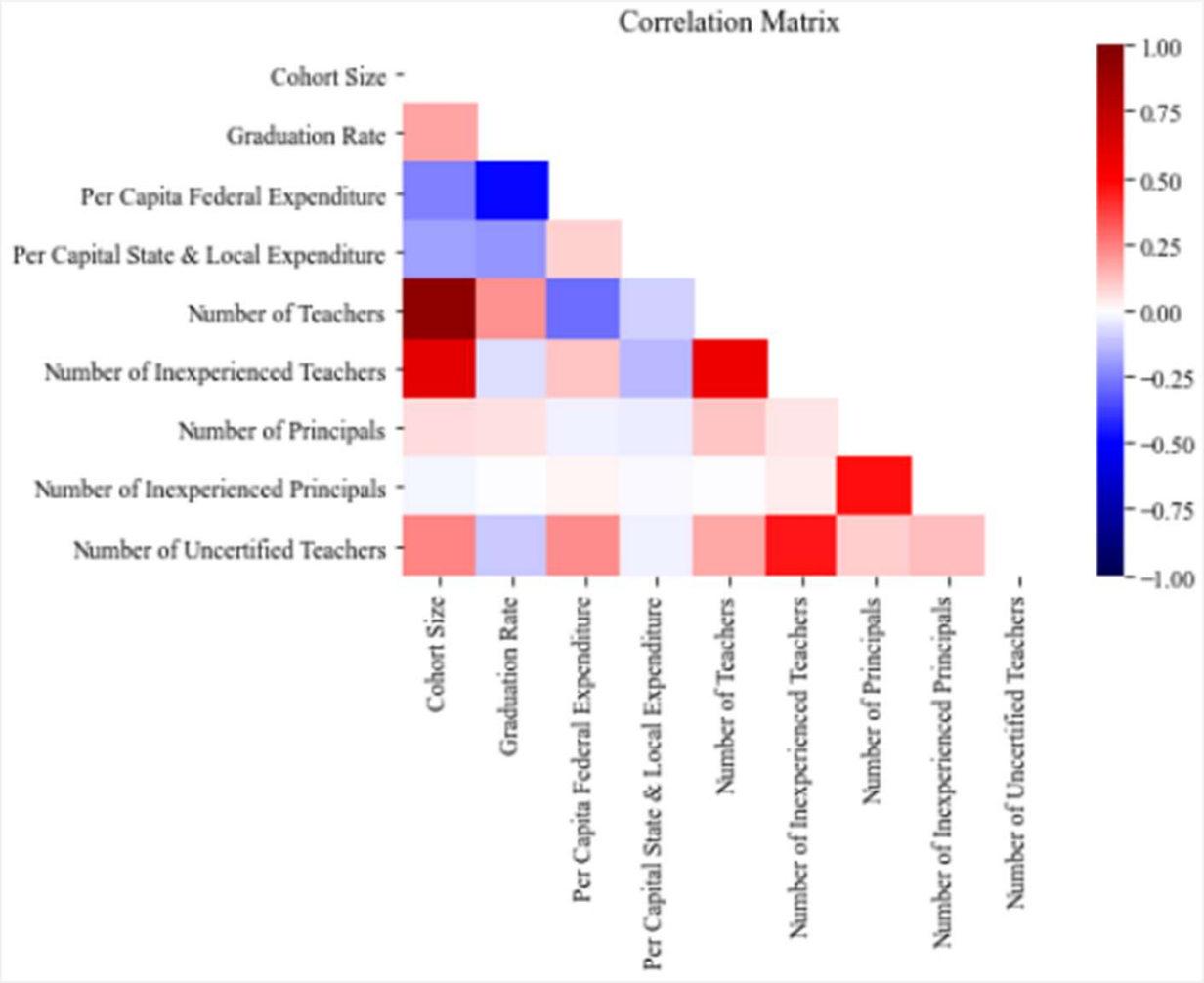
- Graduation Rate
- Inexperienced Teachers and Principal

- Expenditures per Pupil
- Accountability Status
- Institution Grouping
- Teachers Teaching Out of Certification

## DIMENSION TABLES

- Lkp_Cohort
- Lkp_Entity
- Lkp_Instituion
- Lkp_Group
- Lkp_Subgroup
- Lkp_OverallStatus

# CORRELATION MATRIX




## POSITIVE CORRELATIONS

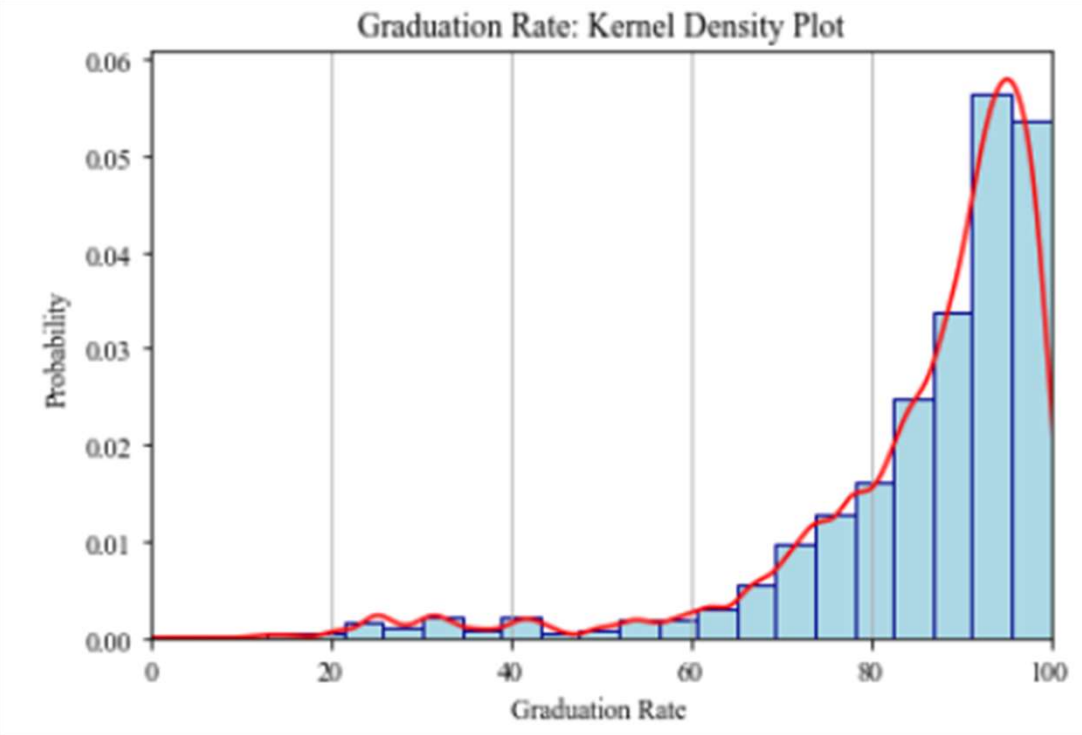
- Number of teachers & cohort size
- Number of inexperienced teachers & cohort size
- Number of inexperienced teachers & number of teachers
- Number of inexperienced teachers & Number of uncertified teachers
- Number of inexperienced principals & number of principals

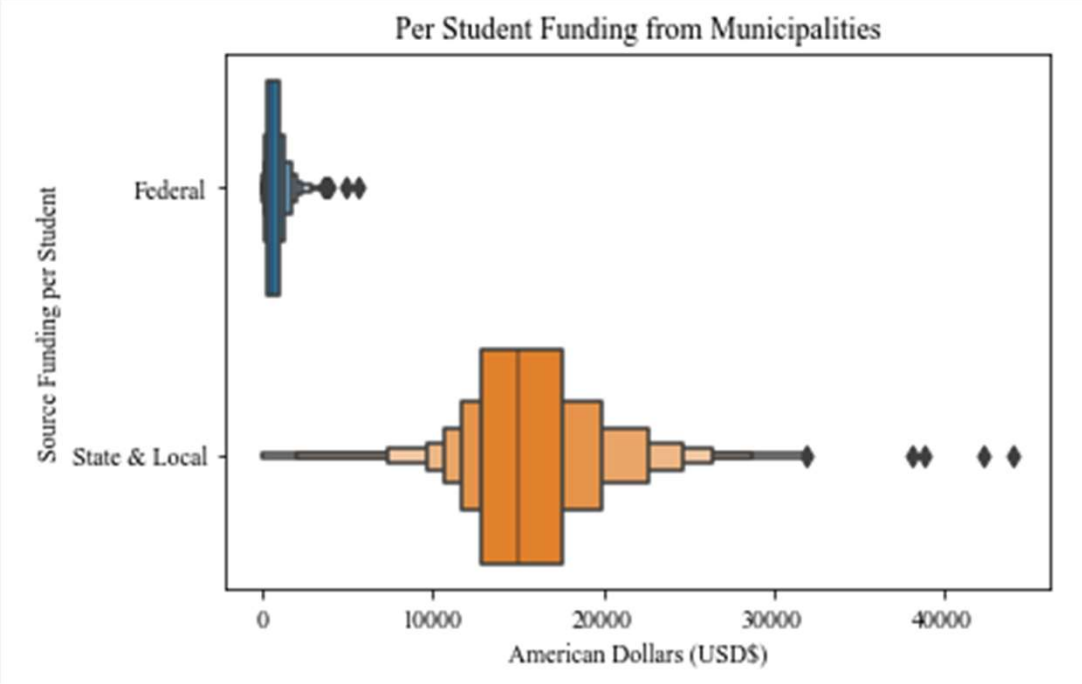
## NEGATIVE CORRELATIONS

- Per capita federal expenditure & cohort size
- Per capita federal expenditure & Number of teachers

# GRADUATION RATE DISTRIBUTION

- Heavily left (negatively) skewed
- Mean < Median < Mode
- Mode at roughly 95-100% graduation rate



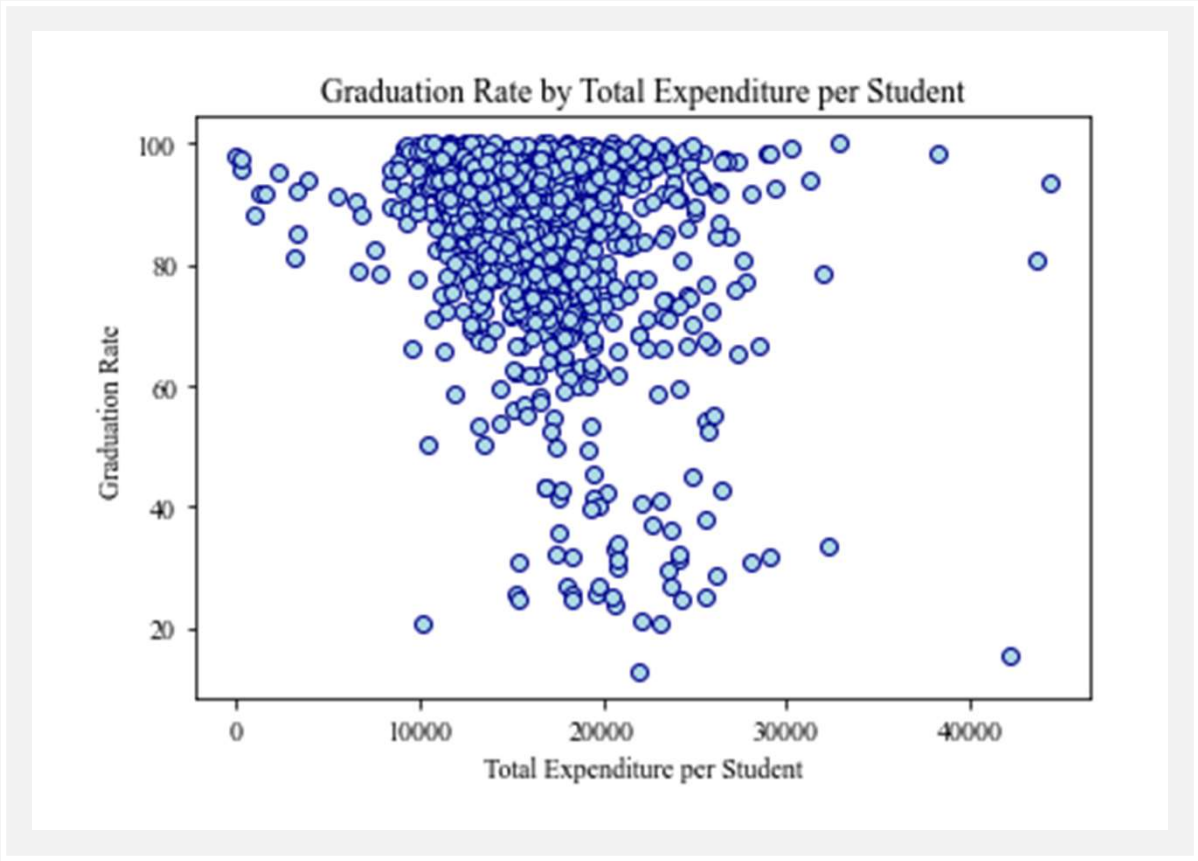



# PER STUDENT FUNDING FROM MUNICIPALITIES

- State & Local municipalities dominate funding for schools
- Average State & Local funding between $10,000 to $20,000 per student
- Federal funding is less than $10,000 per capita

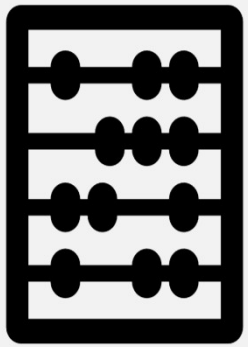
# GRADUATION RATE BY FUNDING

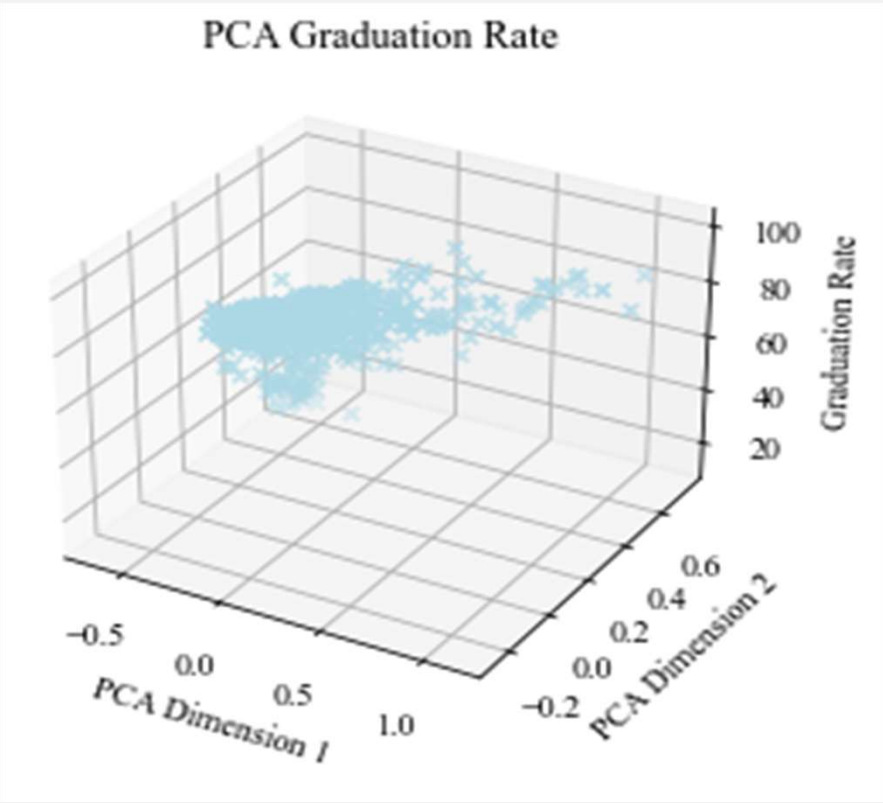
- Most budgets $10,000-$30,000 per student
- $0-$10,000 has a graduation rate between roughly 80 & 100 percent
- Most densely grouped at 60 to 100 percent graduation rate
- $30,000-$40,000 per student does not seem to drastically improve graduation rates
- Potential low graduation rate outliers



# PCA OF GRADUATION RATE

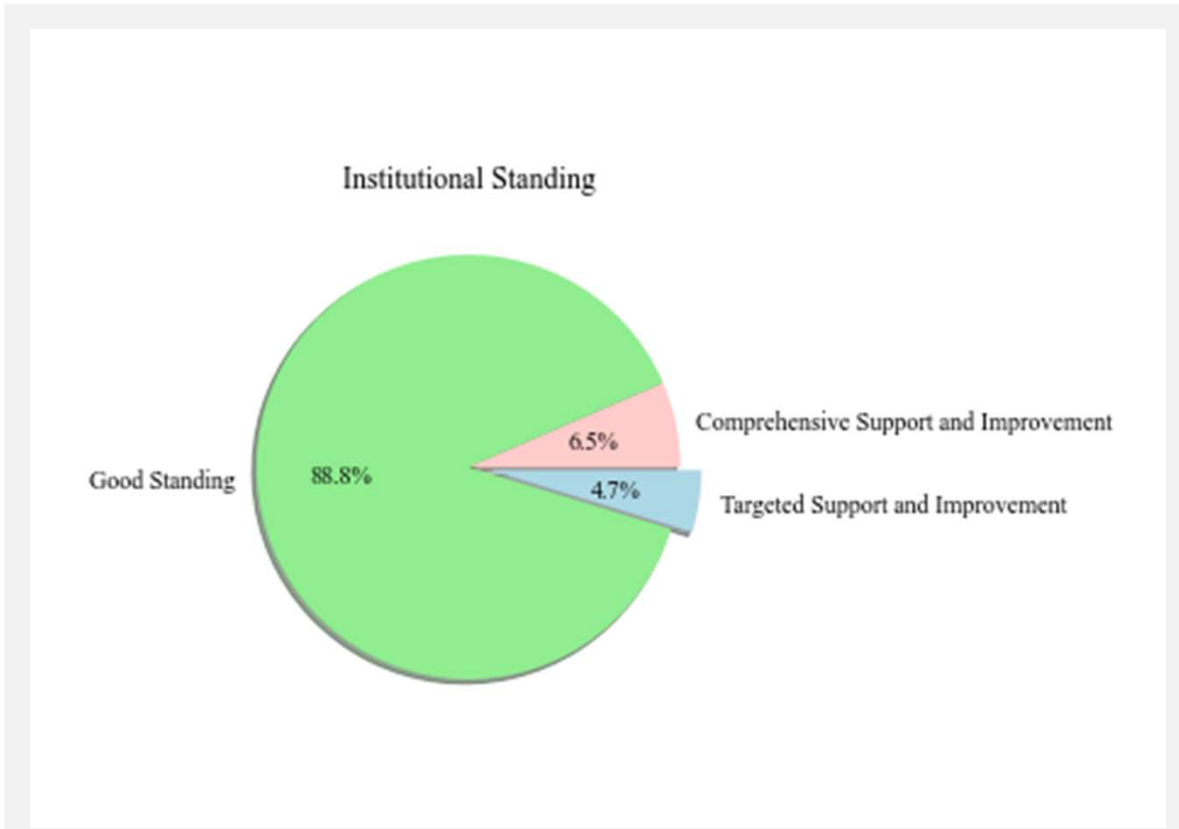
- Graduation rates condensed around 80%
- Explained variance: 55.5%

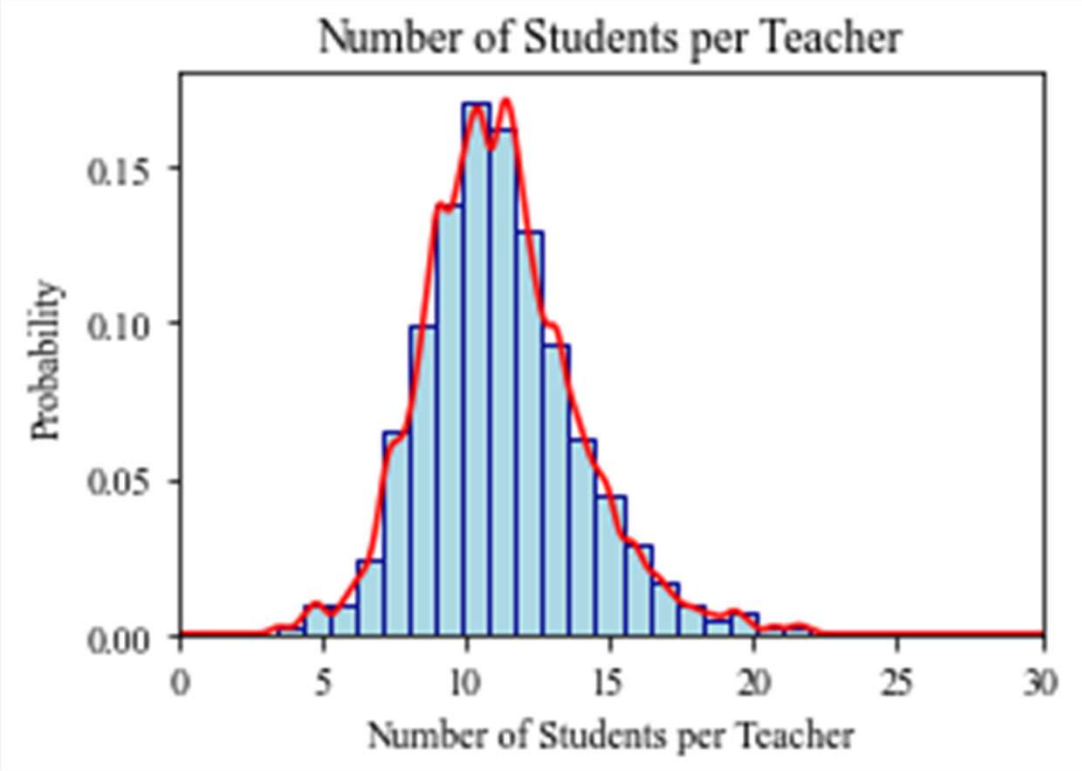




# INSTITUTIONAL STANDING IMBALANCE

- Schools primarily in good standing
- No closing schools
- ~11% currently receiving support & improvement



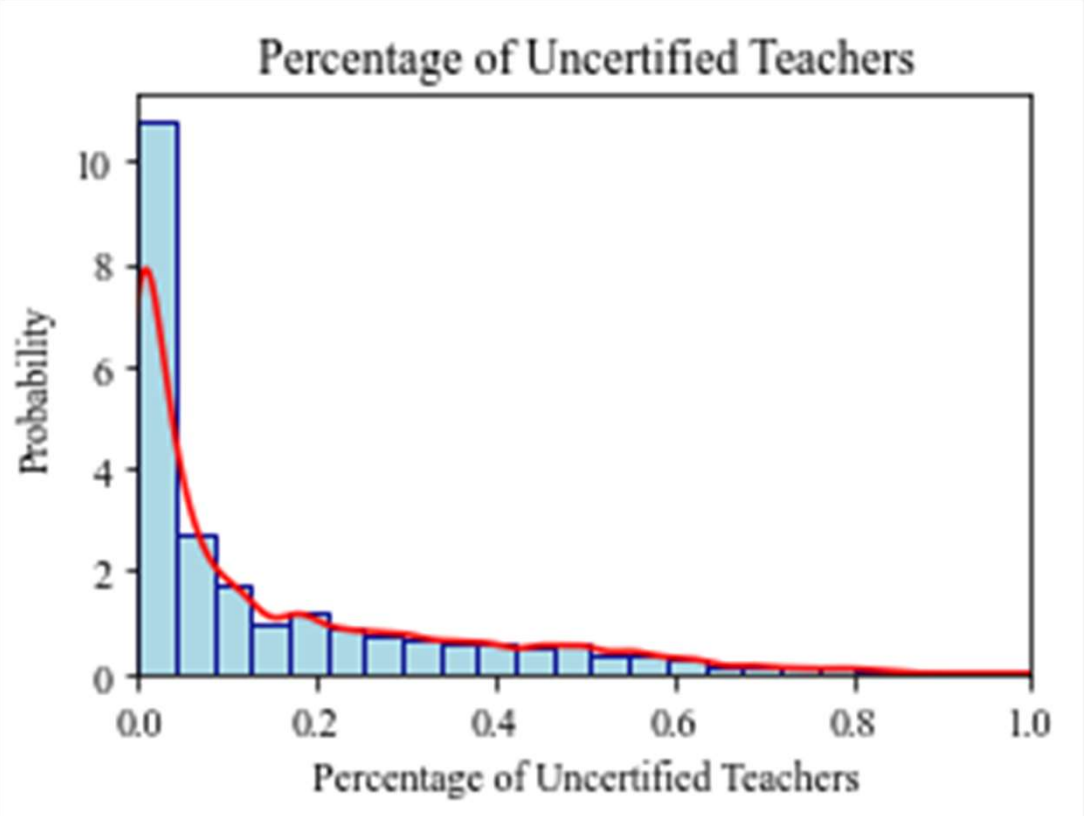



# STUDENTS PER TEACHER

- **Mode**
  - 10 students per teacher
- **Minimum**
  - 3 students per teacher
- **Maximum**
  - 22 students per teacher

# UNCERTIFIED TEACHERS COMPOSITION

- Mode: No uncertified teachers
- More likely to have 0%-20% of teachers in a school teaching outside of their certification
- Few schools have 60% or more teachers outside of certification



# MODEL PIPELINE: FEATURE ENGINEERING

| Feature | Function |
| --- | --- |
| Teachers per Pupil | $\frac{\text{Number of Teachers}}{\text{Cohort Size}}$ |
| Inexperienced Teachers per Pupil | $\frac{\text{Number of Inexperienced Teachers}}{\text{Cohort Size}}$ |
| Inexperienced Teachers per Teacher | $\frac{\text{Number of Inexperienced Teachers}}{\text{Number of Teachers}}$ |
| Principals Per Pupil | $\frac{\text{Number of Principals}}{\text{Cohort Size}}$ |
| Inexperienced Principals per Pupil | $\frac{\text{Number of Inexperienced Principals}}{\text{Cohort Size}}$ |
| Inexperienced Principals per Principal | $\frac{\text{Number of Inexperienced Principals}}{\text{Number of Principals}}$ |

# MODEL PIPELINE: FEATURE PREPROCESSING

## ONE HOT ENCODING

- Overall Status
  - Good Standing
  - Potential Target District
  - Potential Target Support & Improvement
  - Target Support & Improvement
  - Potential Comprehensive Support & Improvement (CSI)
  - Comprehensive Support & Improvement
  - Closing School

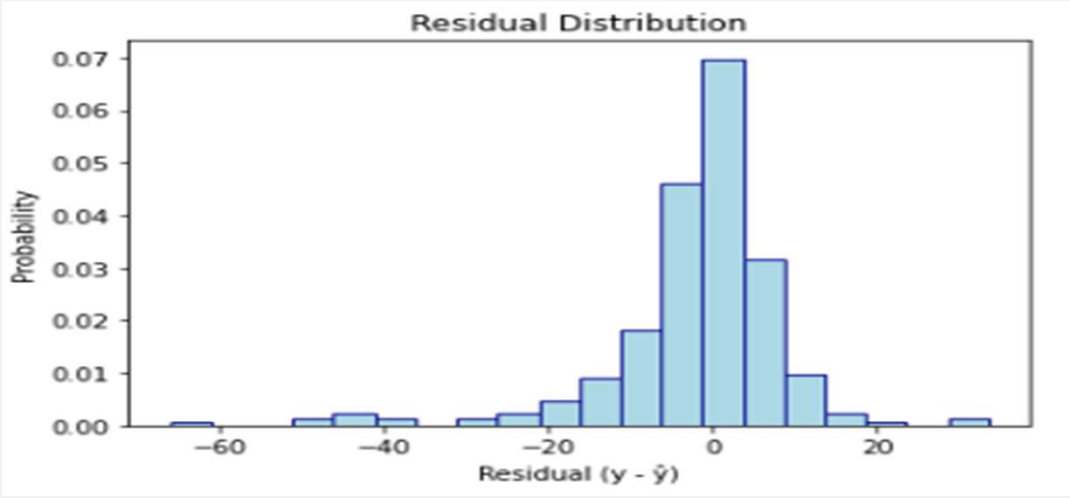
## MIN MAX SCALER

- Cohort Size
- Federal Expenditure per Student
- State & Local Expenditure per Student
- Teachers per Pupil
- Inexperienced Teachers per Pupil
- Inexperienced Teachers per Teacher
- Principals per Pupil
- Inexperienced Principals per Pupil
- Inexperienced Principals per Principal

# MODEL PERFORMANCE

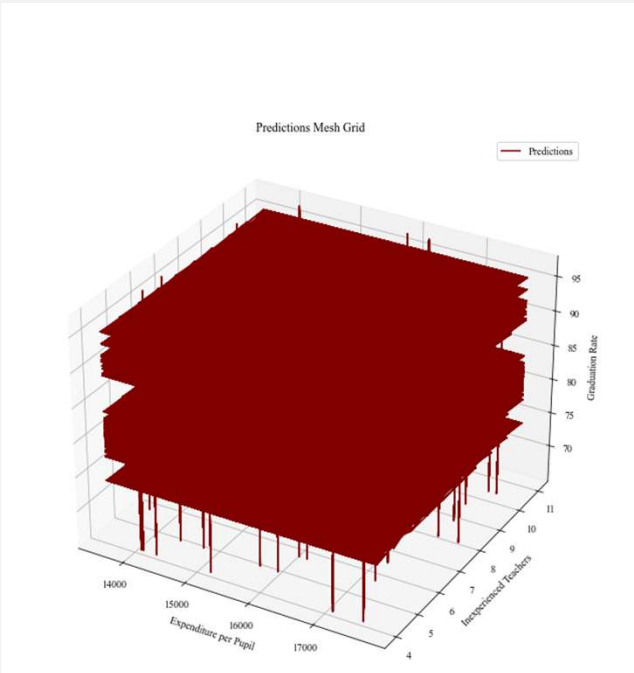
| Metric | Elastic Net | SVR | Bayesian Ridge | AdaBoost Regressor | Random Forest Regressior | Neural Network |
|---|---|---|---|---|---|---|
| Maximum Error | 48.865 | 66.081 | 45.522 | 65.827 | 58.585 | 60.66 |
| Mean Absolute Error | 7.422 | 6.571 | 7.383 | 7.582 | 6.41 | 7.393 |
| Mean Squared Error | 119.415 | 115.972 | 118.08 | 118.254 | 94.777 | 126.727 |
| Median Squared Error | 5.147 | 3.863 | 5.072 | 5.971 | 4.12 | 4.619 |
| *Models Subject to Exhaustive Grid Search* | | | | | | |

- Left (negatively) skewed
- SVR mispredicts underperforming schools (graduation rate < 50%)



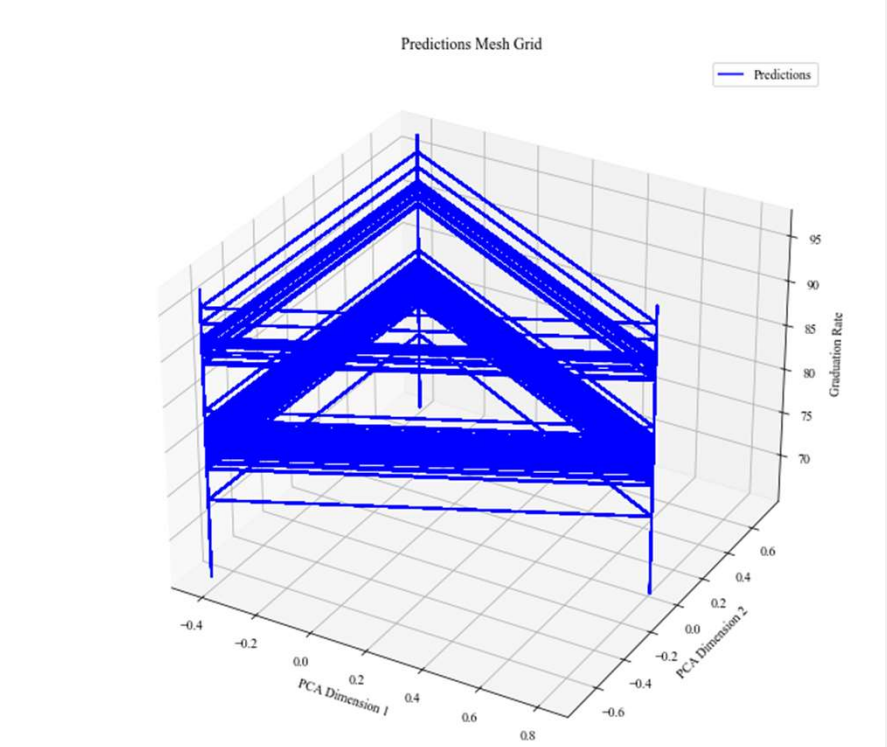
# MESH GRID

- Sample Size: 20,000
- Samples span feature space domain
- Axes:
  - *X*: Expenditure per Pupil
  - *Y: Inexperienced Teachers*
  - *Z*: Graduation Rate
- Analysis
  - Distinct stacked hyperplanes
  - No clear relationship exists



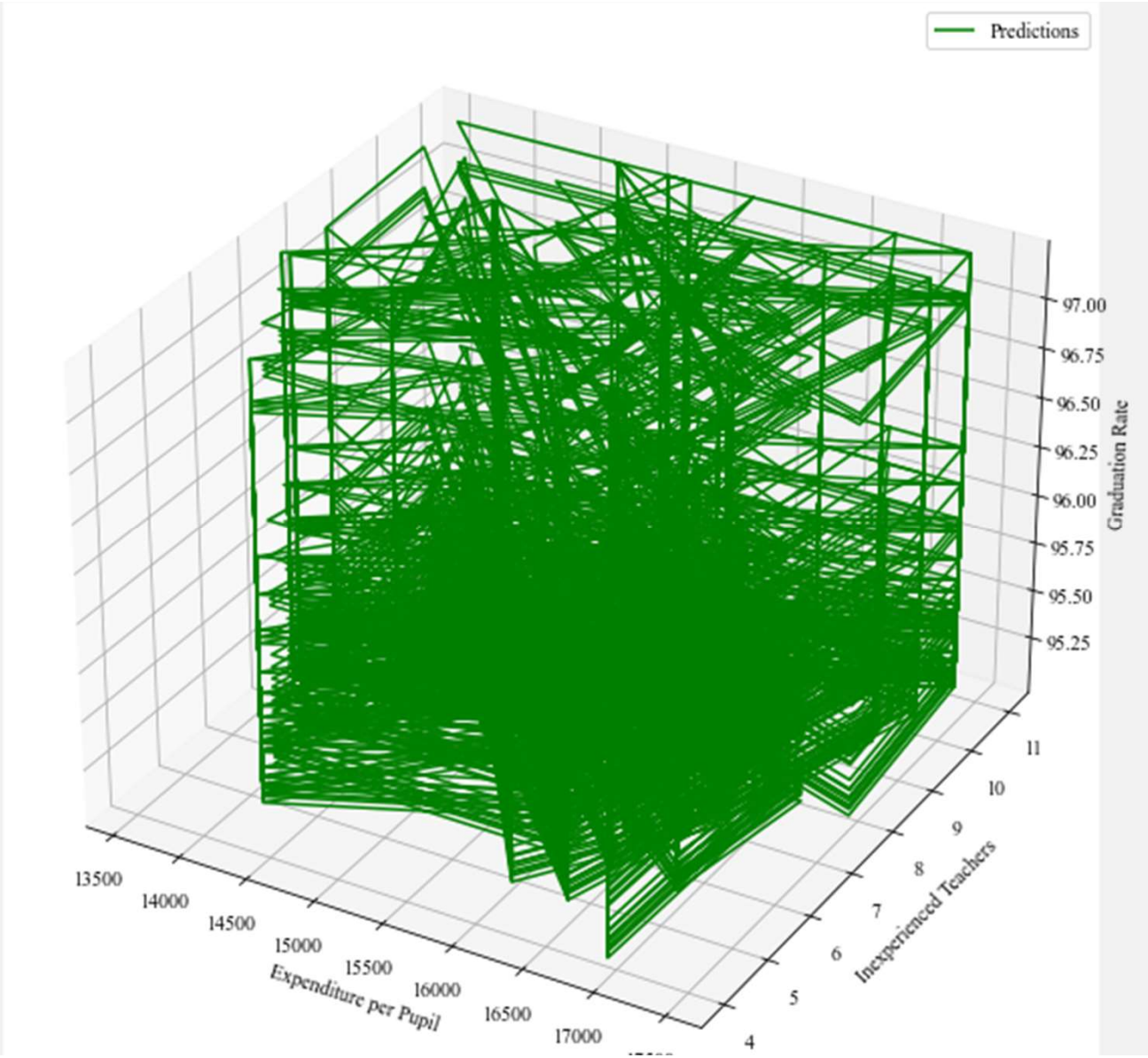
# PCA MESH GRID

- Triangular, stacked structure
- Explained Variance: 93.5%
- Outer Pillars
  - Unobserved Features
    - District employment rates
    - District GDP
    - District crime rate
- Optimal points exist



# CROSS SECTION MESH GRID

- Mesh Grid Subsection
  - 95.25 to 100 Percent
- No clear relationship exists between Expenditure per Pupil & Inexperienced Teacher

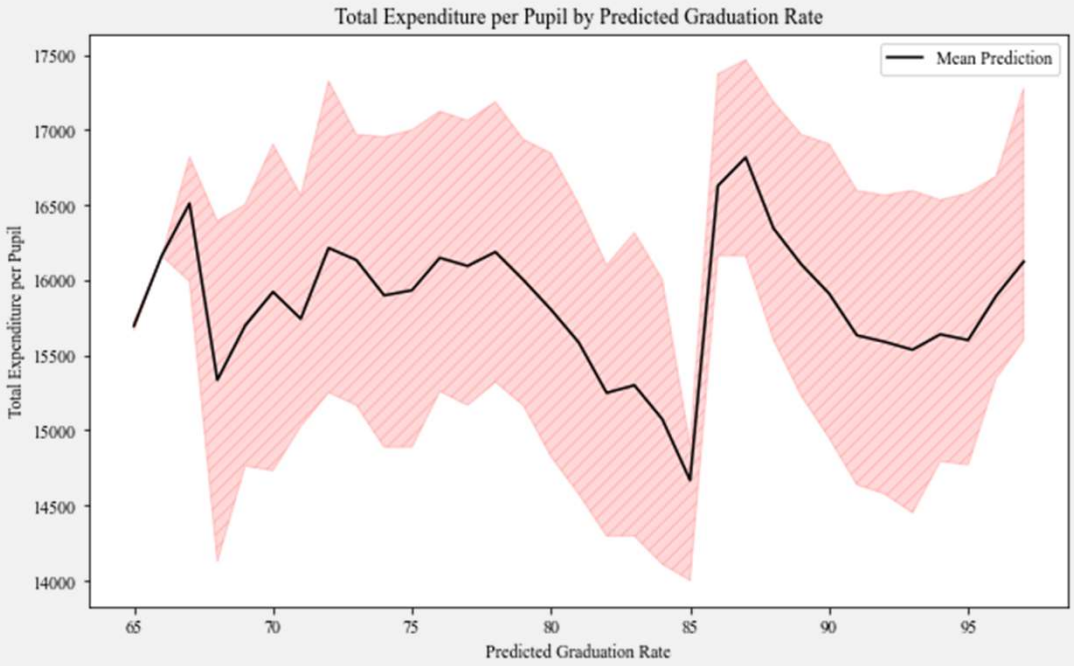


# OPTIMAL COMBINATIONS

| Total Funding per Student | Teachers per Pupil | Teacher per Inexperienced Teacher | Number of Principals | Overall Status | Predicted Graduation Rate |
|---|---|---|---|---|---|
| $13,580 | 5 | 6 | 1 | Good Standing | 96.5 |
| $15,600 | 5 | 8 | 1 | Good Standing | 97.2 |
| $16,810 | 5 | 10 | 1 | Good Standing | 97.0 |
| $17,280 | 5 | 7 | 1 | Good Standing | 97.0 |
| $17,340 | 5 | 6 | 1 | Good Standing | 97.0 |

# TOTAL EXPENDITURE PER PUPIL BY GRADUATION RATE

- *X* Axis
  - Decile Predicted Graduation Rate
- *Y* Axis
  - Total Expenditure per Pupil
- Black Line
  - Mean Total Expenditure per Pupil
- Red Shading
  - Total Expenditure per Pupil IQR
- Analysis
  - Total Expenditure per Pupil **does not** exhibit diminish returns



# RESULTS

| Is there a change in marginal graduate rate gain from incremental government funding? | Does the quality of teachers & principals impact graduation rate? | How can educational resources be better allocated? |
| --- | --- | --- |
| No | Yes | Quality teachers with adequate funding |

# AN EYE TO THE FUTURE

- Use results to inform schools on the impacts of funding, teacher experience, & certification on overall graduation rates
- Continue with new additional data
- Expand analysis to other states
- Compare & contrast different states

# REFERENCES


1. "2020: NY State - Report Card: NYSED Data Site," data.nysed.gov. [Online]. Available: https://data.nysed.gov/
2. D. Boyd, H. Lankford, S. Loeb, J. Rockoff, and J. Wyckoff, "The narrowing gap in New York City teacher qualifications and its implications for student achievement in high-poverty schools," Journal of Policy Analysis and Management, vol. 27, no. 4, pp. 793–818, 2008.
3. NYSED Knowledge, Skill, Opportunity. New York State Education Department [Online]. Available: http://www.nysed.gov/.
4. NYS Schools Logo. New York State Education Department [Online]. Available: http://www.nysed.gov/.
5. "Teaching and Educational Leadership Standards," New York State Education Department. [Online]. Available: http://www.nysed.gov/.