

The state of open data on school bullying and harassment in NYC

Two Sigma Data Clinic

NYC Open Data Week
March 6th, 2018



Important legal information

The views expressed herein are not necessarily the views of Two Sigma Investments, LP or any of its affiliates (collectively, “Two Sigma”). The information presented herein is only for informational and educational purposes and is not an offer to sell or the solicitation of an offer to buy any securities or other instruments. Additionally, the information is not intended to provide, and should not be relied upon for investment, accounting, legal or tax advice. Two Sigma makes no representations, express or implied, regarding the accuracy or completeness of this information, and you accept all risks in relying on this information for any purpose whatsoever.

This presentation shall remain the property of Two Sigma and Two Sigma reserves the right to require the return of this presentation at any time. Copyright © 2018 TWO SIGMA INVESTMENTS, LP. All rights reserved.



TWO SIGMA
DATA CLINIC

est. 2014

pro-bono data and tech for social good

www.twosigma.com/dataclinic

#nyc #opendata #opendataweek
@twosigma



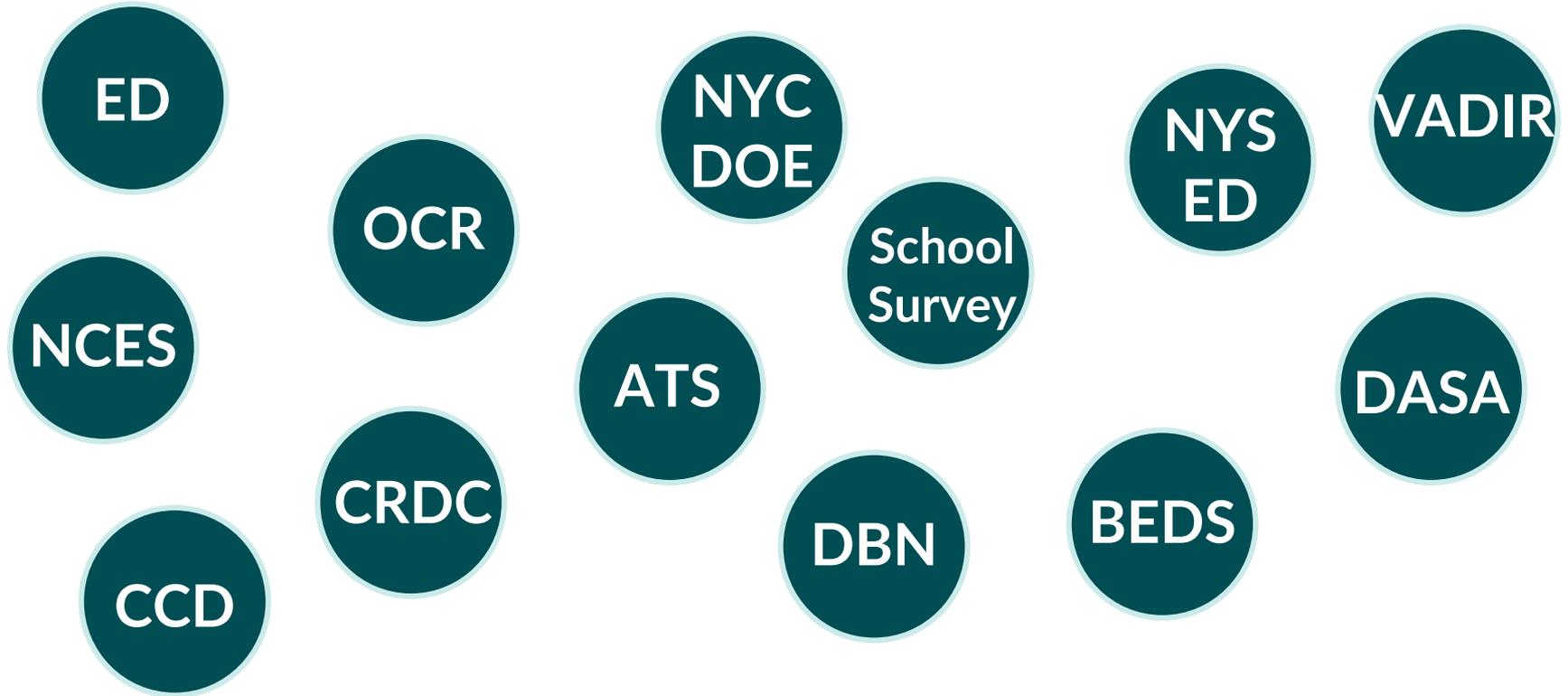
**Open Data
is an invitation for
anyone
anywhere
anytime
to engage
with New York City**

“Open Data for All” (2015)

<http://www1.nyc.gov/assets/home/downloads/pdf/reports/2015/NYC-Open-Data-Plan-2015.pdf>

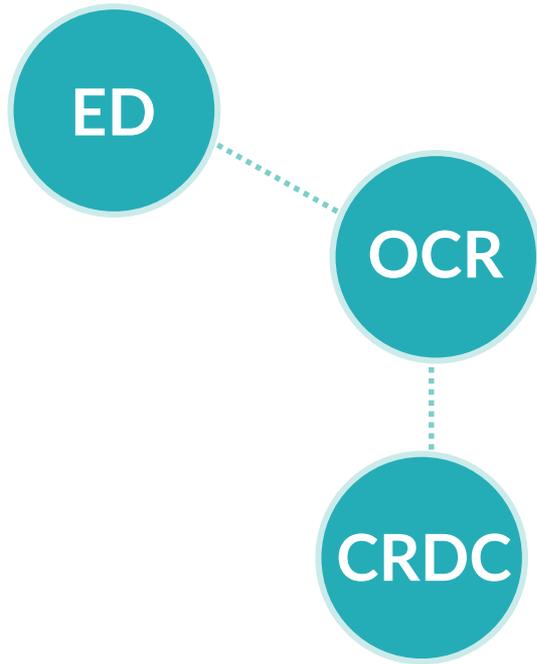
Open data on school bullying and harassment in NYC

Navigating the landscape



Open data on school bullying and harassment in NYC

Federal



- U.S. Department of Education (ED)
 - Office for Civil Rights (OCR)
 - **Civil Rights Data Collection (CRDC)**
95,507 schools nationwide

Open data on school bullying and harassment in NYC Federal

Civil Rights Data Collection (CRDC)
2013-14 school year



Allegations of Harassment or Bullying

Students Reported as Harassed or Bullied

Students Disciplined for Harassment or Bullying

Race-, Sex-,
Disability-
based

Open data on school bullying and harassment in NYC

Local



- New York City Department of Education (NYCDOE)
 - **NYC School Survey**
1,784 schools citywide

Open data on school bullying and harassment in NYC

Local

NYC School Survey
2013-14 school year

Students

At my school students
harass or bully other
students.

At my school students
harass or bully each other
based on differences ...

Open data on school bullying and harassment in NYC

Local

NYC School Survey
2013-14 school year

Students

At my school students
harass or bully other
students.

At my school students
harass or bully each other
based on differences ...

None of
the time



All of
the time

Open data on school bullying and harassment in NYC

Local

NYC School Survey
2013-14 school year

Students

At my school students
**harass or bully other
students.**

At my school students
harass or bully each other
based on differences ...

Teachers

At my school, students are
often harassed or bullied in
school.

At my school, there are
conflicts based on
differences ...

None of the time  All of the time

Open data on school bullying and harassment in NYC

Local

NYC School Survey
2013-14 school year

Students

At my school students
**harass or bully other
students.**

At my school students
harass or bully each other
based on differences ...

Teachers

At my school, students are
often harassed or bullied in
school.

At my school, there are
conflicts based on
differences ...

None of the time  All of the time

Strongly Disagree  Strongly Agree

Open data on school bullying and harassment in NYC

Local

NYC School Survey
2013-14 school year

Students

At my school students
**harass or bully other
students.**

At my school students
harass or bully each other
based on differences ...

Parents

At my child's school students
**harass or bully other
students.**

At my child's school students
harass or bully each other
based on differences ...

Teachers

At my school, students are
often harassed or bullied in
school.

At my school, there are
conflicts based on
differences ...

None of the time  All of the time

Strongly Disagree  Strongly Agree

Open data on school bullying and harassment in NYC

Local

NYC School Survey
2013-14 school year

Students

At my school students
**harass or bully other
students.**

At my school students
harass or bully each other
based on differences ...

None of the time  All of the time

Parents

At my child's school students
**harass or bully other
students.**

At my child's school students
harass or bully each other
based on differences ...

Strongly Disagree  Strongly Agree

Teachers

At my school, students are
often harassed or bullied in
school.

At my school, there are
conflicts based on
differences ...

Strongly Disagree  Strongly Agree

Open data on school bullying and harassment in NYC

Local

NYC School Survey
2013-14 school year

Students

At my school students
**harass or bully other
students.**

At my school students
harass or bully each other
based on differences ...

None of the time  All of the time

Parents

At my child's school students
**harass or bully other
students.**

At my child's school students
harass or bully each other
based on differences ...

Strongly Disagree  Strongly Agree
+ Don't Know

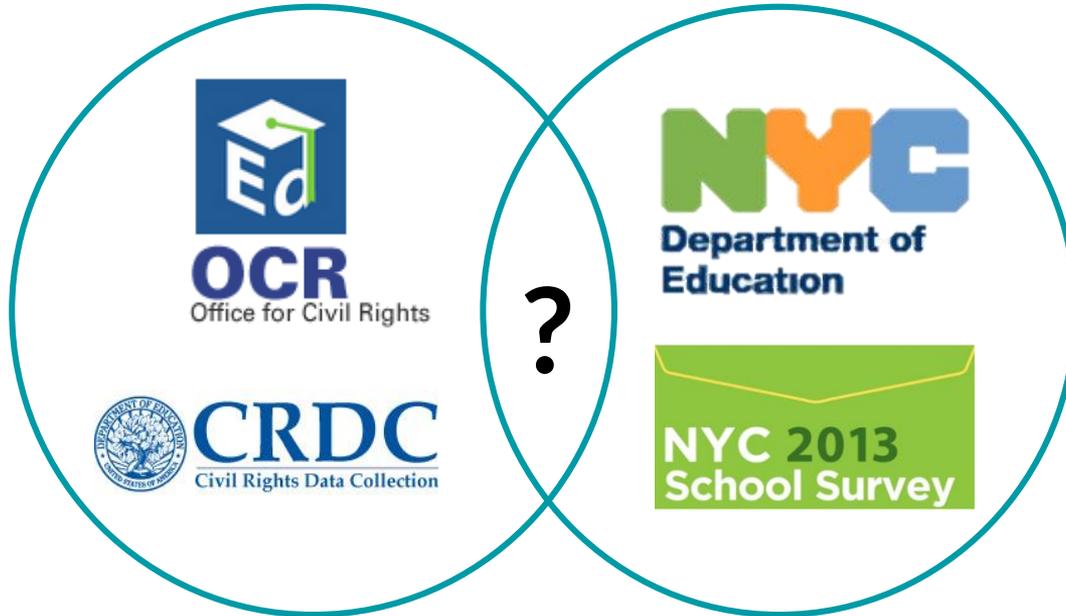
Teachers

At my school, students are
often harassed or bullied in
school.

At my school, there are
conflicts based on
differences ...

Strongly Disagree  Strongly Agree

Bringing **federal** and **local** together



Bringing federal and local together



NYC
DOE

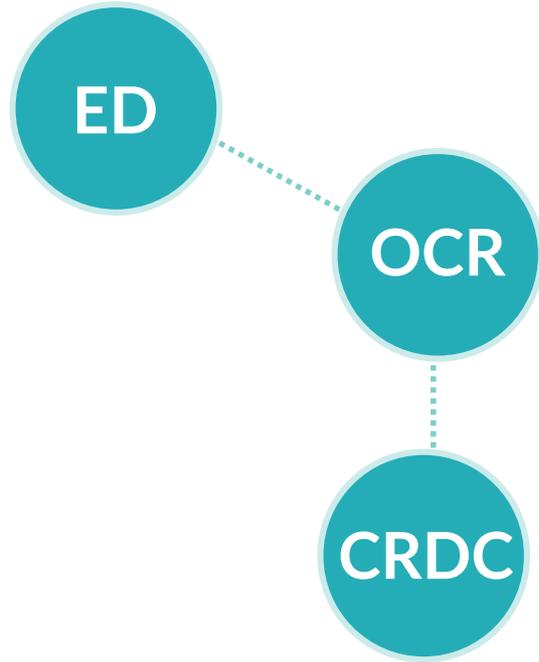
School
Survey

DBN

Schools in the NYC School Survey are identified by a **District Borough Number (DBN)** ...

DBN	School Name
24Q290	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus

Bringing federal and local together



SCH_NAME	COMBOKEY	DBN	School Name
PS 290	362058006223	24Q290	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus

... while schools in the Office for Civil Rights' Civil Rights Data Collection are identified by a **12-digit numeric ID** called the "COMBOKEY"

Bringing **federal** and **local** together: what we didn't do

```
def levenshteinDistance(str1, str2):
    m = len(str1)
    n = len(str2)
    lensum = float(m + n)
    d = []
    for i in range(m+1):
        d.append([i])
    del d[0][0]
    for j in range(n+1):
        d[0].append(j)
    for j in range(1,n+1):
        for i in range(1,m+1):
            if str1[i-1] == str2[j-1]:
                d[i].insert(j,d[i-1][j-1])
            else:
                minimum = min(d[i-1][j]+1, d[i][j-1]+1, d[i-1][j-1]+2)
                d[i].insert(j, minimum)
    ldist = d[-1][-1]
    ratio = (lensum - ldist)/lensum
    return {'distance':ldist, 'ratio':ratio}

print(levenshteinDistance("A.C.E. Academy for Scholars at the Geraldine Ferraro Campus",
    "PS 290"))
```

- P.S. 188 The Island School
- PS 188 ISLAND SCHOOL (THE)

```
{'distance': 61, 'ratio': 0.06153846153846154}
```



NYC School Survey

DBN	School Name
24Q290	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus

Office for Civil Rights

SCH_NAME	COMBOKEY
PS 290	362058006223

Bringing **federal** and **local** together: what we did



NYC Open Data Portal: School Locations

ATS SYSTEM CODE	LOCATION_NAME	BEDS NUMBER
24Q290	A.C.E. Academy for Scholars at the Geraldine Fe	342400010290



Bringing federal and local together: what we did

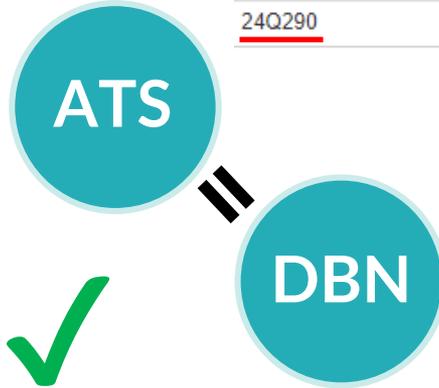
NYC School Survey

DBN	School Name
<u>24Q290</u>	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus

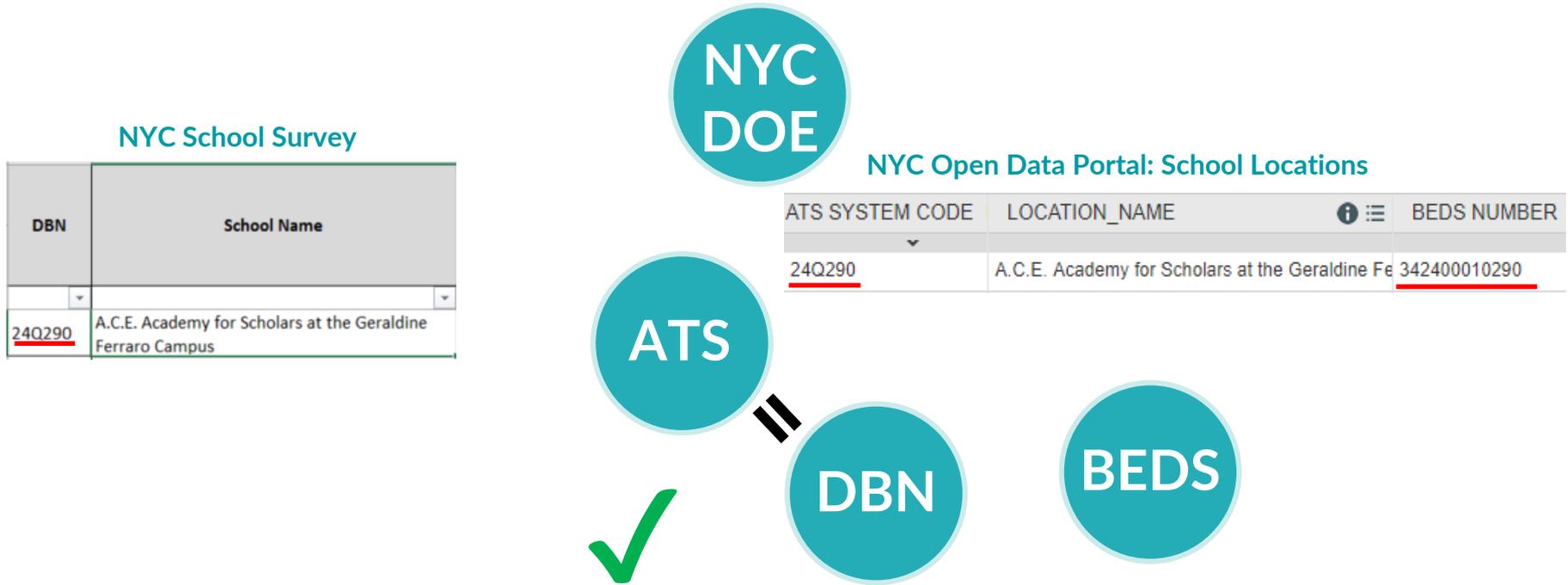
NYC
DOE

NYC Open Data Portal: School Locations

ATS SYSTEM CODE	LOCATION_NAME	BEDS NUMBER
<u>24Q290</u>	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus	342400010290



Bringing federal and local together: what we did



Bringing federal and local together: what we did

NYC School Survey

DBN	School Name
<u>240290</u>	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus

Office for Civil Rights

SCH_NAME	COMBOKEY
PS 290	<u>362058006223</u>

NYC
DOE

NYC Open Data Portal: School Locations

ATS SYSTEM CODE	LOCATION_NAME	BEDS NUMBER
<u>24Q290</u>	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus	<u>342400010290</u>

ATS

=

DBN

BEDS

X



Bringing federal and local together: what we did



NYC School Survey

DBN	School Name
24Q290	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus

NYC Open Data Portal: School Locations

ATS SYSTEM CODE	LOCATION_NAME	BEDS NUMBER
24Q290	A.C.E. Academy for Scholars at the Geraldine Fe	342400010290



Bringing federal and local together: what we did



NYC School Survey

DBN	School Name
24Q290	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus

NYC Open Data Portal: School Locations

ATS SYSTEM CODE	LOCATION_NAME	BEDS NUMBER
24Q290	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus	342400010290



Bringing federal and local together: what we did



Common Core of Data

NCESSCH	SEASCH	SCHNAM
360009806223	<u>342400010290</u>	PS 290



NYC School Survey

DBN	School Name
24Q290	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus

NYC Open Data Portal: School Locations

ATS SYSTEM CODE	LOCATION_NAME	BEDS NUMBER
<u>24Q290</u>	A.C.E. Academy for Scholars at the Geraldine Fe	<u>342400010290</u>



Bringing federal and local together: what we did

NYC School Survey

DBN	School Name
24Q290	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus

NYC Open Data Portal: School Locations

ATS SYSTEM CODE	LOCATION_NAME	BEDS NUMBER
24Q290	A.C.E. Academy for Scholars at the Geraldine Fe	342400010290

ED

NYC
DOE

Common Core of Data

NCESSCH	SEASCH	SCHNAM
360009806223	342400010290	PS 290

NCES

CCD

DBN

BEDS

Bringing federal and local together: what we did

NYC School Survey

DBN	School Name
24Q290	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus

NYC Open Data Portal: School Locations

ATS SYSTEM CODE	LOCATION_NAME	BEDS NUMBER
24Q290	A.C.E. Academy for Scholars at the Geraldine Ferraro Campus	342400010290

Common Core of Data

NCESSCH	SEASCH	SCHNAM
360009806223	342400010290	PS 290

Office for Civil Rights

SCH_NAME	COMBOKEY
PS 290	362058006223

ED

NYC DOE

NCES

CCD

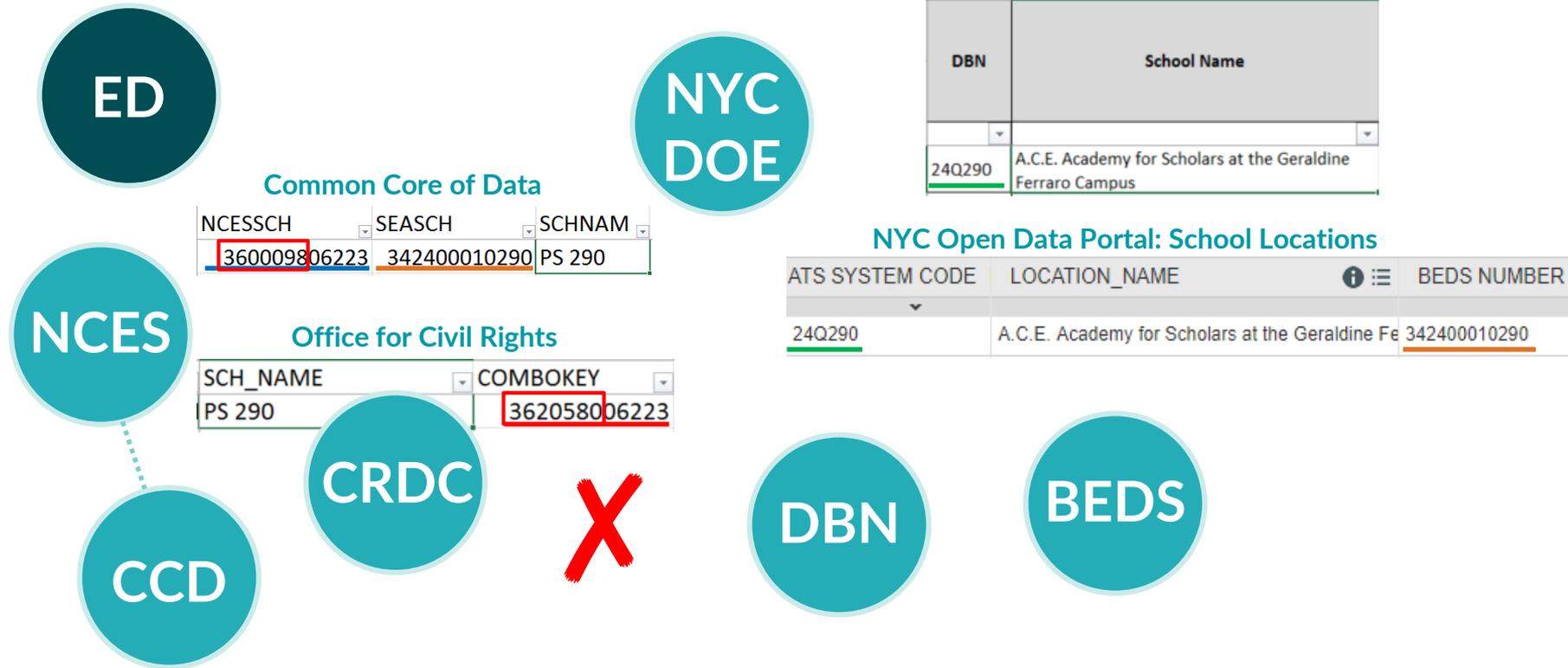
CRDC

?

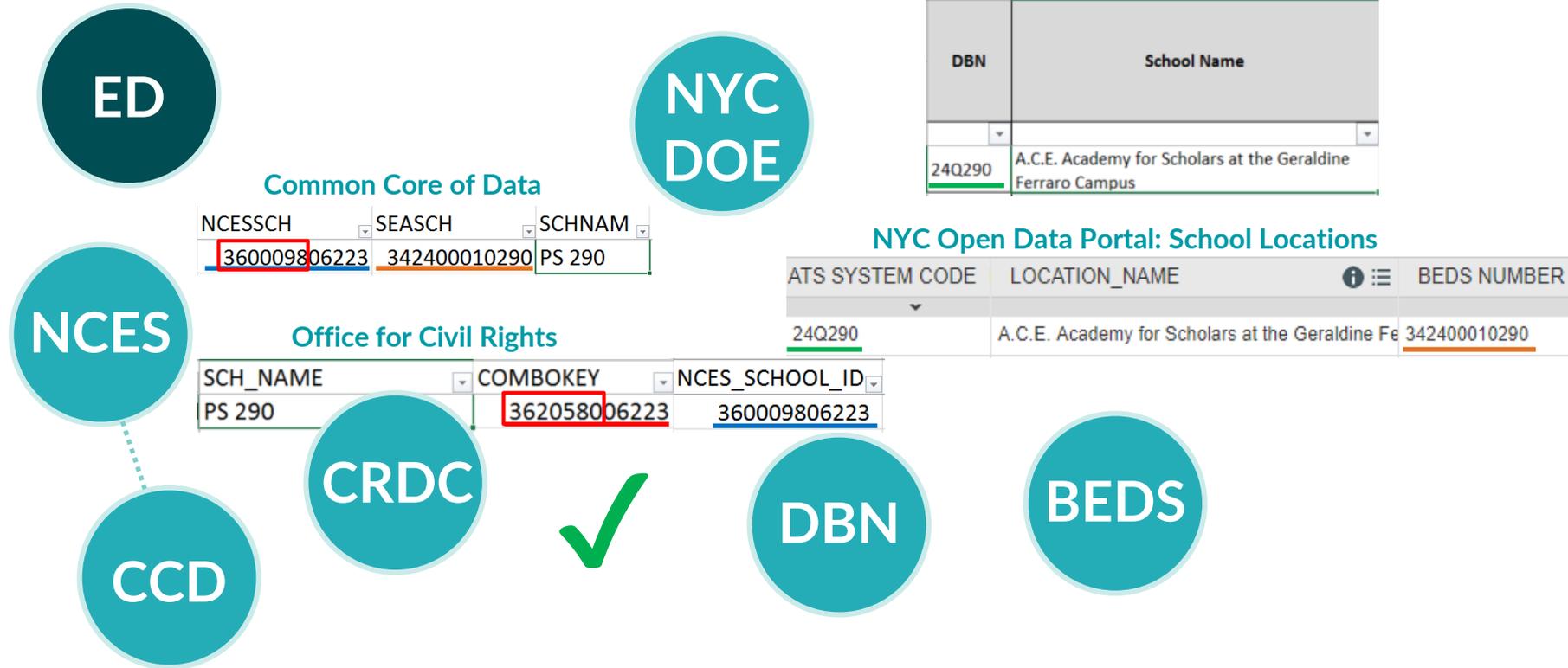
DBN

BEDS

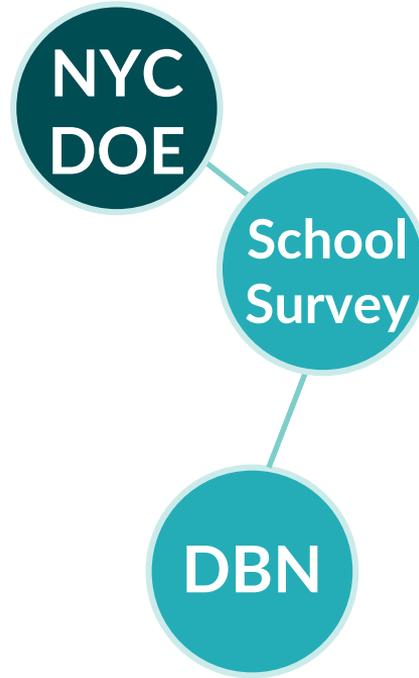
Bringing federal and local together: what we did



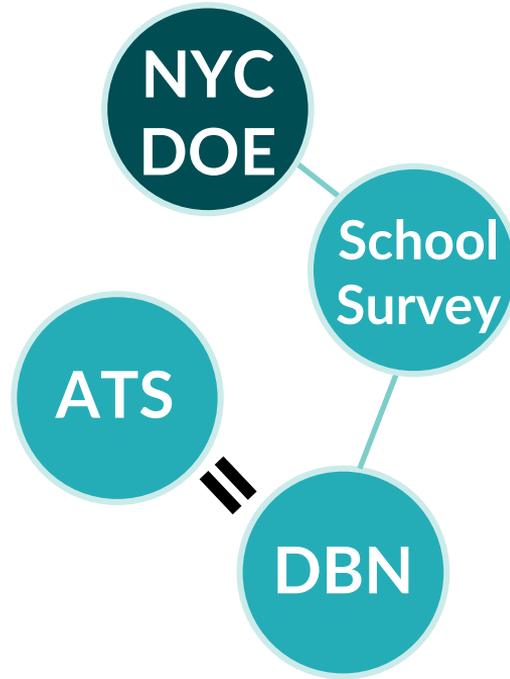
Bringing federal and local together: what we did



Local

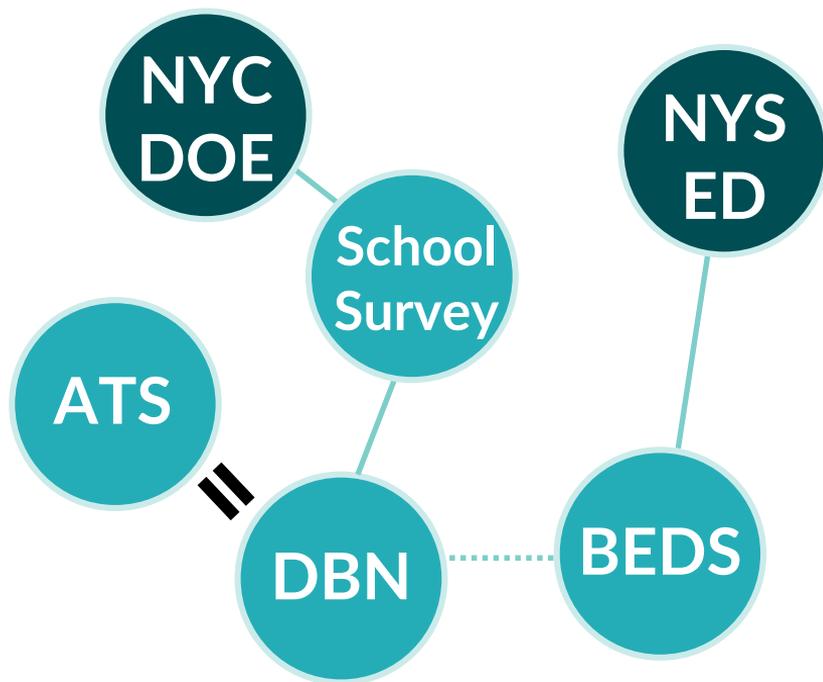


Local



Local

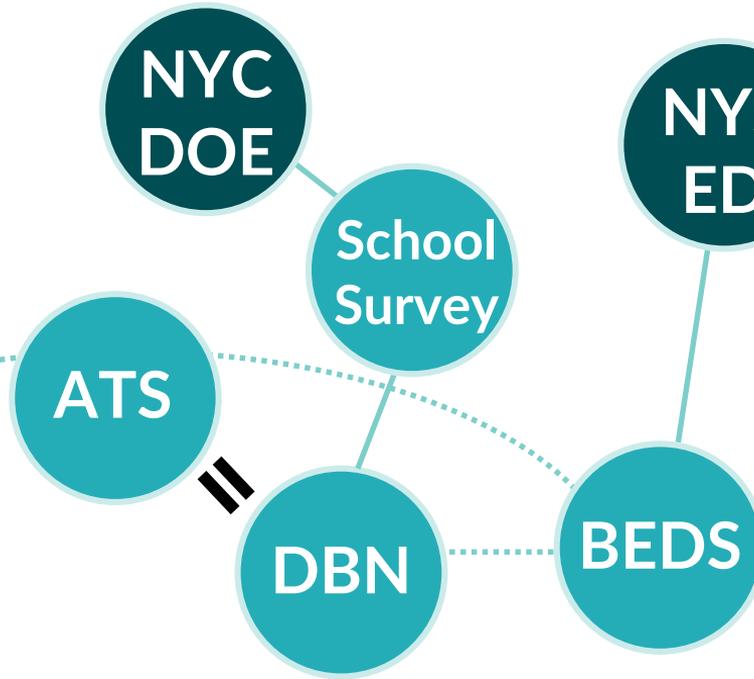
State



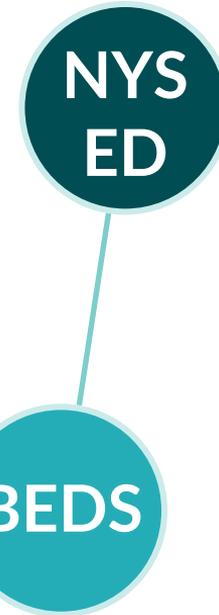
Federal



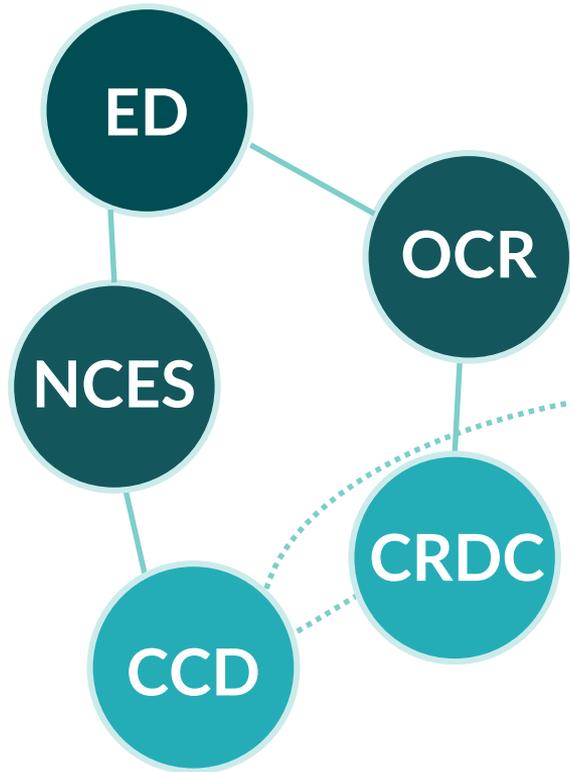
Local



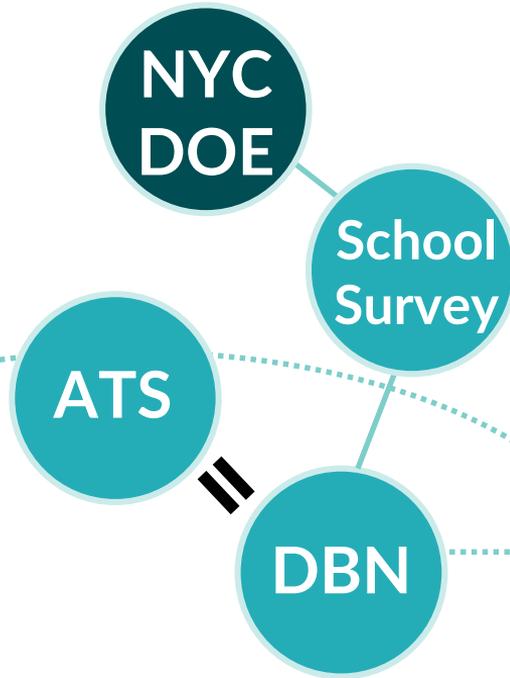
State



Federal



Local



State



Bringing federal and local together: what we did

Local (DBN) → State (BEDS) → Federal (NCESSCH) → Federal (COMBOKEY)

one school, many “unique” school IDs:

dbn	beds	ncessch	combokey	school_name
24Q290	342400010290	360009806223	362058006223	A.C.E. Academy for Scholars at the Geraldine F...



How to Navigate New York's Murky System for Tracking Bullying in Schools

By Amy Zimmer and Nigel Chiwaya | September 9, 2015 7:42am



DNAinfo New York

Bullying in New York City Public Schools

A.C.E. Academy for Scholars at the Geraldine Ferraro Campus

Grade Level: **Elementary**
 This school reported:
0 incidents to VADIR. (The citywide median is 2.)^a
0 incidents to DASA.^b
[Find Another School](#)

a. VADIR: Violent and Disruptive Incidents Database, NYSED, 2013-14
 b. DASA: Bias and Cyberbullying Database created under the Dignity for All
 Survey Results are from the NYC Department of Education Student Survey, are not surveyed, and schools with response rates below 38 percent are ex

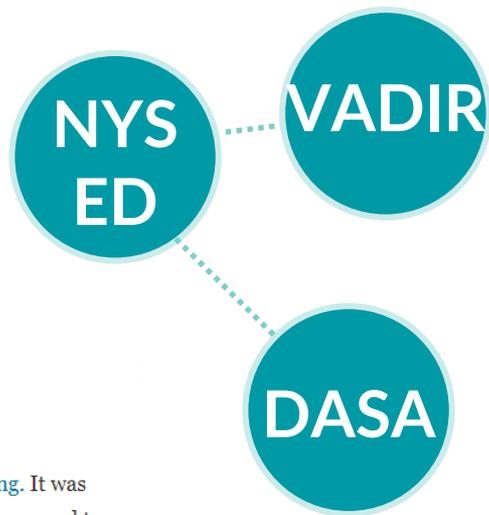
► How city schools are (or aren't) tracking bullying

Two separate state Education Department databases track bullying.

The first is called VADIR — or Violent and Disruptive Incident Reporting. It was created in 2000 and tracks bullying using 20 categories. A formula is then used to crunch the data from the various categories to determine whether a school should be on a "persistently dangerous" list.

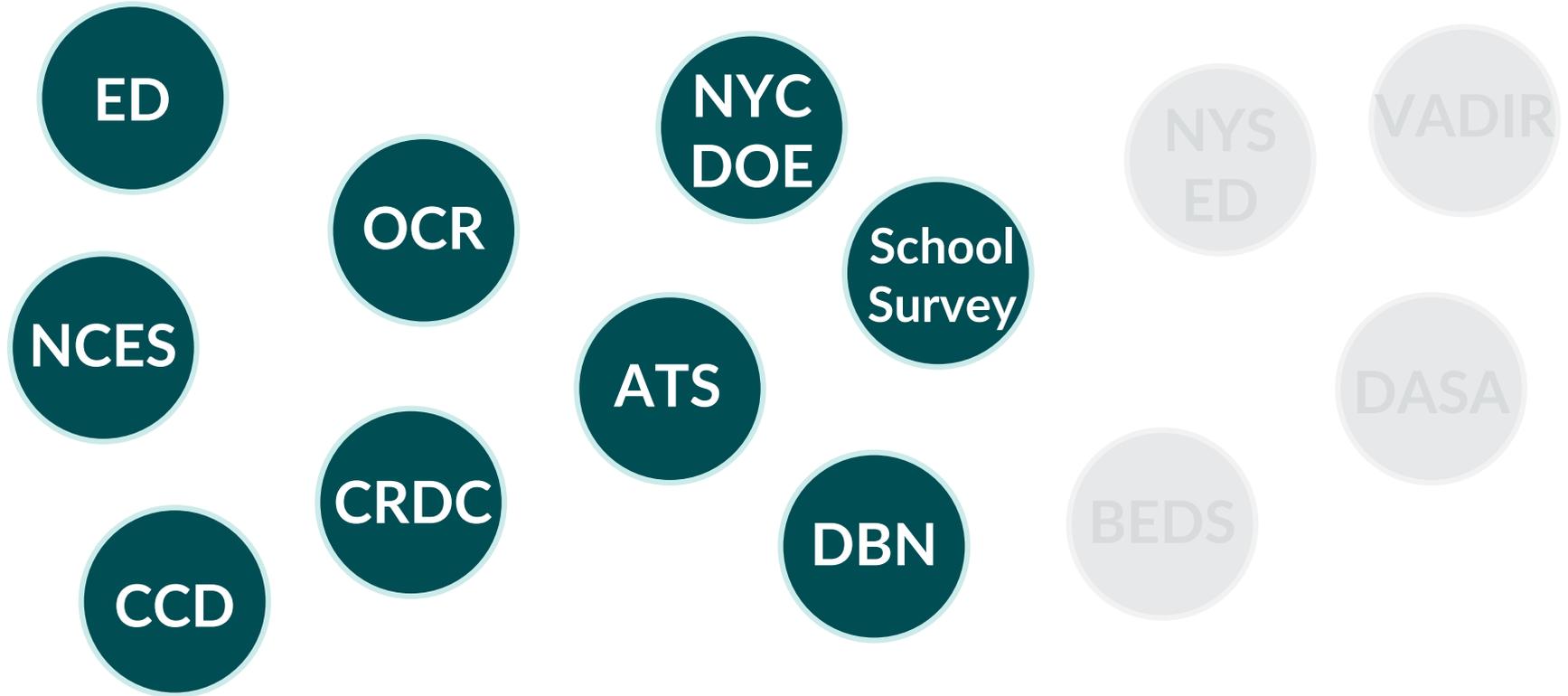
The second, created by the Dignity for all Students Act, was implemented by the state in 2012 to address cyberbullying and bias-related bullying (targeting race, religion, gender or weight, for example) as part of a bigger effort to help make schools safe environments free from discrimination, harassment and bullying.

State



Open data on school bullying and harassment in NYC

Federal + Local

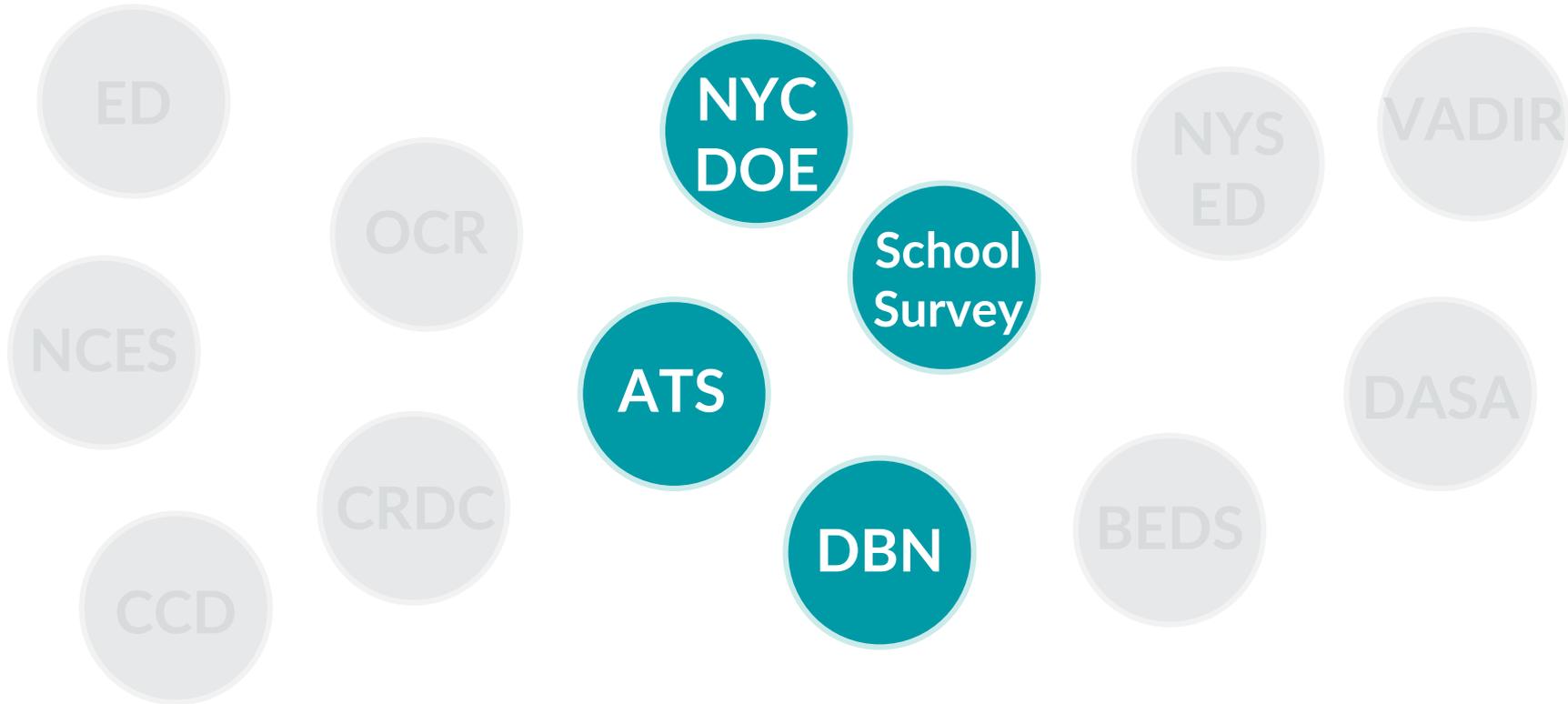


Goals of Analysis

- Local **NYC School Survey** – differences in perceptions of bullying and harassment between **students, parents, and teachers**.
- Federal **Civil Rights Data** – prevalence of bullying and harassment **allegations**.
- Local-Federal **survey agreement** – investigating school characteristics associated with **(dis)agreement**.

Open data on school bullying and harassment in NYC

Local





NYC School Survey

2013-14 school year

bullying/harassment

Students

At my school students
harass or bully other
students.

At my school students
harass or bully each other
based on differences ...

Parents

At my child's school students
harass or bully other
students.

At my child's school students
harass or bully each other
based on differences ...

Teachers

At my school, students are
often **harassed or bullied** in
school.

At my school, there are
conflicts **based on**
differences ...



NYC School Survey

2013-14 school year

bullying/harassment based on differences

Students

At my school students
harass or bully other
students.

At my school students
harass or bully each other
based on differences ...

Parents

At my child's school students
harass or bully other
students.

At my child's school students
harass or bully each other
based on differences ...

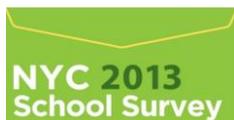
Teachers

At my school, students are
often **harassed or bullied** in
school.

At my school, there are
conflicts **based on
differences ...**

NYC School Survey

2013-14 school year



Students

bullying/harassment

bullying/harassment
based on differences

Parents

bullying/harassment

bullying/harassment
based on differences

Teachers

bullying/harassment

bullying/harassment
based on differences

NYC School Survey

2013-14 school year



Students

bullying/harassment

bullying/harassment
based on differences

None of the time  All of the time

Parents

bullying/harassment

bullying/harassment
based on differences

Strongly Disagree  Strongly Agree
+ Don't Know

Teachers

bullying/harassment

bullying/harassment
based on differences

Strongly Disagree  Strongly Agree

NYC School Survey

2013-14 school year



Students

bullying/harassment

bullying/harassment
based on differences

None of the time

1



All of the time

4

Parents

bullying/harassment

bullying/harassment
based on differences

Strongly Disagree

1



+ Don't Know

Strongly Agree

4

Teachers

bullying/harassment

bullying/harassment
based on differences

Strongly Disagree

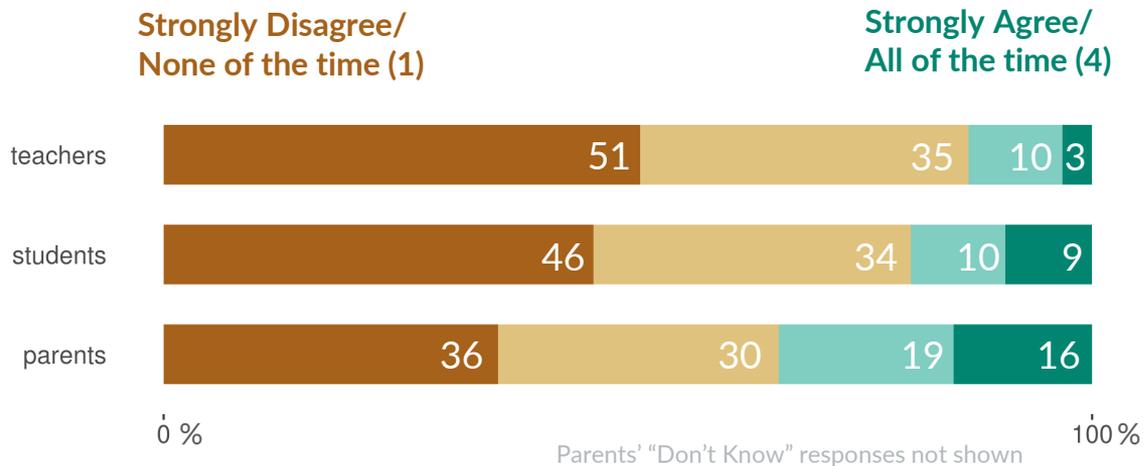
1



Strongly Agree

4

Students bully or harass each other / there are conflicts based on differences



- The vast majority of teachers say they disagree (35%) or strongly disagree (51%) there are conflicts based on differences.
- Parents are the most likely (35%) to feel that students harass/bully each other based on differences.
- Students' feelings about bullying/harassment are "in the middle," between parents' and teachers'.

NYC School Survey

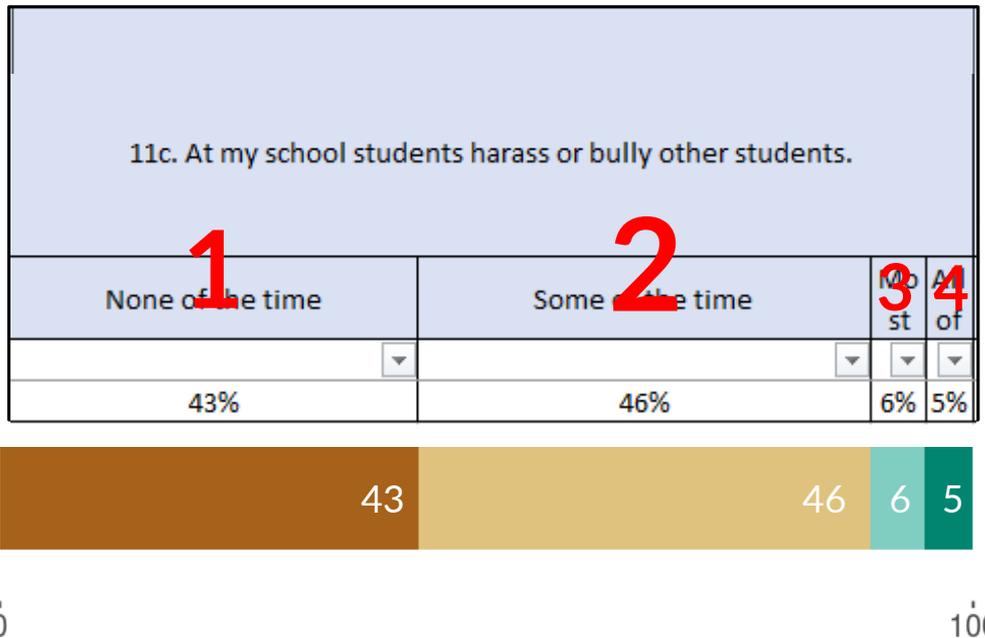
Creating a common scale for comparing schools to each other

11c. At my school students harass or bully other students.

1 None of the time	2 Some of the time	3 Most of the time	4 All of the time
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
43%	46%	6%	5%

NYC School Survey

Creating a common scale for comparing schools to each other

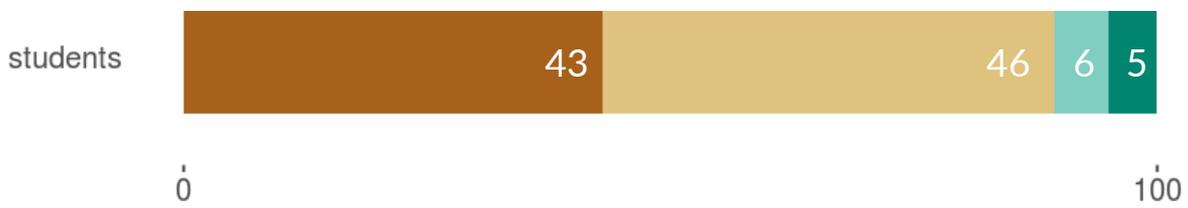


NYC School Survey

Creating a common scale for comparing schools to each other

11c. At my school students harass or bully other students.

None of the time	Some of the time	Most of the time	All of the time
43%	46%	6%	5%



Average school "score" =
1.73

NYC School Survey

Creating a common scale for comparing schools to the citywide average

11c. At my school students harass or bully other students.			
1 None of the time	2 Some of the time	3 Most of the time	4 All of the time
43%	46%	6%	5%

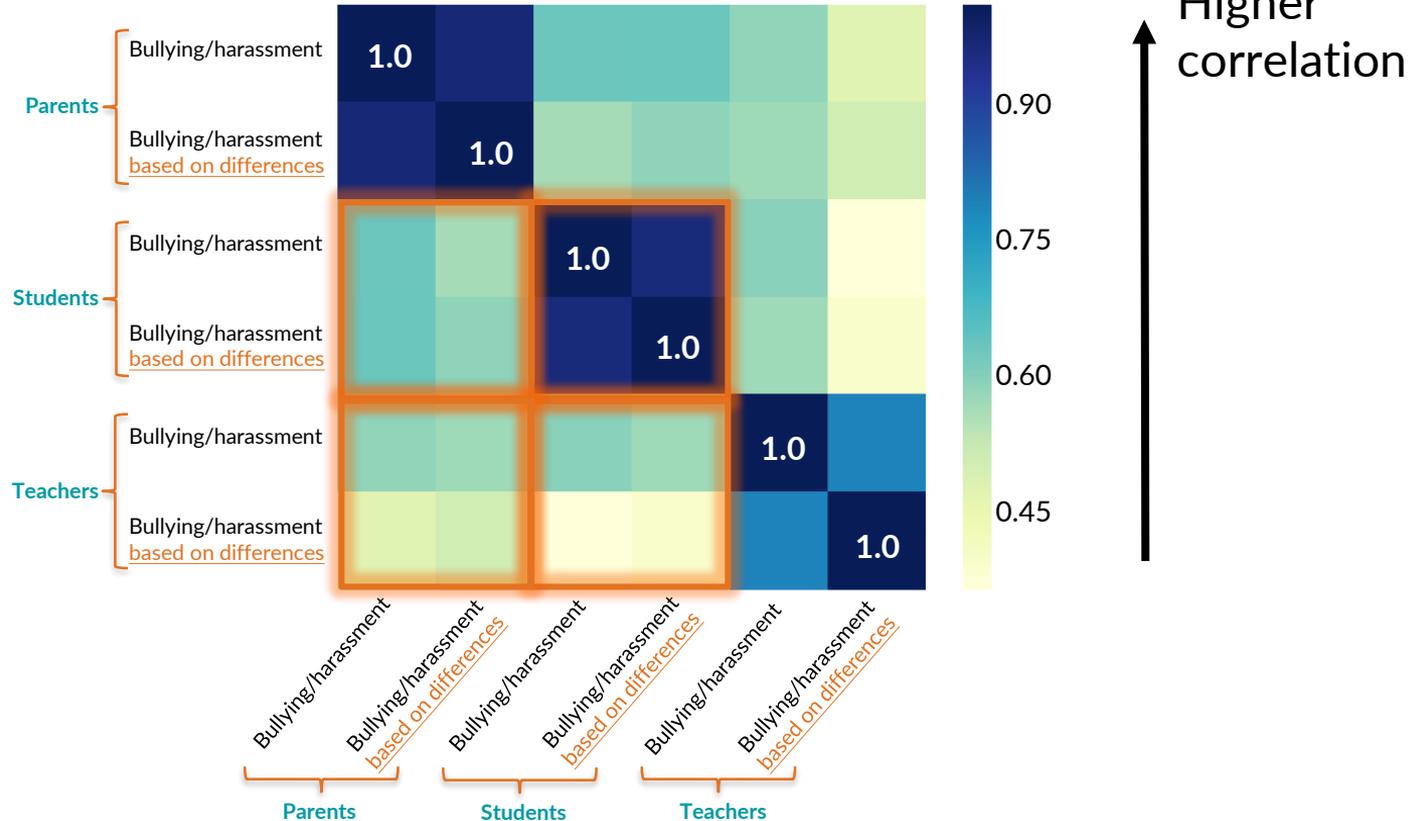
average school
"score"
1.73

average NYC
"score"
1.96

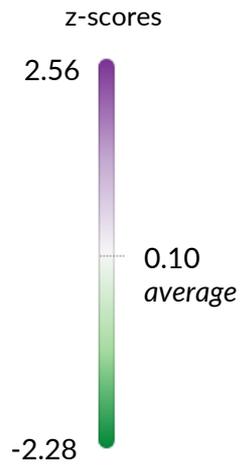
school "z-
score"
-0.72

- Given the average school "score" for each school...
- Calculate the citywide average "score"
- Compare each school's "score" to the citywide average
- We do this for both bullying/harassment questions
 - Parents
 - Students
 - Teachers

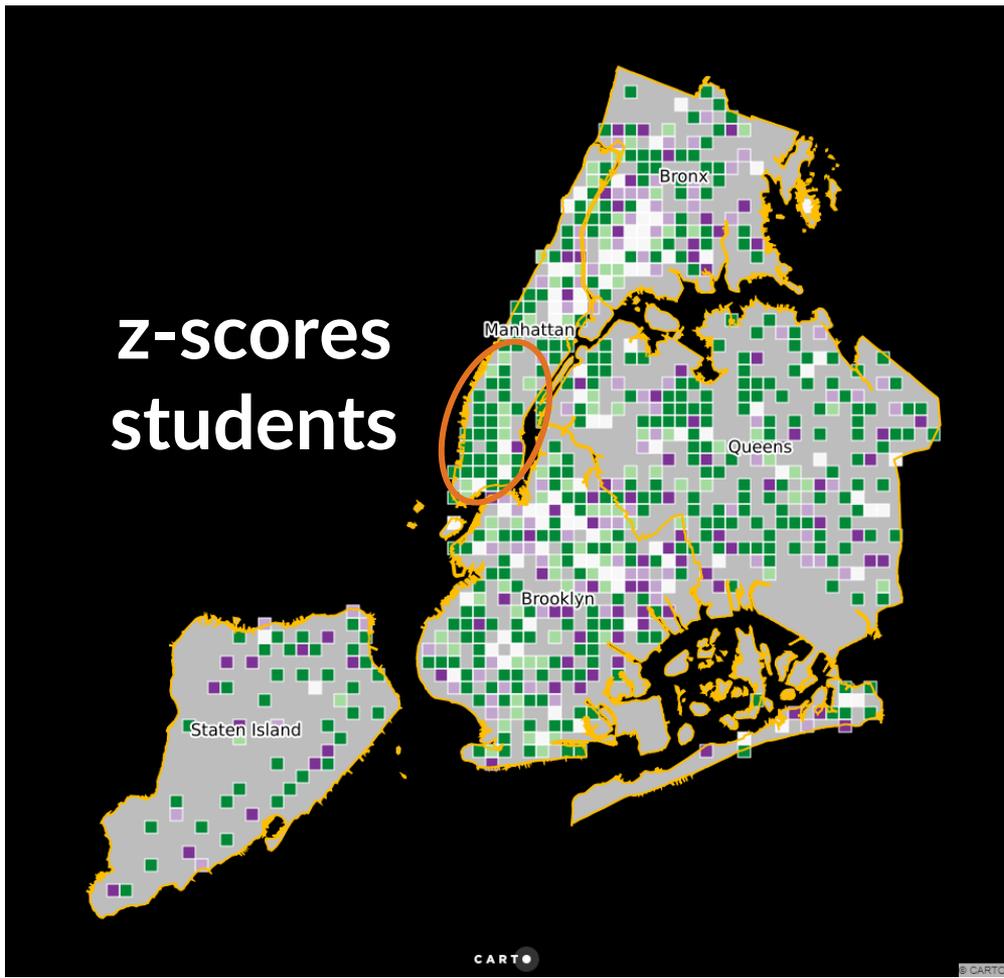
Responses are highly correlated



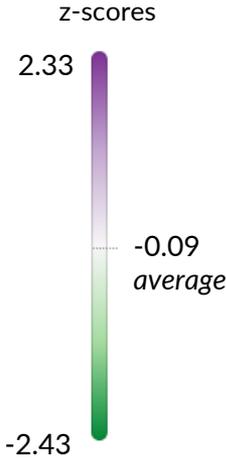
Bullying/harassment based on differences:
student responses



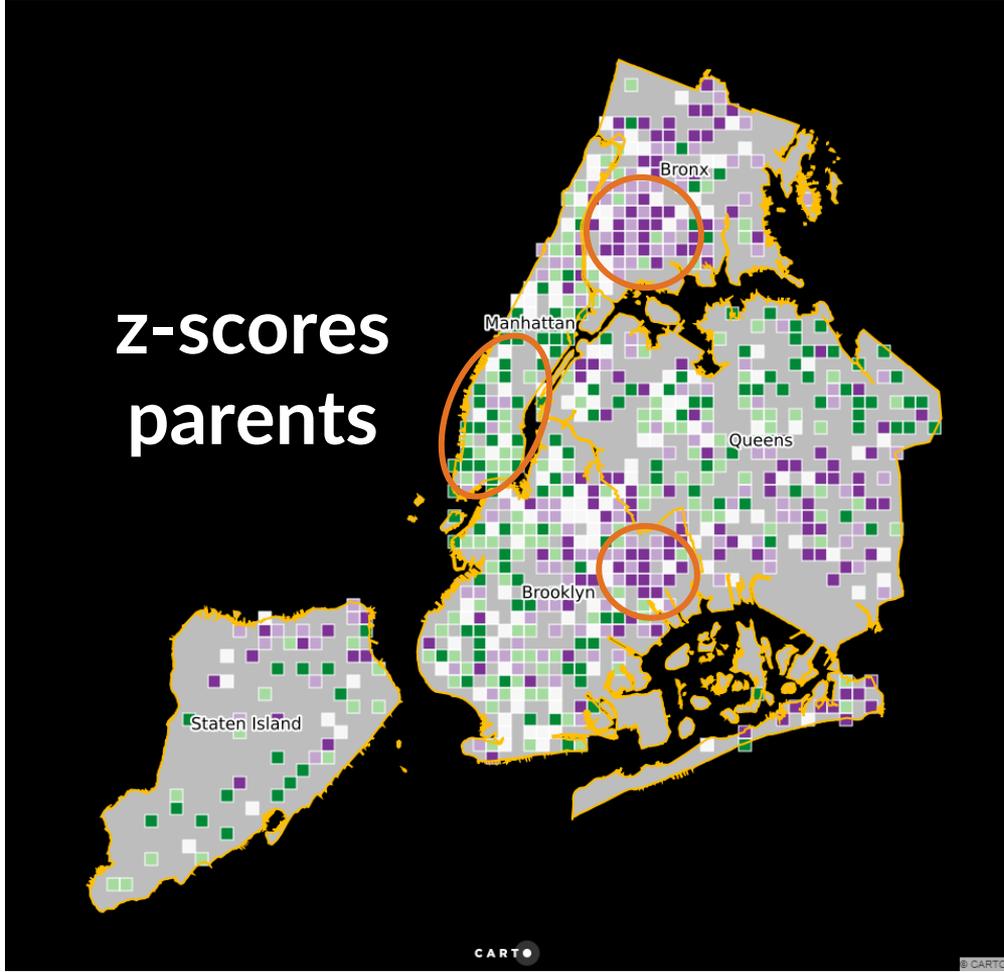
**z-scores
students**



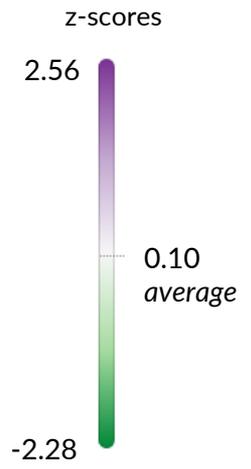
Bullying/harassment based on differences:
parent responses



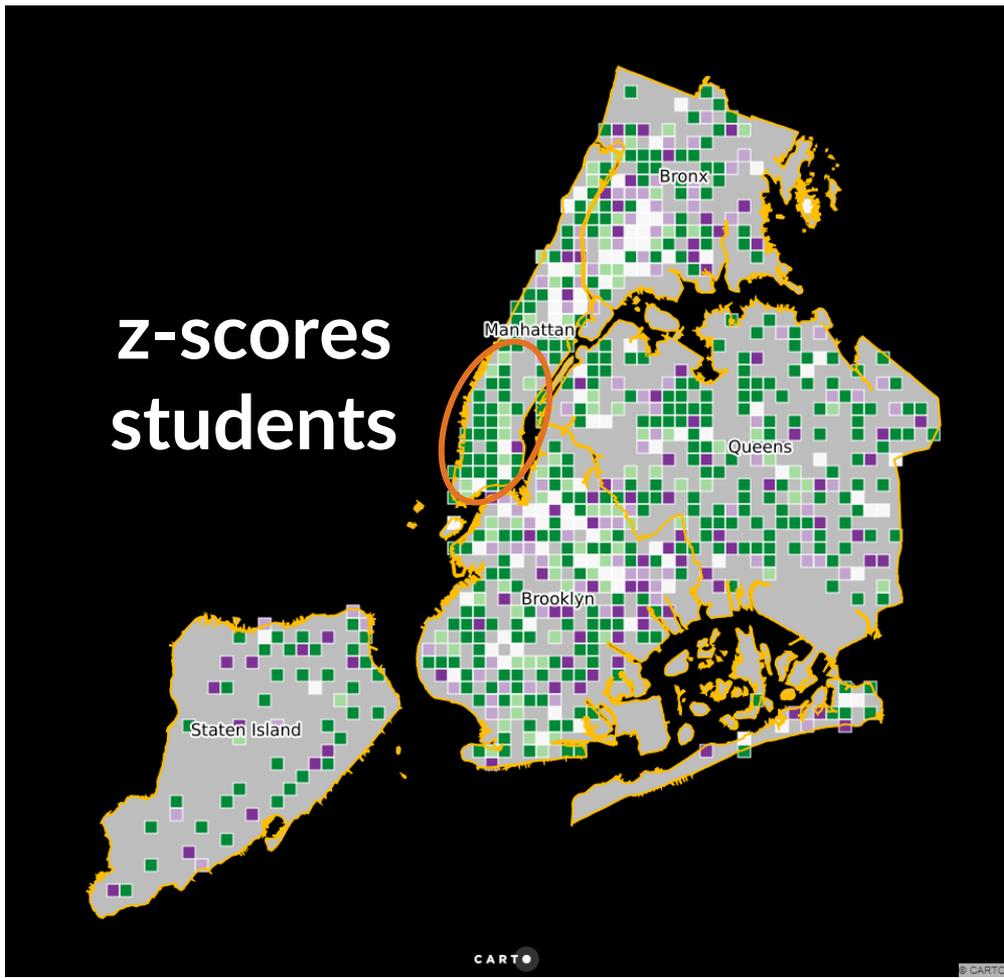
**z-scores
parents**



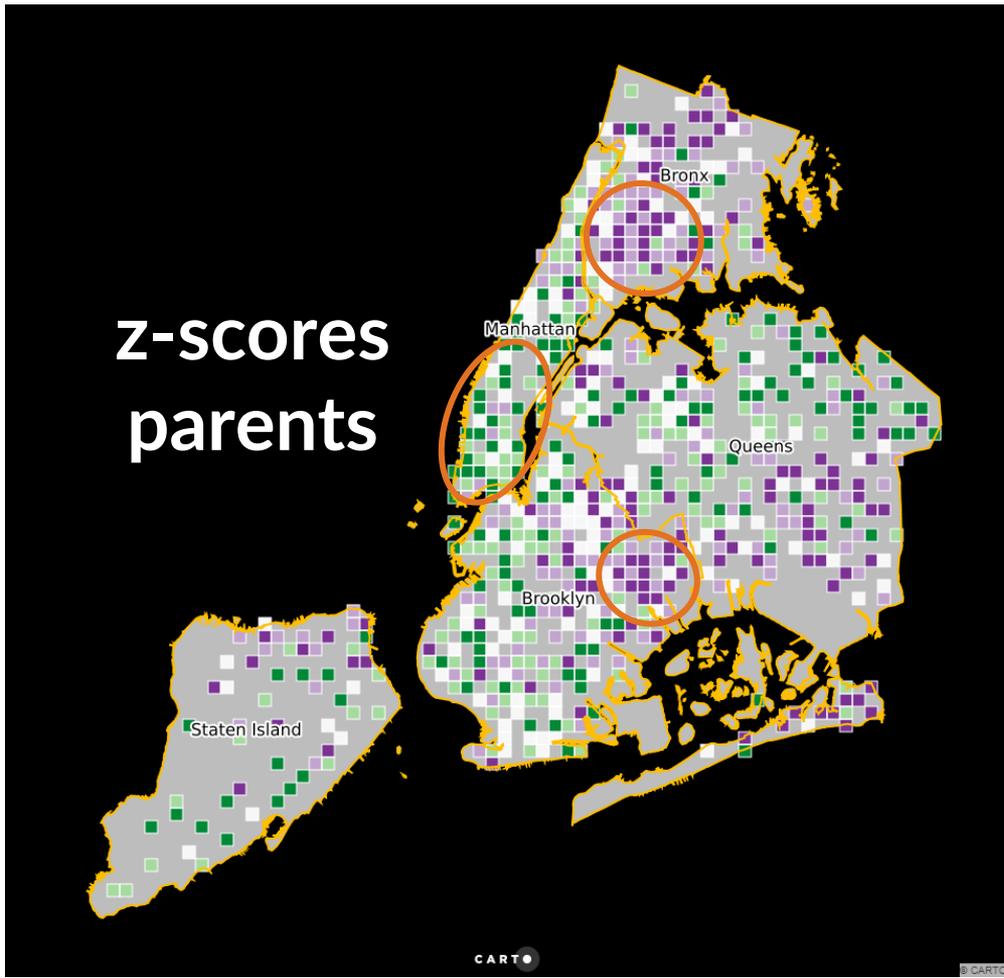
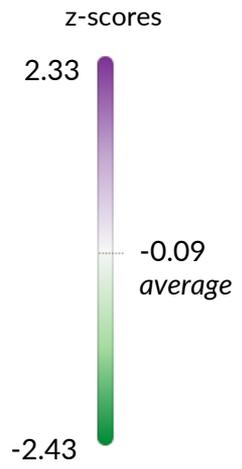
Bullying/harassment based on differences:
student responses



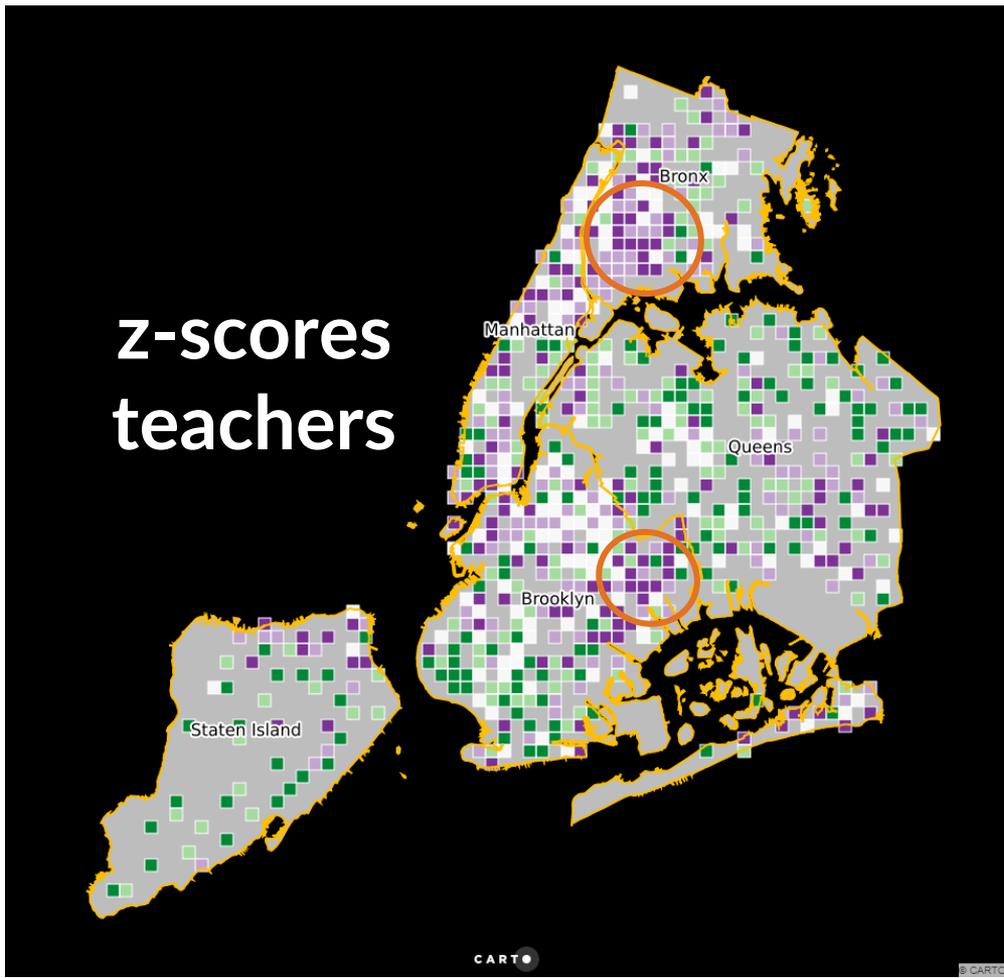
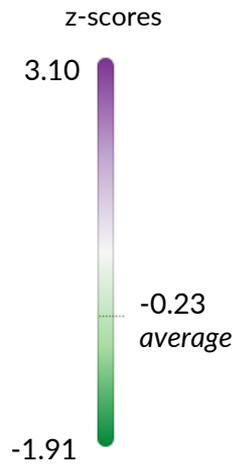
**z-scores
students**



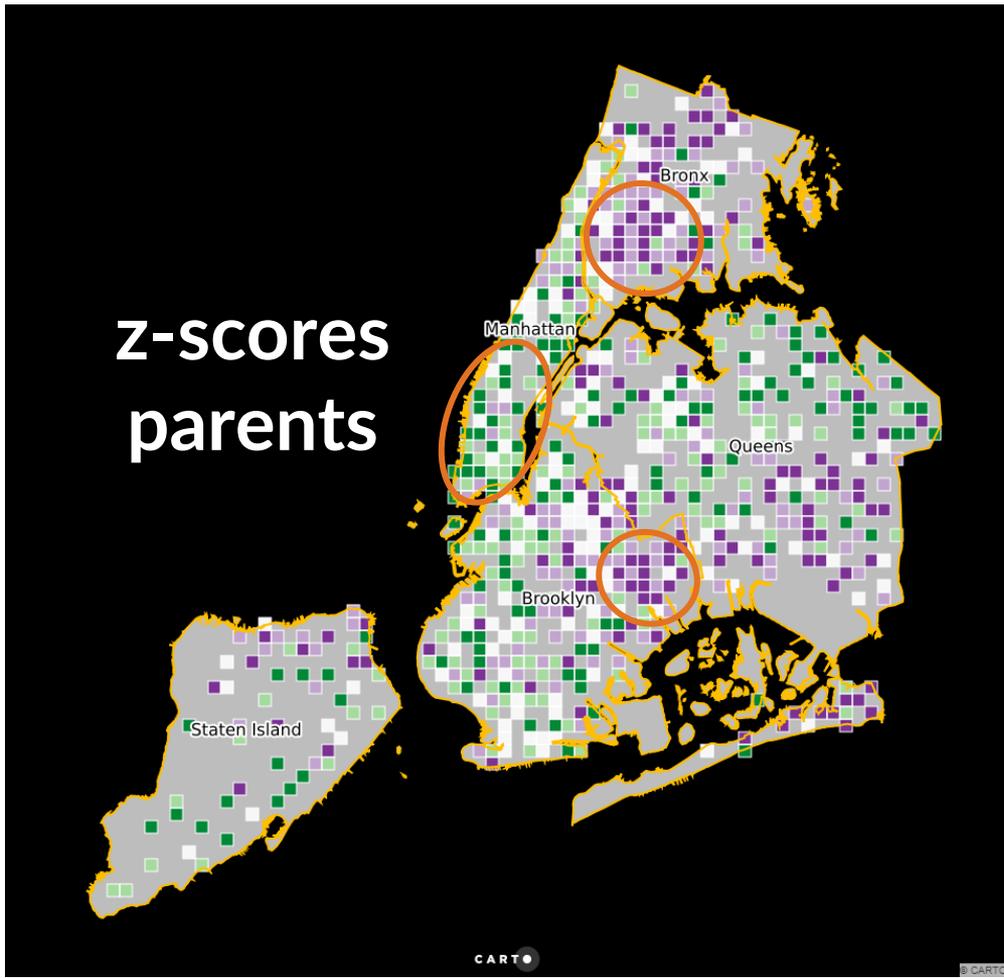
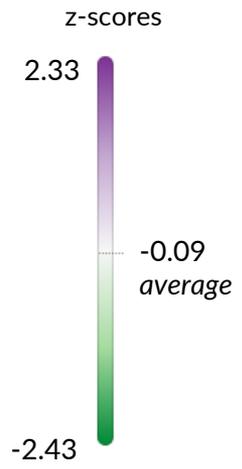
Bullying/harassment based on differences:
parent responses



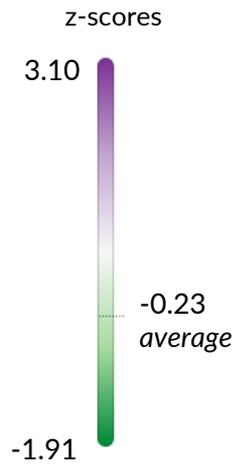
Bullying/harassment based on differences: teacher responses



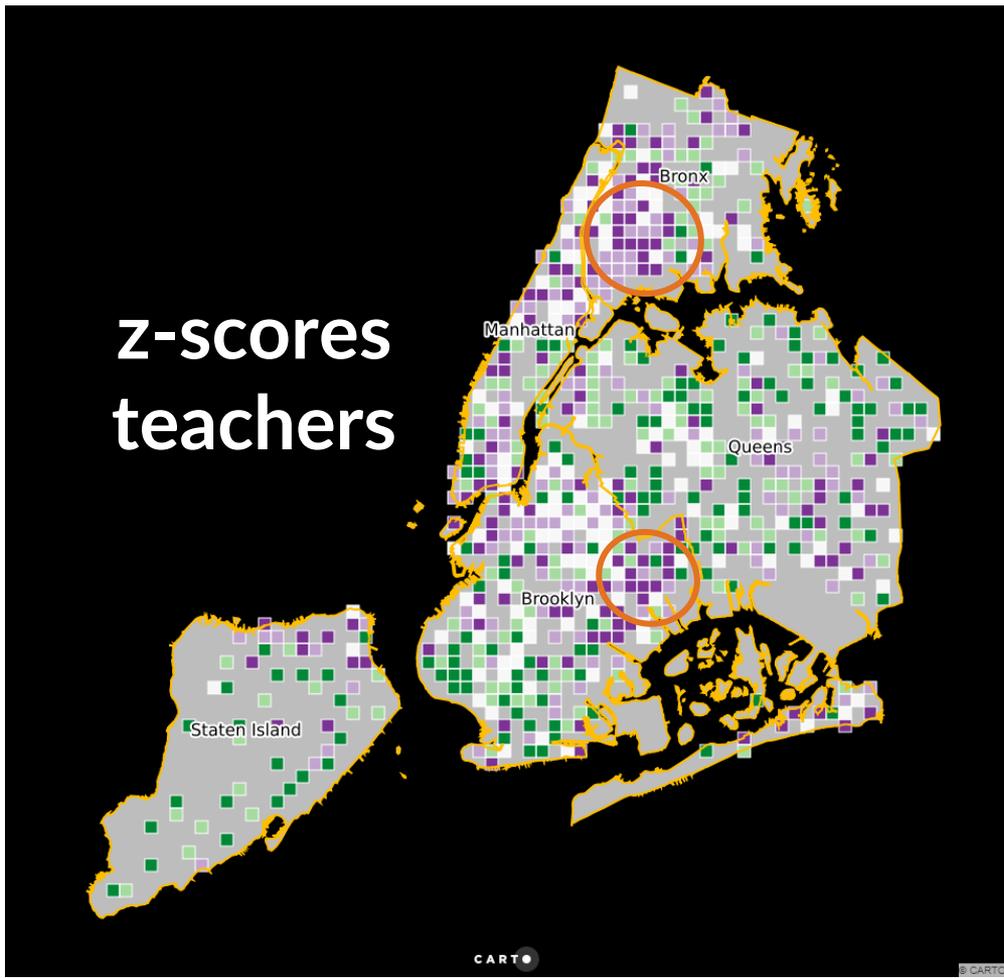
**Bullying/harassment based on differences:
parent responses**



Bullying/harassment based on differences:
teacher responses



**z-scores
teachers**

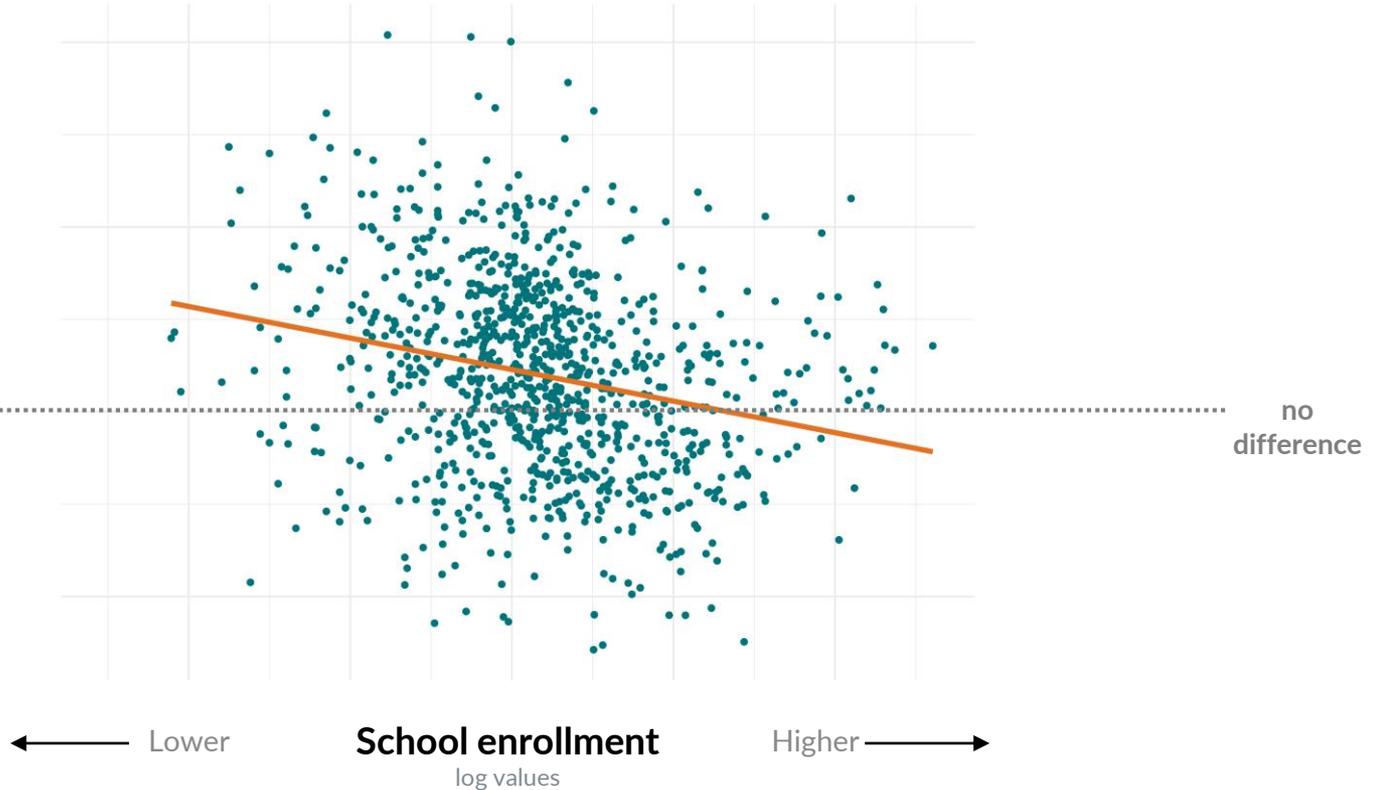


Teachers and students appear to disagree in the largest and smallest schools (though this correlation is small)

Feelings about the prevalence
of bullying/harassment
based on differences

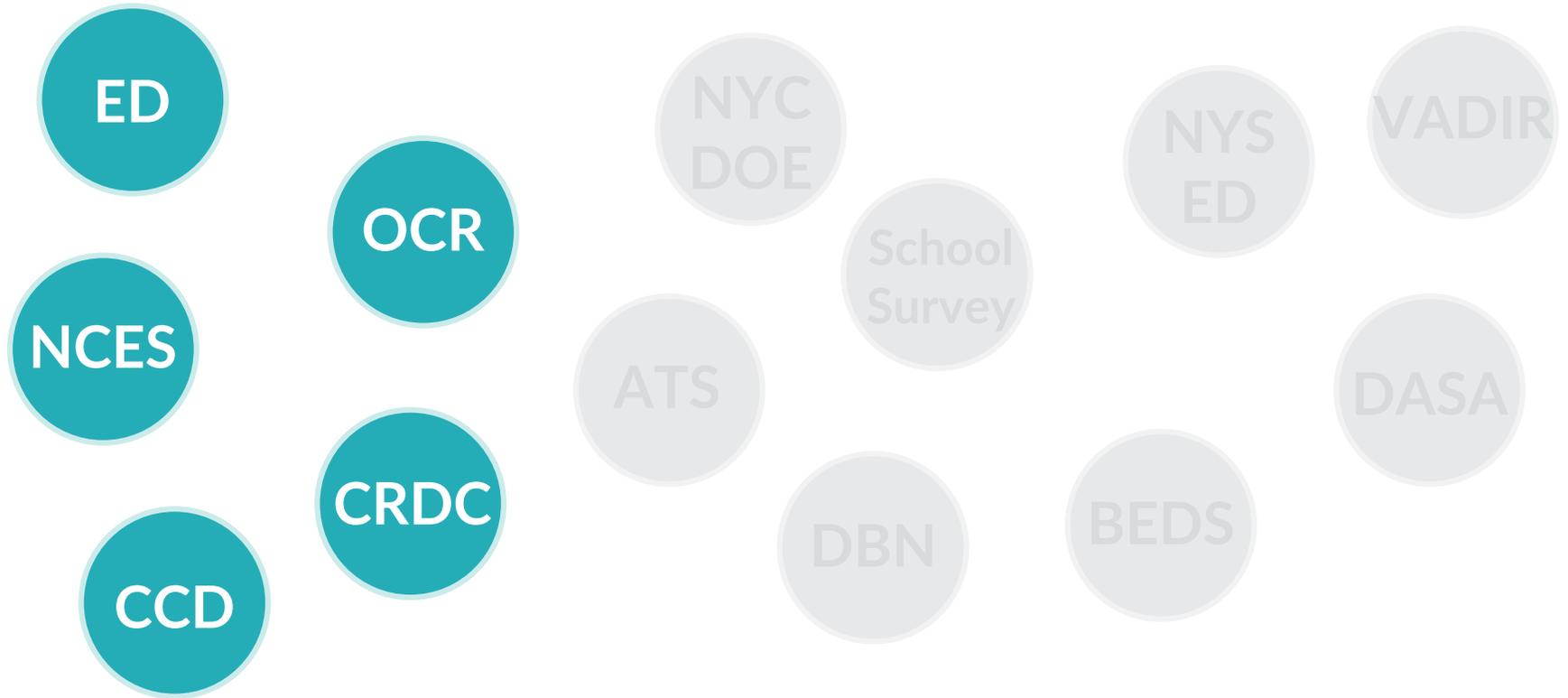
Teachers feel
more strongly

Students feel
more strongly



Open data on school bullying and harassment in NYC

Federal



Open data on school bullying and harassment in NYC Federal

Civil Rights Data Collection (CRDC)
2013-14 school year



Allegations of Harassment or Bullying

Students Reported as Harassed or Bullied

Students Disciplined for Harassment or Bullying

Race-, Sex-,
Disability-
based

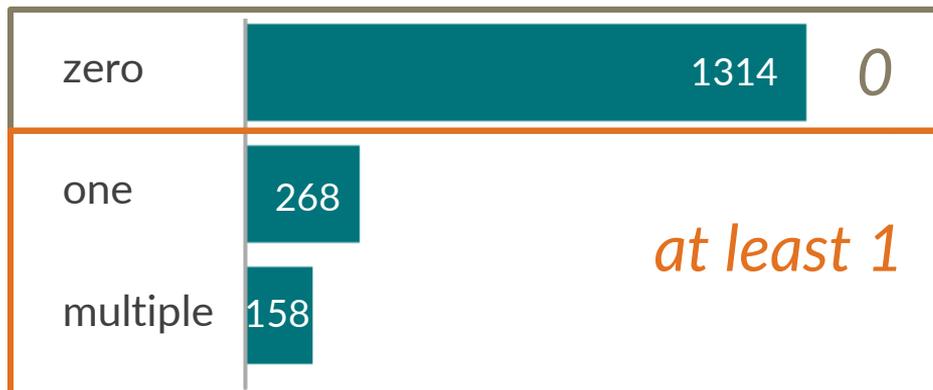


Civil Rights Data Collection

Office for Civil Rights

2013-14 school year

of schools reporting _____
allegations of bullying/harassment



- Sex-based bullying/harassment is the most commonly alleged category, disability-based is the least
- 75% of NYC schools report zero allegations (of any type) of bullying/harassment to the federal OCR. Nationally, this number is 80%.

Characteristics of schools with ...

	<u>0 allegations</u>	<u>1+ allegations</u>
Total Enrollment	576	713
Students with disabilities (%)	18	19
English Language Learners (%)	14	12
Parent responses z-scores*	-0.12	0.37
Student responses z-scores*	-0.15	0.27
Teacher responses z-scores*	-0.15	0.45



1314 schools (75%) reporting 0 allegations

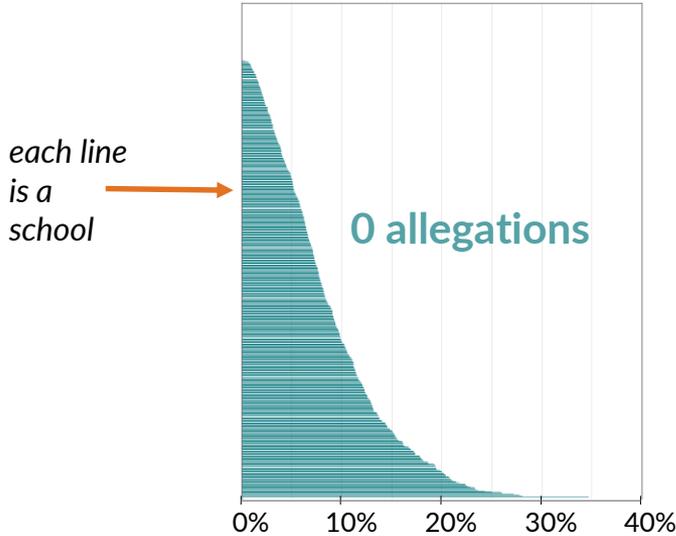
426 schools reporting 1+ allegations

Additional characteristics not shown as there were no significant differences. *z-scores are shown for the bullying/harassment due to differences question (results for responses the general question were similar)

How much do the two surveys (dis)agree?



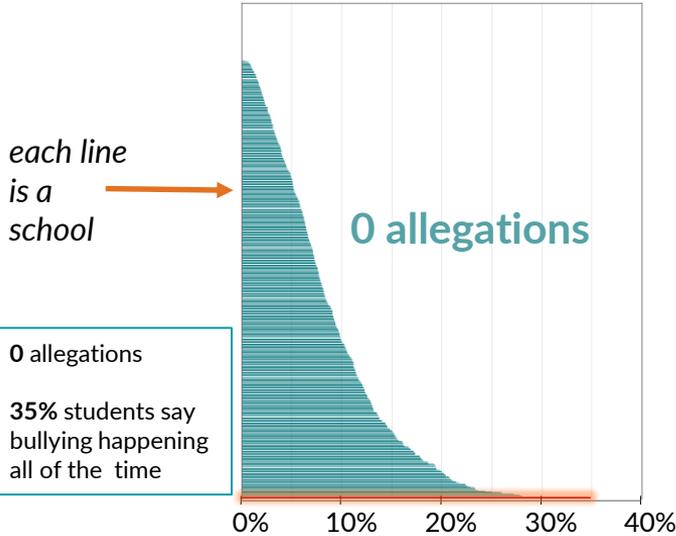
How much do the two surveys (dis)agree?



% of **students** who say bullying/
harassment based on differences
happens **all of the time** in ...

- Schools that reported zero allegations

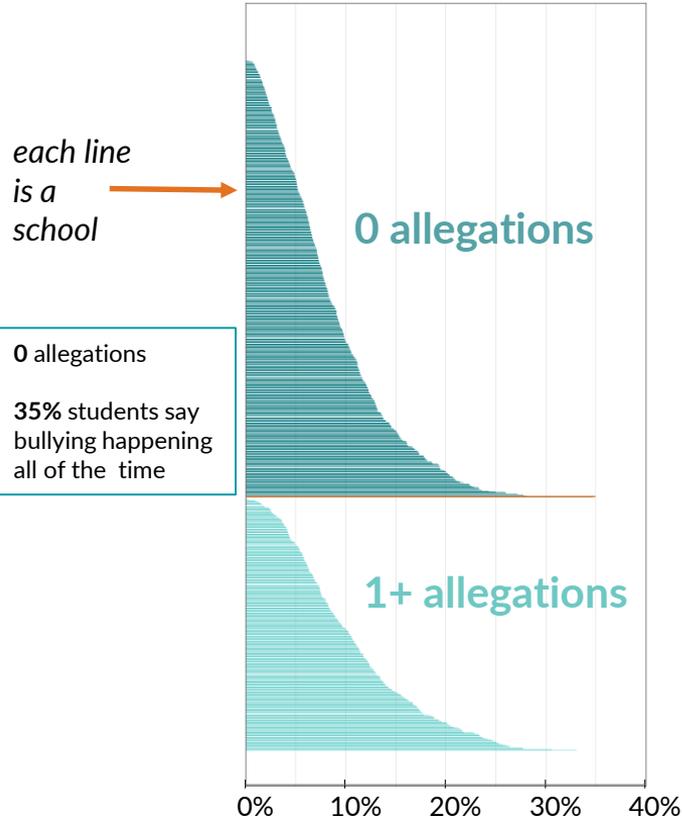
How much do the two surveys (dis)agree?



% of **students** who say bullying/harassment based on differences happens **all of the time** in ...

- Schools that reported **zero** allegations

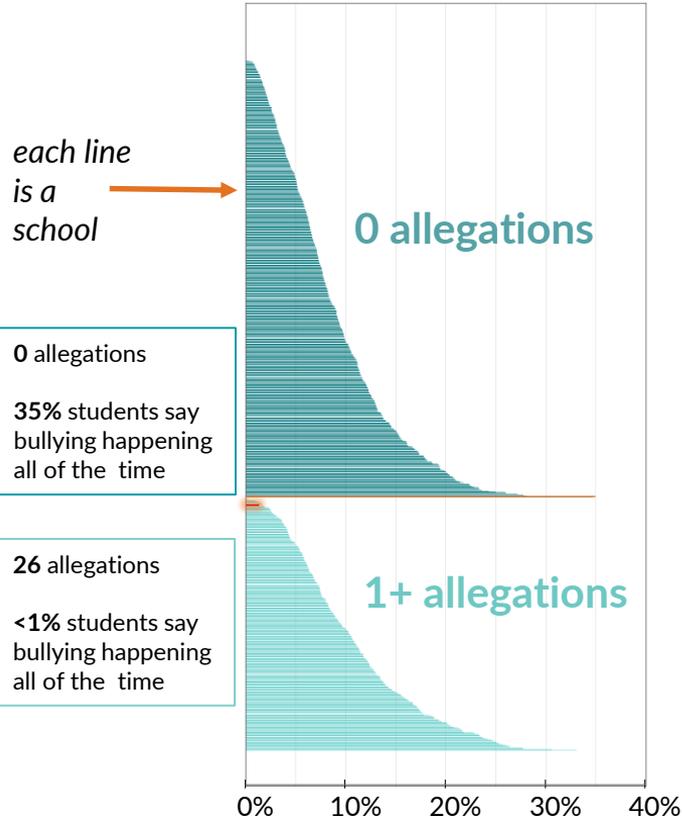
How much do the two surveys (dis)agree?



% of **students** who say bullying/harassment based on differences happens **all of the time** in ...

- Schools that reported **zero** allegations
- Schools that reported **at least 1** allegation

How much do the two surveys (dis)agree?



% of **students** who say bullying/harassment based on differences happens **all of the time** in ...

- Schools that reported **zero** allegations
- Schools that reported **at least 1** allegation

Zones of (dis)agreement



zero
allegations



at least one
allegation

Zones of (dis)agreement

high perception of bullying/
harassment based on differences

zero
allegations

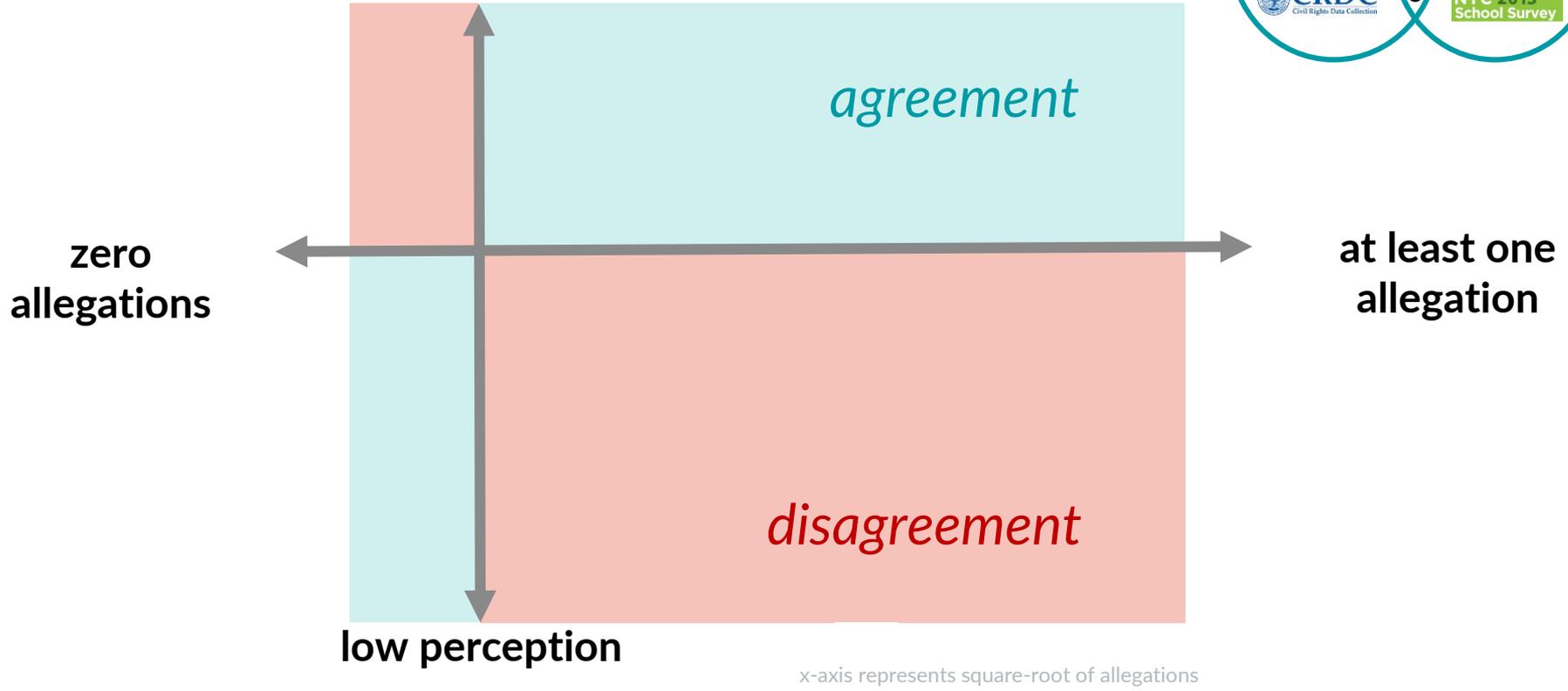
at least one
allegation

low perception



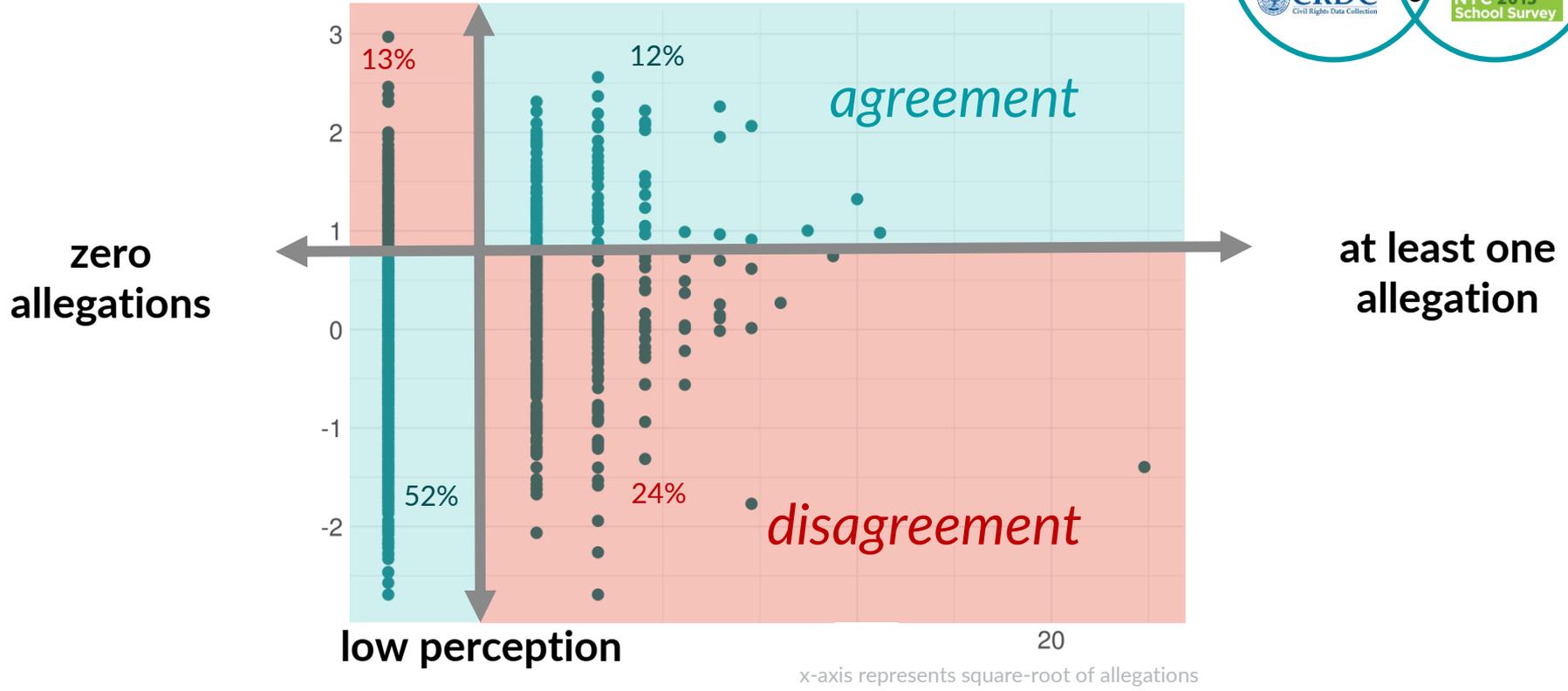
Zones of (dis)agreement

high perception of bullying/
harassment based on differences



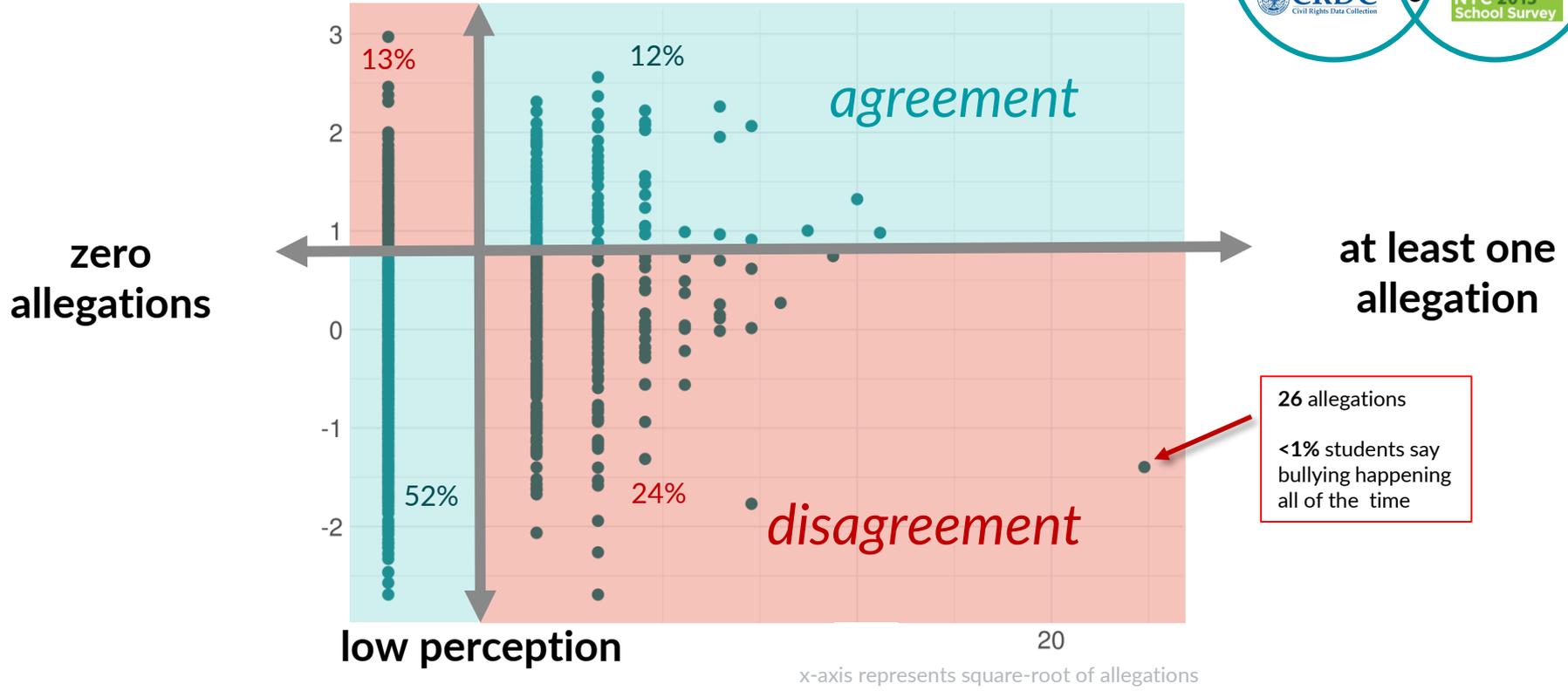
Zones of (dis)agreement

high perception of bullying/
harassment based on differences



Zones of (dis)agreement

high perception of bullying/
harassment based on differences



x-axis represents square-root of allegations

Which factors are associated with survey *dis*agreement?



Parents

- High Schools
vs. Elementary Schools
- High Schools
vs. Middle Schools
- High Schools
vs. Mixed-Grade Schools
- General Education Schools
vs. Career/Technical/etc. Schools
- Regular Schools
vs. Charter Schools
- Total Enrollment
- Female (%)
- Black (%)
- Hispanic (%)
- Students with Disabilities (%)
- English Language Learners (%)
- Free Lunch (%)

Students

- High Schools
vs. Elementary Schools
- High Schools
vs. Middle Schools
- High Schools
vs. Mixed-Grade Schools
- General Education Schools
vs. Career/Technical/etc. Schools
- Regular Schools
vs. Charter Schools
- Total Enrollment
- Female (%)
- Black (%)
- Hispanic (%)
- Students with Disabilities (%)
- English Language Learners (%)
- Free Lunch (%)

Teachers

- High Schools
vs. Elementary Schools
- High Schools
vs. Middle Schools
- High Schools
vs. Mixed-Grade Schools
- General Education Schools
vs. Career/Technical/etc. Schools
- Regular Schools
vs. Charter Schools
- Total Enrollment
- Female (%)
- Black (%)
- Hispanic (%)
- Students with Disabilities (%)
- English Language Learners (%)
- Free Lunch (%)

Which factors are associated with survey *dis*agreement?



Parents

High Schools
vs. Elementary Schools

High Schools
vs. Middle Schools

High Schools
vs. Mixed-Grade Schools

General Education Schools
vs. Career/Technical/etc. Schools

Regular Schools
vs. Charter Schools

Total Enrollment

Female (%)

Black (%)

Hispanic (%)

Students with Disabilities (%)

English Language Learners (%)

Free Lunch (%)

Students

High Schools
vs. Elementary Schools

High Schools
vs. Middle Schools

High Schools
vs. Mixed-Grade Schools

General Education Schools
vs. Career/Technical/etc. Schools

Regular Schools
vs. Charter Schools

Total Enrollment

Female (%)

Black (%)

Hispanic (%)

Students with Disabilities (%)

English Language Learners (%)

Free Lunch (%)

Teachers

High Schools
vs. Elementary Schools

High Schools
vs. Middle Schools

High Schools
vs. Mixed-Grade Schools

General Education Schools
vs. Career/Technical/etc. Schools

Regular Schools
vs. Charter Schools

Total Enrollment

Female (%)

Black (%)

Hispanic (%)

Students with Disabilities (%)

English Language Learners (%)

Free Lunch (%)

positive association with
OCR-NYC School Survey
disagreement

not significantly associated
with OCR-NYC School Survey
(dis)agreement

Which factors are associated with survey *dis*agreement?



Parents

High Schools
vs. Elementary Schools

High Schools
vs. Middle Schools

High Schools
vs. Mixed-Grade Schools

General Education Schools
vs. Career/Technical/etc. Schools

Regular Schools
vs. Charter Schools

Total Enrollment

Female (%)

Black (%)

Hispanic (%)

Students with Disabilities (%)

English Language Learners (%)

Free Lunch (%)

Students

High Schools
vs. Elementary Schools

High Schools
vs. Middle Schools

High Schools
vs. Mixed-Grade Schools

General Education Schools
vs. Career/Technical/etc. Schools

Regular Schools
vs. Charter Schools

Total Enrollment

Female (%)

Black (%)

Hispanic (%)

Students with Disabilities (%)

English Language Learners (%)

Free Lunch (%)

Teachers

High Schools
vs. Elementary Schools

High Schools
vs. Middle Schools

High Schools
vs. Mixed-Grade Schools

General Education Schools
vs. Career/Technical/etc. Schools

Regular Schools
vs. Charter Schools

Total Enrollment

Female (%)

Black (%)

Hispanic (%)

Students with Disabilities (%)

English Language Learners (%)

Free Lunch (%)

positive association with
OCR-NYC School Survey
disagreement
not significantly associated
with OCR-NYC School Survey
(dis)agreement

Caveats and considerations

- Our data is not perfect
 - 2013-14 school year, the latest year available on the Civil Rights Data Collection (Office for Civil Rights)
 - Student sample is smaller than teachers and parents
 - Between-survey (local vs. federal) & within-survey (parent vs. teacher vs. student) comparisons required assumptions
- We only looked at federal and citywide open data, not state
- We are not education experts



The screenshot shows the NYC OpenData website. At the top, it says "NYC OpenData". Below that, there's a section for "2017 NYC School Survey" with a sub-category "Education". The text describes the survey's purpose: "The Survey helps school leaders understand what key members of the school community say about the learning environment at each school. The information captured by the survey is designed to support a dialogue among all members of the school community about how to make the school a better place to learn." There is a "More" link at the bottom of the text.



The screenshot shows the DATA.GOV website. The header includes "DATA.GOV" and navigation links for "DATA", "TOPICS", "IMPACT", "APPLICATIONS", "DEVELOPERS", and "CONTACT". Below the header, there's a "DATA CATALOG" section with "Datasets" and "Organizations" tabs. The main content area shows a dataset titled "Civil Rights Data Collection, 2013-14" with a "Federal" tag. The description states: "The Civil Rights Data Collection, 2013-14 (CRDC 2013-14) is part of the Civil Rights Data Collection (CRDC) program; program data are available beginning with the 2000 collection at. CRDC 2013-14 is a cross-sectional survey that collects data on key education and civil rights issues in the nation's public schools, which include student enrollment and educational programs and services, disaggregated by race/ethnicity, sex, limited English proficiency, and disability. LEAs submit administrative records about schools in the district. CRDC 2013-14 is a universe survey. Key statistics produced from CRDC 2013-14 can provide information about critical civil rights issues as well as contextual information on the state of civil rights in the nation, including enrollment demographics, advanced placement, school discipline, and special education services." The publisher is listed as "Office for Civil Rights".

What we learned about bullying and harassment in NYC schools

- Teachers' responses are less correlated with students' and parents'. This varies by school enrollment.
- In aggregate, perceptions of school bullying/harassment are related to federal reporting of bullying/harassment.
- But perceptions can vary dramatically, even in schools that report zero allegations—which is the vast majority of schools.
- Bigger schools tend to have more disagreement between local and federal school surveys.



What we learned about NYC open data

- Having a “unique” school ID might not be enough ...
 - DBN = “District Borough Number” [Local]
 - ATS = “Automate the Schools” [Local]
 - BEDS = “Basic Education Data System” [State]
 - NCESSCH = “National Center of Education Statistics School ID” [Federal]
 - COMBOKEY = Office for Civil Rights [Federal]



What we learned about NYC open data



- Having a “unique” school ID might not be enough ...

- DBN = “District Borough Number” [Local]
- ATS = “Automate the Schools” [Local]
- BEDS = “Basic Education Data System” [State]
- NCESSCH = “National Center of Education Statistics School ID” [Federal]
- COMBOKEY = Office for Civil Rights [Federal]



What we learned about NYC open data

- Having a “unique” school ID might not be enough ...
 - DBN = “District Borough Number” [Local]
 - ATS = “Automate the Schools” [Local]
 - BEDS = “Basic Education Data System” [State]
 - NCESSCH = “National Center of Education Statistics School ID” [Federal]
 - COMBOKEY = Office for Civil Rights [Federal]



Crosswalks (files that link all the IDs) can help!



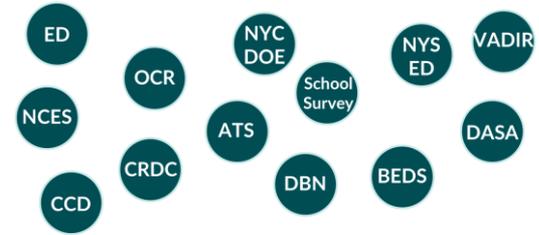
What we learned about NYC open data

- Having a “unique” school ID might not be enough ...
- Being publicly available != being “open”
 - Necessary, but not sufficient

Secure | <https://sunlightfoundation.com/policy/documents/ten-open-data-principles/>



What we learned about NYC open data



- Having a “unique” school ID might not be enough ...
- Being publicly available != being “open”
 - Necessary, but not sufficient

Secure | <https://sunlightfoundation.com/policy/documents/ten-open-data-principles/>

SUNLIGHT FOUNDATION

NYC School Survey

2017 School Survey

For the second year in a row, survey participants, students, and teachers completing the 2017 survey at the citywide level and also summarized and reported.

- View the [2017 citywide analysis of survey results](#). (PDF)
- Download the 2017 [parent](#), [teacher](#), and [student](#) survey data for all New York City public schools. (Excel)
- View the [2017 NYC School Survey Data Protocol](#). (PDF)

7. Commonly owned or open Standards

Commonly owned or open standards refer to who owns the format in which data is stored. For example, if only one company manufactures the program that can read a file where data is stored, access to that information is dependent upon use of the company's processing program. Sometimes that program is unavailable to the public at any cost, or is available, but for a fee. For example, Microsoft Excel is a fairly commonly-used spreadsheet program which costs money to use. Freely available alternative formats often exist by which stored data can be accessed without the need for a software license. Removing this cost makes the data available to a wider pool of potential users.

What we learned about NYC open data



- Having a “unique” school ID might not be enough ...
- Being publicly available != being “open”
- Limitations in the availability of open data may hinder timely analyses

NYC OpenData

2017 NYC School Survey Education

The Survey helps school leaders understand what key stakeholders in the community say about the learning environment at each school. The data captured by the survey is designed to support a dialogue between the school community about how to make the school a better place to learn.

[More](#)

DATA.GOV DATA TOPICS - IMPACT APPLICATIONS DEVELOPERS CONTACT

DATA CATALOG / Datasets Organizations ?

Department of Education / Office for Civil Rights Submit Data Story Report Data Issue

ED.gov Federal

Civil Rights Data Collection, 2013-14
Metadata Updated: February 4, 2018

on, 2013-14 (CRDC 2013-14) is part of the Civil Rights Data Collection (CRDC) available beginning with the 2000 collection at . CRDC 2013-14 is a cross-district data on key education and civil rights issues in the nation's public schools, which includes data on educational programs and services, disaggregated by race/ethnicity, sex, and disability. LEAs submit administrative records about schools in the district survey. Key statistics produced from CRDC 2013-14 can provide information on school discipline, advanced placement, school discipline, and special education services, as well as contextual information on the state of civil rights in the nation, such as school discipline, advanced placement, school discipline, and special education services.

3. Timeliness

Datasets released by the government should be available to the public in a timely fashion. Whenever feasible, information collected by the government should be released as quickly as it is gathered and collected. Priority should be given to data whose utility is time sensitive. Real-time information updates would maximize the utility the public can obtain from this information.

What we learned about NYC open data

- Having a “unique” school ID might not be enough ...
- Being publicly available != being “open”
- Limitations in the availability of open data may hinder timely analyses
- Understanding the original purpose of different data sets can guide research



What we learned about NYC open data



- Having a “unique” **school ID** might not be enough ...
- Being publicly available != being “**open**”
- Limitations in the availability of open data may hinder **timely** analyses
- Understanding the **original purpose** of different data sets can guide research
- Someone else probably has (or had) the same data problem as you!



What we learned about NYC open data



- Having a “unique” **school ID** might not be enough ...
- Being publicly available != being “open”
- Limitations in the availability of open data may hinder **timely** analyses
- Understanding the **original purpose** of different data sets can guide research
- Someone else probably has (or had) the same data problem as you!

The Research Alliance for
New York City Schools

About Us Our Team Research News Contact

- **Identifiers that link** school records over time and that link New York City data to other public sources such as those available from the New York State Education Department, the Census Bureau, and the U.S. Department of Education

What we learned about NYC open data



- Having a “unique” **school ID** might not be enough ...
- Being publicly available != being “open”
- Limitations in the availability of open data may hinder **timely** analyses
- Understanding the **original purpose** of different data sets can guide research
- Someone else probably has (or had) the same data problem as you!



- **Identifiers that link** school records over time

Education civil rights data: The good, the bad, the dirty data to other

2018 CHICAGO MARCH 8-11

CAR CONFERENCE

Event: 2018 CAR Conference
Speakers: Jennifer LaFleur, Alex Har
Date/Time: Thursday, March 8 at 10:
Location: Addison
Audio file: No audio file available.

Panelists with a range of expertise will discuss the **federal civil rights dataset**, how to make the most of it and avoid the pitfalls. We also will talk about how to use **state data** to tell stories about disparities in schools. This session will provide lots of story

ent,
U.S.



John Paul Farmer ✓
@johnpaulfarmer

Following



"#OpenData may be the Central Park of data"
— an amazing public resource that we
shouldn't take for granted. - @MiguelGamino
at #nycSoData on #OpenDataDay



11:16 AM - 3 Mar 2018 from [Manhattan, NY](#)

www.twosigma.com/dataclinic

Rachael Weiss Riley & Christine Zhang
Dave Santin
Tiffany Chang
Chris Mulligan

Special thanks to:

Ben Wellington (Two Sigma), Erin Stein (Overdeck Family Foundation)
Matt Chingos & Erica Blom (Urban Institute)
Adrienne Schmoeker (NYC MODA)
Nick Chmura & Two Sigma Q4 2017 Hack Day participants

NYC OpenData



Julia Bloom-Weltman (julia.bloom-weltman@aemcorp.com)

Director of Research
Applied Engineering Management



Meghan McCormick (meghan.mccormick@mdrc.org)

Research Associate
MDRC



Johanna Miller (@johannawithanh)

Advocacy Director
New York Civil Liberties Union



Jasmine Soltani (jasmine.soltani@nyu.edu)

Data Manager
Research Alliance for NYC Schools



Amy Zimmer (@the_zim | amyz@localize.city)

Journalist, Content Strategist
Localize.city