New York City Education Data

Aaron Kohn Milestone 1

# NYC Education Data

## Data Sets

1. HTML data table of active schools in New York State. Including information regarding location of school, community type, and school type. [Public & Charter Schools: Active Public and Charter Schools with Grade Org and Community type - IBM Cognos Viewer (nysed.gov)](https://cognos.nysed.gov/ibmcognos/bi/v1/disp?b_action=cognosViewer&ui.tool=CognosViewer&ui.action=run&ui.object=%2Fcontent%2Ffolder%5B%40name%3D%27NYSED+Reports%27%5D%2Ffolder%5B%40name%3D%27SEDREF%27%5D%2Ffolder%5B%40name%3D%27SEDREF+Reports+for+Public+Website%27%5D%2Freport%5B%40name%3D%27Public+%26+Charter+Schools%3A+Active+Public+and+Charter+Schools+with+Grade+Org+and+Community+type%27%5D&m_passportID=101:9c25e60c-5b69-f3e9-8c02-f7d02fb354c1:3907078367&cv.header=false&cv.toolbar=false&run.outputFormat=HTML) https://cognos.nysed.gov/ibmcognos/bi/v1/disp?b\_action=cognosViewer&ui.tool=CognosViewer&ui.action=run&ui.object=%2Fcontent%2Ffolder%5B%40name%3D%27NYSED+Reports%27%5D%2Ffolder%5B%40name%3D%27SEDREF%27%5D%2Ffolder%5B%40name%3D%27SEDREF+Reports+for+Public+Website%27%5D%2Freport%5B%40name%3D%27Public+%26+Charter+Schools%3A+Active+Public+and+Charter+Schools+with+Grade+Org+and+Community+type%27%5D&m\_passportID=101:b6a4ecc7-c128-fd01-38e3-dc96dfd82342:3458594141&cv.header=false&cv.toolbar=false&run.outputFormat=HTML
2. API for demographics in NYC schools. Information on enrollment in each grade, race and gender. [2013 - 2018 Demographic Snapshot School | NYC Open Data (cityofnewyork.us)](https://data.cityofnewyork.us/Education/2013-2018-Demographic-Snapshot-School/s52a-8aq6)
3. Downloadable data to csv file. Quality review ratings for NYC schools. Provides access to year by year ratings for NYC schools. [2005 - 2020 Quality Review Ratings | NYC Open Data (cityofnewyork.us)](https://data.cityofnewyork.us/Education/2005-2020-Quality-Review-Ratings/3wfy-sn5g)

## Relationship Between Data

The relationship between the datasets seems to be obvious to link by school. The identification system appears different in each data set. Further study is required if identification can be confirmed. It is also feasible that relationship by school district should be used. If relationship by school is established than the relationship is one to one. In the event that a one to one relationship can not be confirmed, using a one to many relationship depending on what data is available.

## Work Required

The HTML data will provide information regarding school location and district. The data would need to be limited to include only instances of NYC schools. School names would need to be processed to match school identification used in other demographics.

For the API data a query will be set to extract the relevant data. School names are provided, which will hopefully be useful to establish relationships. The DBN column is unclear (even after reading the information regarding columns. The values don’t match the ones provided).

Quality review ratings has been downloaded in csv format. Column names provide no information regarding the data. Reading and understanding the accompanying information is required to understand the data. Renaming columns to reflect what each variable contains is necessary. Schools are identified by code for borough and school number. Translating the code to school names will be required in order to merge with other data sets. The rating system used also needs explanation or recoding to make it understandable.

Looking at the data one final step would be required in order to merge the data is identifying years. There are multiple rows per school for each school year. Merging the data would require matching years. Not all schools have data collected for all years. The API and CSV data sets both have multiple years of data. Data would need to be merged based on school and year.

Much learning and study would be required for me to successfully extract and combine the data. The steps outlined above are based on the knowledge as of now. It is in no way intended to be a complete overview.