PROJECT 3: BEHAVIOR OF SCHOOL METRICS AND CRIME RATES IN PHILADELPHIA

Presented by:
Elizabeth Vandergrift
Juan Camilo Bohorquez Rozo
Eric Croston

THE DATA

PhilaSD.org

- SPREE 2023 School Performance

OpenDataPhilly.org

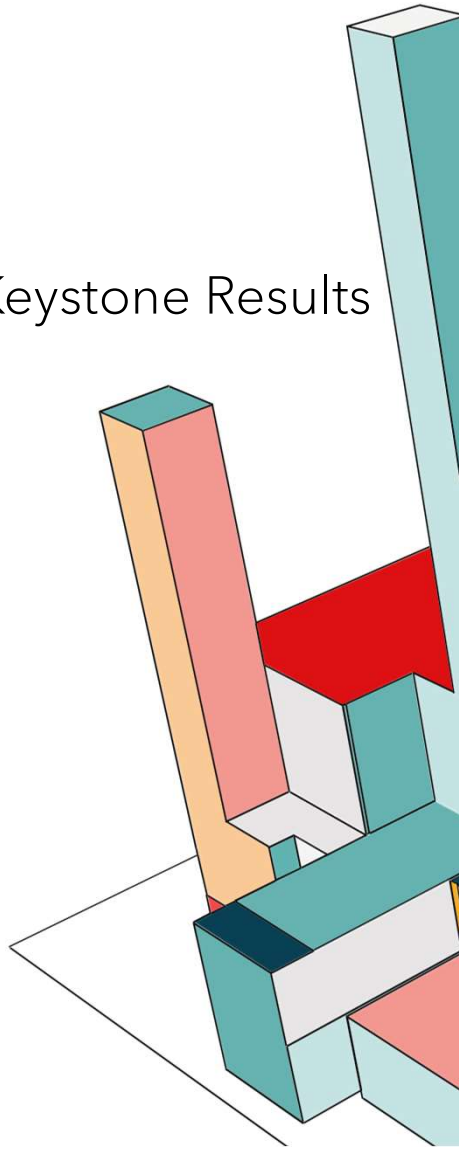
- Police Stations
- Schools
- Zip Code Boundaries
- PSA Boundaries
- Crime Incidents

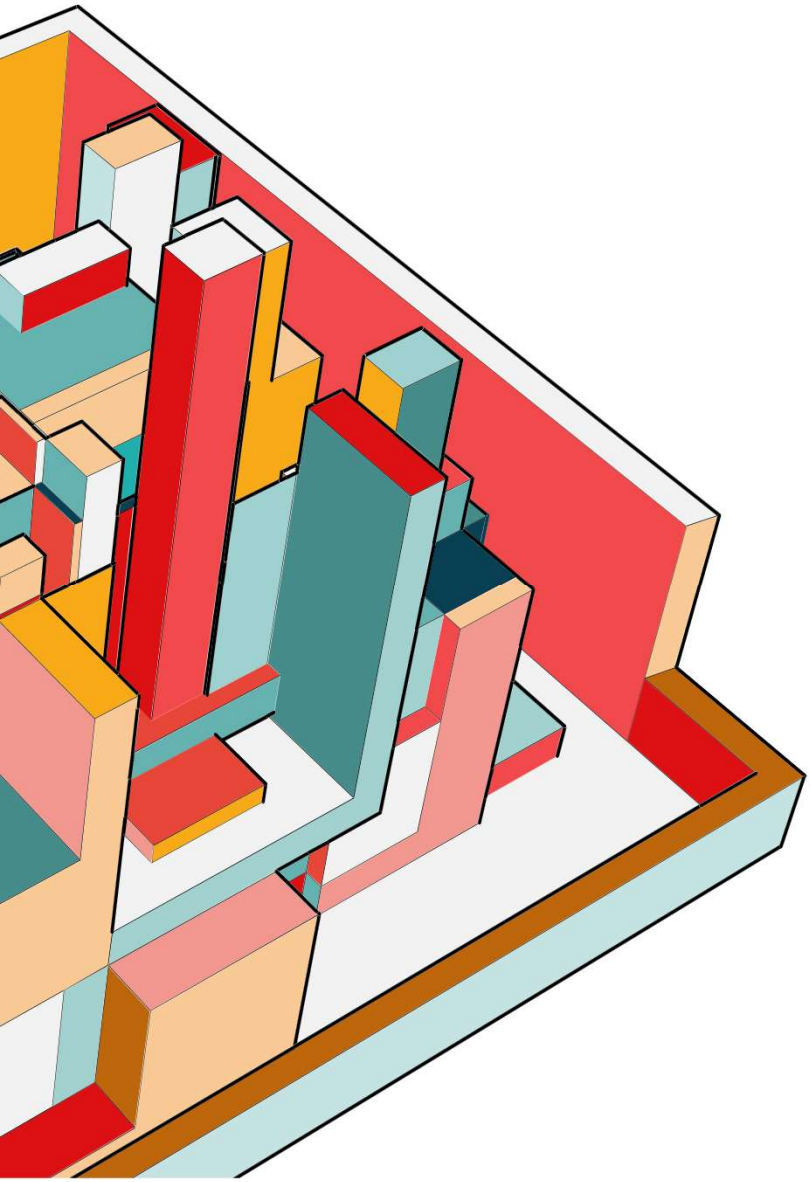
SCHOOL METRICS

School Report on Education and Equity (SPREE)

Example Metrics: Attendance, Graduation Rates and PSSA/Keystone Results

- 294 School IDs and 323 Total Schools
- 260 Metrics per School
- 4 Report/School Types: High School, Middle School, K-8 Schools and Elementary Schools
- 84 High Schools
- 29 Middle Schools
- 154 K-8 Schools
- 56 Elementary Schools





SCHOOL METRICS

Data Cleaning and Prep for Database Import

- Create a dataframe for each report/school type
- Create lists of metrics to keep for analysis
- Reduce the number of rows per school
- Pivot the dataframes to turn the duplicate school rows into columns
- Rename the pivoted columns
- Change the data types for analysis
- Export CSV file to import into our database

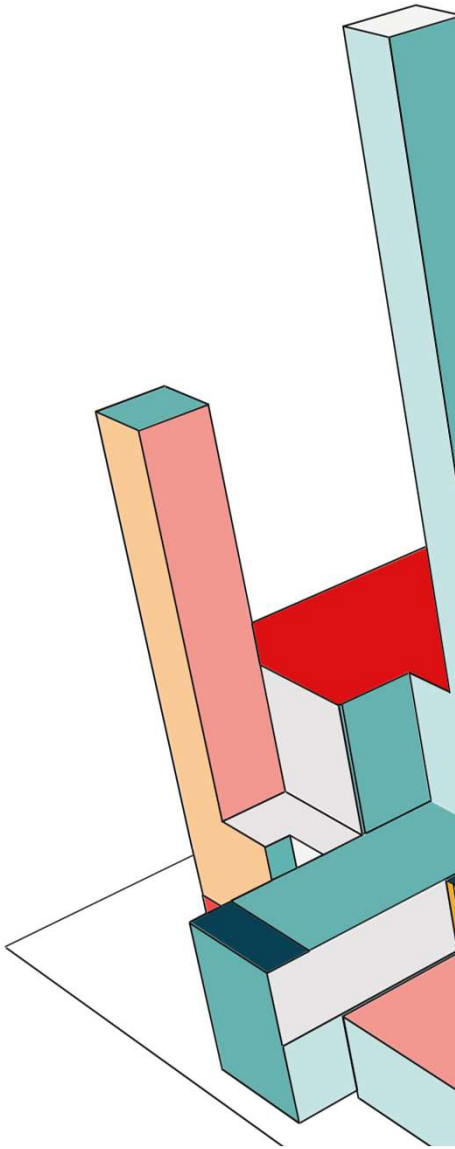
CRIME INCIDENTS

There were 85,529 Crime Incidents Documented by the Philadelphia Police Department From January 1st, 2024 until July 17th, 2024.

For Each Crime there is a Numerical Code and Incident Description.

There are 26 different unique incident types with specific codes.

	UCR General Code	Incidents
0	100	[Homicide - Criminal, Homicide - Justifiable]
1	200	[Rape]
2	300	[Robbery Firearm, Robbery No Firearm]
3	400	[Aggravated Assault No Firearm, Aggravated Ass...
4	500	[Burglary Non-Residential, Burglary Residential]
5	600	[Thefts, Theft from Vehicle]
6	700	[Motor Vehicle Theft]
7	800	[Other Assaults]
8	900	[Arson]
9	1000	[Forgery and Counterfeiting]
10	1100	[Fraud]
11	1200	[Embezzlement]
12	1300	[Receiving Stolen Property]
13	1400	[Vandalism/Criminal Mischief]
14	1500	[Weapon Violations]
15	1600	[Prostitution and Commercialized Vice]
16	1700	[Other Sex Offenses (Not Commercialized)]
17	1800	[Narcotic / Drug Law Violations]
18	1900	[Gambling Violations]
19	2000	[Offenses Against Family and Children]
20	2100	[DRIVING UNDER THE INFLUENCE]
21	2200	[Liquor Law Violations]
22	2300	[Public Drunkenness]
23	2400	[Disorderly Conduct]
24	2500	[Vagrancy/Loitering]
25	2600	[All Other Offenses]



CRIME INCIDENTS

UCR_Code	General_Crime_Category	Crime_Group
100	[Homicide - Criminal, Homicide - Justifiable]	[Violent Crimes]
200	[Rape]	[Violent Crimes]
300	[Robbery Firearm, Robbery No Firearm]	[Violent Crimes]
400	[Aggravated Assault No Firearm, Aggravated Ass...	[Violent Crimes]
500	[Burglary Non-Residential, Burglary Residential]	[Property Crimes]
600	[Thefts, Theft from Vehicle]	[Property Crimes]
700	[Motor Vehicle Theft]	[Property Crimes]
800	[Other Assaults]	[Violent Crimes]
900	[Arson]	[Property Crimes]
1000	[Forgery and Counterfeiting]	[Property Crimes]
1100	[Fraud]	[Other Crimes]
1200	[Embezzlement]	[Other Crimes]
1300	[Receiving Stolen Property]	[Property Crimes]
1400	[Vandalism/Criminal Mischief]	[Property Crimes]
1500	[Weapon Violations]	[Other Crimes]
1600	[Prostitution and Commercialized Vice]	[Drug and Vice Crimes]
1700	[Other Sex Offenses (Not Commercialized)]	[Drug and Vice Crimes]
1800	[Narcotic / Drug Law Violations]	[Drug and Vice Crimes]
1900	[Gambling Violations]	[Drug and Vice Crimes]
2000	[Offenses Against Family and Children]	[Other Crimes]
2100	[DRIVING UNDER THE INFLUENCE]	[Public Order Crimes]
2200	[Liquor Law Violations]	[Public Order Crimes]
2300	[Public Drunkenness]	[Public Order Crimes]
2400	[Disorderly Conduct]	[Public Order Crimes]
2500	[Vagrancy/Loitering]	[Public Order Crimes]
2600	[All Other Offenses]	[Other Crimes]

Crime Incident Types were grouped into related crime categories to be able to make comparisons.

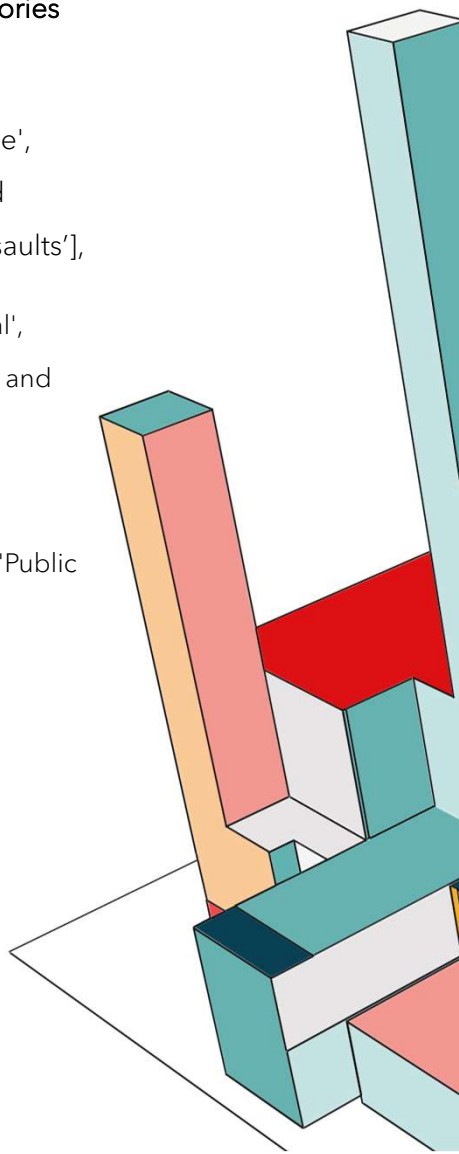
'Violent Crimes': ['Homicide - Criminal', 'Homicide - Justifiable', 'Rape', 'Robbery No Firearm', 'Robbery Firearm', 'Aggravated Assault Firearm', 'Aggravated Assault No Firearm', 'Other Assaults'],

'Property Crimes': ['Burglary Residential', 'Burglary Non-Residential', 'Thefts', 'Theft from Vehicle', 'Motor Vehicle Theft', 'Arson', 'Forgery and Counterfeiting', 'Receiving Stolen Property', 'Vandalism/Criminal Mischief'],

Public Order Crimes': ['Disorderly Conduct', 'Vagrancy/Loitering', 'Public Drunkenness', 'Liquor Law Violations', 'DRIVING UNDER THE INFLUENCE'],

'Drug and Vice Crimes': ['Prostitution and Commercialized Vice', 'Narcotic / Drug Law Violations', 'Other Sex Offenses (Not Commercialized)', 'Gambling Violations'],

'Other Crimes': ['Offenses Against Family and Children', 'Fraud', 'Embezzlement', 'Weapon Violations', 'All Other Offenses'



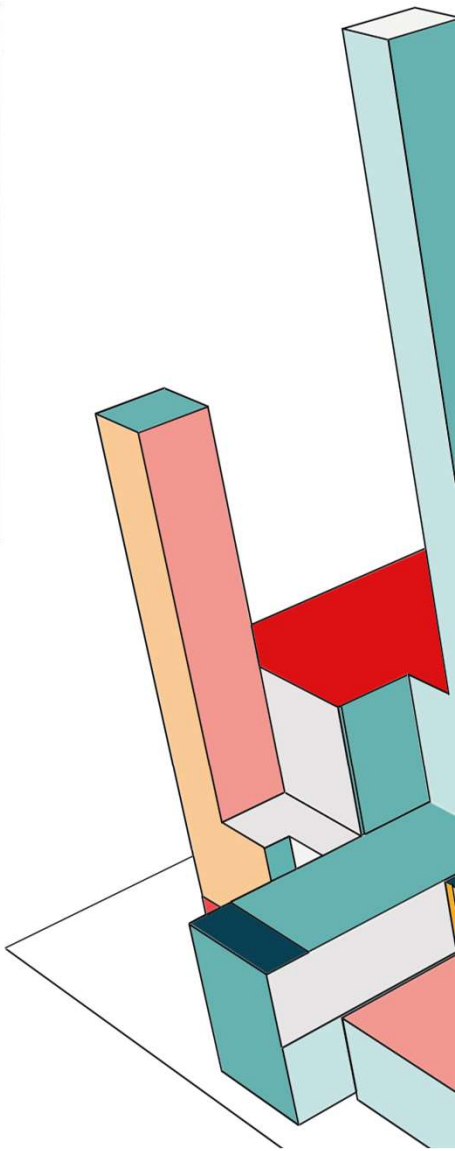
CRIME INCIDENTS

Creating a composite key for District PSA including the District and the Police Service Area number within that District

District PSA composite added to Police Station Location and Police Incident locations

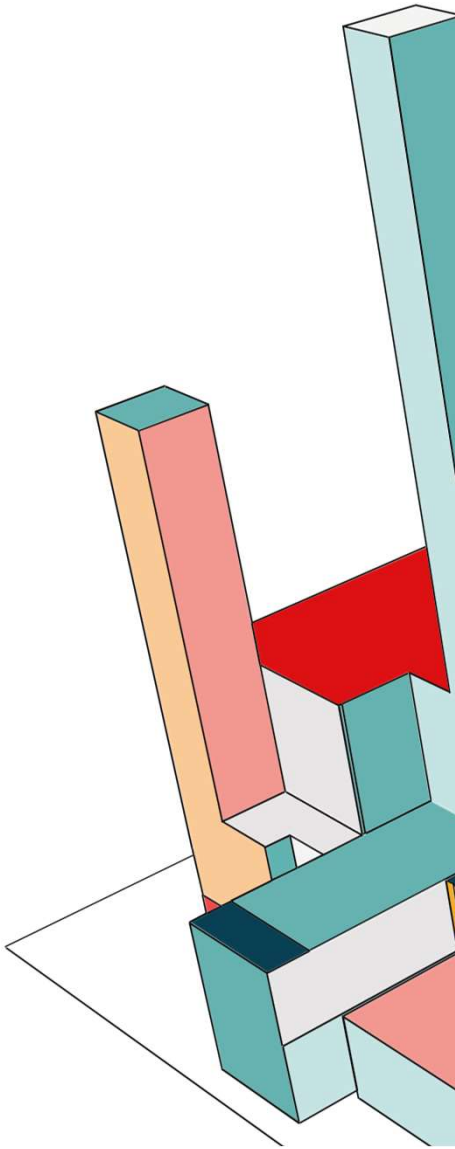
District	Police_Service_Area	District_PSA
18	3	183
17	1	171
5	1	51
3	3	33
15	2	152
...	...	...
24	1	241
22	4	224
35	3	353
18	3	183
8	2	82

District_PSA	District	Police_Service_Area	Police_Location_ID	Location	ZIP_CODE	Telephone_Number	Geometry
11	1	1	21	24th St & Wolf St	19145	686-3010	POINT (-75.18693 39.92407)
12	1	2	21	24th St & Wolf St	19145	686-3010	POINT (-75.18693 39.92407)
21	2	1	4	Harbison Ave & Levick St	19149	686-3020	POINT (-75.06325 40.03149)
22	2	2	4	Harbison Ave & Levick St	19149	686-3020	POINT (-75.06325 40.03149)
23	2	3	4	Harbison Ave & Levick St	19149	686-3020	POINT (-75.06325 40.03149)
...	...	...	...	...	...	...	...
352	35	2	5	N Broad St & Champlost St	19141	686-3350	POINT (-75.14365 40.04450)
353	35	3	5	N Broad St & Champlost St	19141	686-3350	POINT (-75.14365 40.04450)
391	39	1	8	22nd St & Hunting Park Ave	19140	686-3390	POINT (-75.16425 40.01090)
392	39	2	8	22nd St & Hunting Park Ave	19140	686-3390	POINT (-75.16425 40.01090)
393	39	3	8	22nd St & Hunting Park Ave	19140	686-3390	POINT (-75.16425 40.01090)



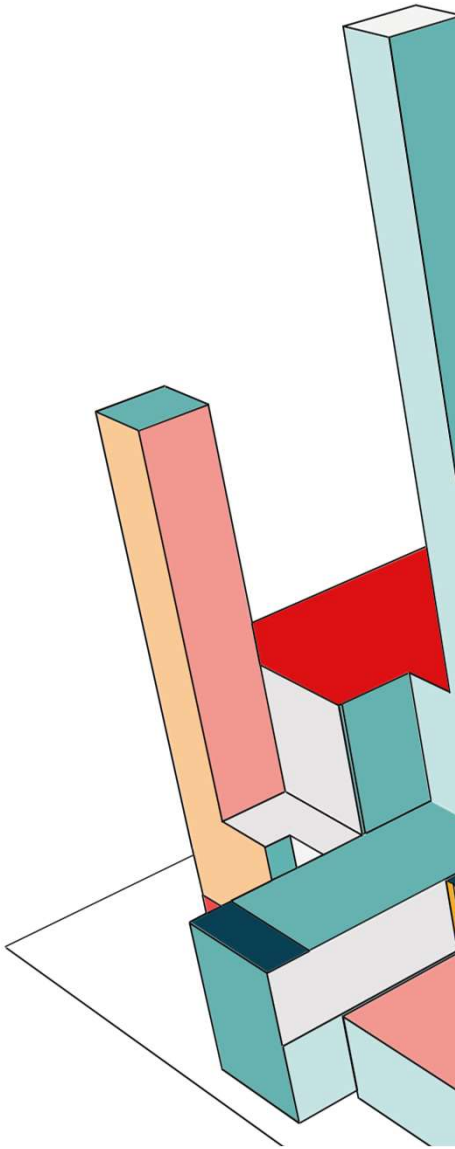
HOW ARE WE GOING TO CONNECT OVER 85K INCIDENT LOCATIONS WITH SCHOOLS AND THEIR METRICS????

Our plan was to use ZIP Codes….



BUT THAT DID NOT WORK....

- Only some files had ZIP Codes, and the largest file with Police Incidents (85K rows) did not!
- We also found looping through 85K rows of data to determine if a school point is within the same polygon Zip Code of a crime was not easily possible.
- Possibly, we could find the common ZIP Codes by using the Police Station Coordinates and School Coordinates in the ZIP Code?
- However, this does not account for Crimes handled by the Police District in Different ZIP Codes. ZIP Codes and Police Districts- Police Service Area are not aligned.
- Then we found another plan….



POLICE SERVICE AREA BOUNDARIES GEOJSON FILE AND SCHOOL COORDINATES


```
{
"type": "FeatureCollection",
"name": "Boundaries_PSA",
"crs": { "type": "name", "properties": { "name": "urn:ogc:def:crs:OGC:1.3:CRS84" } },
"features": [
{ "type": "Feature", "properties": { "OBJECTID": 1, "AREA": 136213942.09201899, "PERIMETER": 56995.903427600002, "PSACOV_": null, "PSACOV_ID": null, "ID": 16, "DISTRICT__": "08", "PSA_NUM": "081", "
{ "type": "Feature", "properties": { "OBJECTID": 2, "AREA": 168027734.49678701, "PERIMETER": 66474.245601000002, "PSACOV_": null, "PSACOV_ID": null, "ID": 17, "DISTRICT__": "08", "PSA_NUM": "082", "
{ "type": "Feature", "properties": { "OBJECTID": 3, "AREA": 169959461.95717099, "PERIMETER": 66663.49849, "PSACOV_": null, "PSACOV_ID": null, "ID": 18, "DISTRICT__": "08", "PSA_NUM": "083", "OLD_SEC
{ "type": "Feature", "properties": { "OBJECTID": 4, "AREA": 32508297.425106101, "PERIMETER": 15689.0698557, "PSACOV_": null, "PSACOV_ID": null, "ID": 19, "DISTRICT__": "09", "PSA_NUM": "091", "OLD_S
{ "type": "Feature", "properties": { "OBJECTID": 5, "AREA": 29214097.941779301, "PERIMETER": 15689.0698557, "PSACOV_": null, "PSACOV_ID": null, "ID": 19, "DISTRICT__": "09", "PSA_NUM": "092", "OLD_S
{ "type": "Feature", "properties": { "OBJECTID": 6, "AREA": 19008246.7181146, "PERIMETER": 15689.0698557, "PSACOV_": null, "PSACOV_ID": null, "ID": 19, "DISTRICT__": "09", "PSA_NUM": "093", "OLD_SEC
```

```
"coordinates": [ [ [ -75.008288806677399, 40.052724593837198 ], [ -75.007888777577605, 40.052390318161898 ], [ -75.006161033850304, 40.053722286591501 ], [ -75.004797266474597, 40.054756771948497 ]
rdinates": [ [ [ -74.998757023896999, 40.081329577131797 ], [ -74.997236249321006, 40.080032169182502 ], [ -74.995804752140899, 40.080854128908697 ], [ -74.9943506036728, 40.081690977236697 ], [ -74
es": [ [ [ -74.975672816224005, 40.120198122085903 ], [ -74.975567719945602, 40.120179624105198 ], [ -74.975391210557007, 40.120197133386498 ], [ -74.975227563188099, 40.120223658870501 ], [ -74.975
ates": [ [ [ -75.159704916655201, 39.970211165970099 ], [ -75.159978922015, 39.969004349499599 ], [ -75.160074982712501, 39.968608897080699 ], [ -75.160161815143795, 39.968186373384 ], [ -75.160215
ates": [ [ [ -75.135237801501304, 39.950695126877399 ], [ -75.135237712268093, 39.950695742045603 ], [ -75.134608455059904, 39.952333445575697 ], [ -75.133721198377501, 39.9545774237411 ], [ -75.133
ates": [ [ [ -75.167164573243198, 39.958701224529896 ], [ -75.167991700952797, 39.955018015857597 ], [ -75.168769883043396, 39.951477074693898 ], [ -75.170106921668904, 39.945276622417801 ], [ -75.1
```


```
DataFrame saved to new_geo_schools_df.csv
```

OBJECTID	school_id	SCHOOL_NAME	SCHOOL_NAME_LABEL	STREET_ADDRESS	ZIP_CODE	PHONE_NUMBER	GRADE_LEVEL	GRADE_ORG	TYPE	TYPE_SPECIFIC	geometry	District	District_PSA
70	2670	CONSTITUTION HIGH SCHOOL	CONSTITUTION HIGH SCHOOL	18 S 7TH ST	19106	(215) 400-7850	HIGH SCHOOL	9-12	1	DISTRICT	POINT (-75.15234 39.95025)	09	095
71	5230	CONWELL, RUSSELL MIDDLE SCHOOL	RUSSELL H. CONWELL SCHOOL	1849 E CLEARFIELD ST	19134	(215) 400-7210	MIDDLE SCHOOL	5-8	1	DISTRICT	POINT (-75.11482 39.99442)	24	242

CONNECTING SCHOOL LOCATIONS AND POLICE SERVICE AREAS INCIDENT LOCATIONS USING A COMBINATION KEY AND POINTS WITHIN THE SERVICE AREA POLYGON

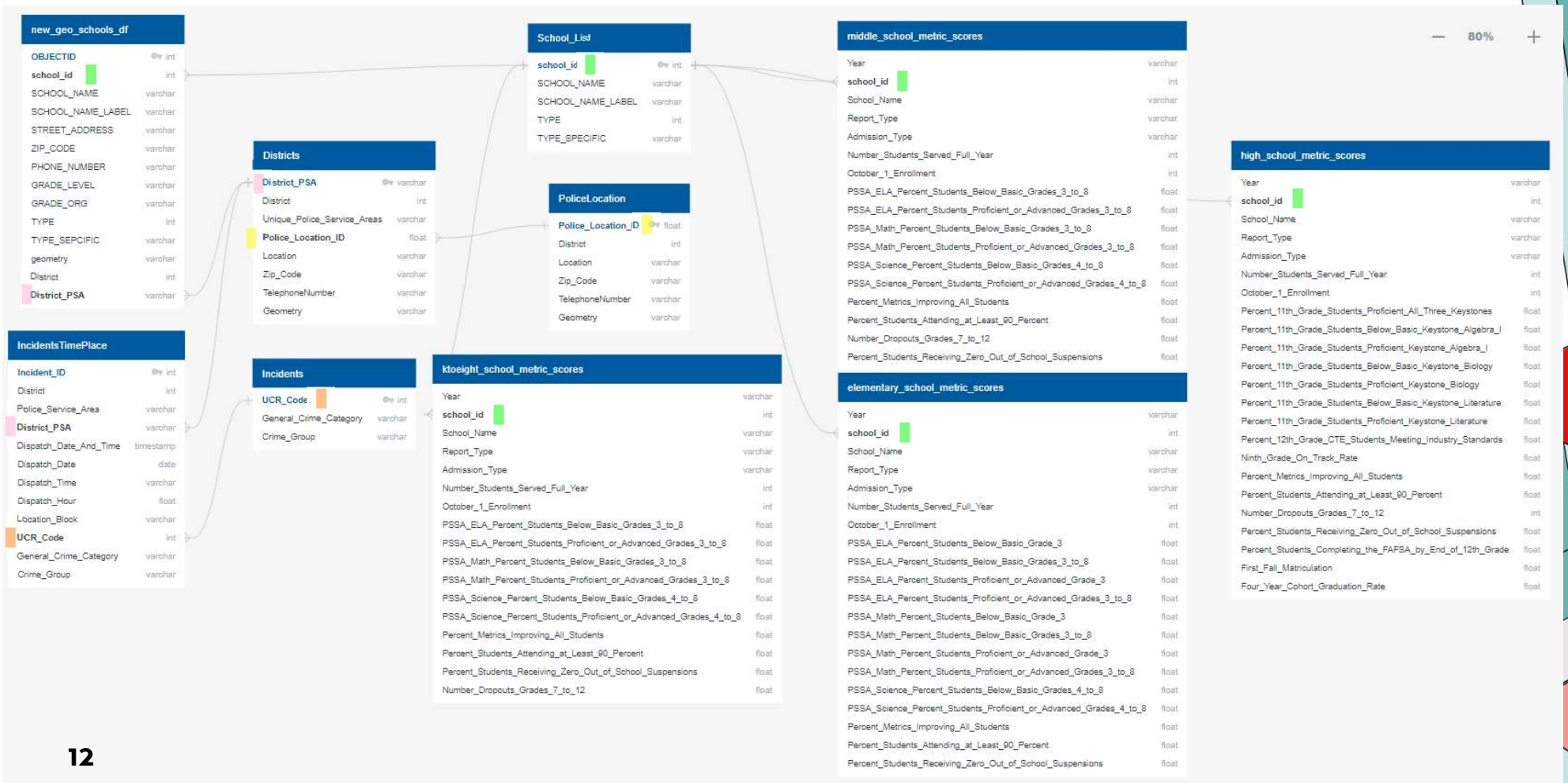
Using Two smaller datasets to be able to connect the larger data sets

Incident_ID	District	Police_Service_Area	District_PSA	Dispatch_Date_and_Time	Dispatch_Date	Dispatch_Time	Dispatch_Hour	Location_Block	UCR_Code	General_Crime_Categ
13965196	18	3	183	2024-02-02 16:30:00+00:00	2024-02-02	11:30:00	11.0	300 BLOCK S 34TH ST	600	Th
17545922	17	1	171	2024-06-15 12:10:00+00:00	2024-06-15	08:10:00	8.0	2100 BLOCK SOUTH ST	600	Th
13065664	5	1	51	2024-01-06 20:35:00+00:00	2024-01-06	15:35:00	15.0	300 BLOCK ROCHELLE AV	600	Th

```
DataFrame saved to new_geo_schools_df.csv
```

OBJECTID	school_id	SCHOOL_NAME	SCHOOL_NAME_LABEL	STREET_ADDRESS	ZIP_CODE	PHONE_NUMBER	GRADE_LEVEL	GRADE_ORG	TYPE	TYPE_SPECIFIC	geometry	District	District_PSA
70	2670	CONSTITUTION HIGH SCHOOL	CONSTITUTION HIGH SCHOOL	18 S 7TH ST	19106	(215) 400-7850	HIGH SCHOOL	9-12	1	DISTRICT	POINT (-75.15234 39.95025)	09	095
71	5230	CONWELL, RUSSELL MIDDLE SCHOOL	RUSSELL H. CONWELL SCHOOL	1849 E CLEARFIELD ST	19134	(215) 400-7210	MIDDLE SCHOOL	5-8	1	DISTRICT	POINT (-75.11482 39.99442)	24	242

Entity Relationship Diagram (ERD)



CRIME INCIDENTS LOCATIONS AND SCHOOL LOCATIONS

```
14
15  SELECT * FROM "IncidentsTimePlace"
16
```

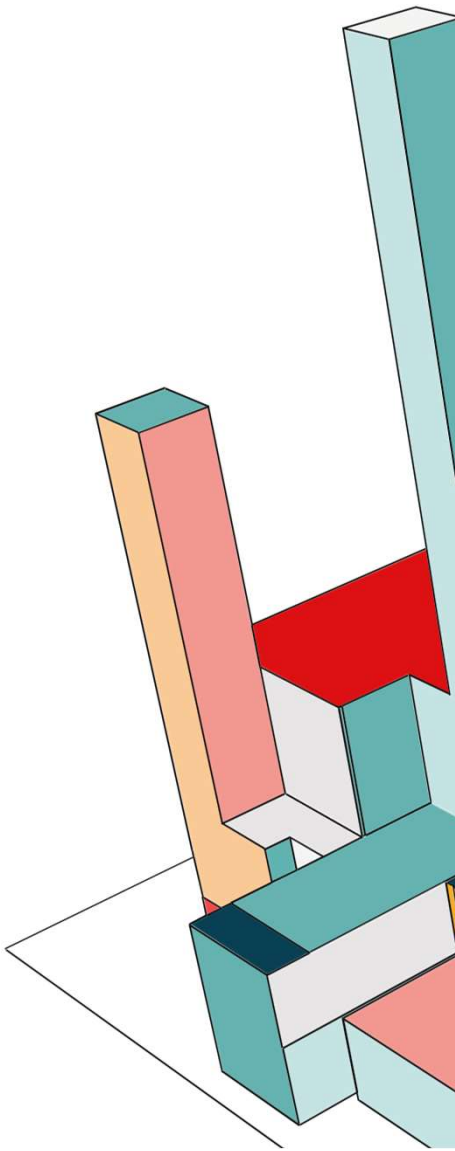
Data Output | Messages | Notifications

	Incident_ID [PK] integer	District integer	Police_Service_Area character varying	District_PSA character varying	Dispatch_Date_And_Time timestamp without time zone	Dispatch_Date date	Dispatch_Time character varying	Dispatch_Hour double precision	Location_Block character varying	UCR_C integer
1	13965196	18	3	183	2024-02-02 16:30:00	2024-02-02	11:30:00	11	300 BLOCK S 34TH ST	
2	17545922	17	1	171	2024-06-15 12:10:00	2024-06-15	08:10:00	8	2100 BLOCK SOUTH ST	
3	13065664	5	1	51	2024-01-06 20:35:00	2024-01-06	15:35:00	15	300 BLOCK ROCHELLE AV	
4	15737696	3	3	33	2024-04-26 21:13:00	2024-04-26	17:13:00	17	1100 BLOCK ELLSWORTH ST	
5	14865984	15	2	152	2024-03-15 11:17:00	2024-03-15	07:17:00	7	5500 BLOCK TULIP ST	
6	15983040	9	2	92	2024-05-03 12:39:00	2024-05-03	08:39:00	8	N 2ND ST & SPRING GARDEN ST	
7	16146833	24	1	241	2024-04-24 19:27:00	2024-04-24	15:27:00	15	3900 BLOCK CASTOR AV	
8	13376825	9	3	93	2024-02-05 16:23:00	2024-02-05	11:23:00	11	600 BLOCK N BROAD ST	
9	15985222	19	3	193	2024-05-03 21:34:00	2024-05-03	17:34:00	17	1500 BLOCK N 52ND ST	
10	13129838	15	1	151	2024-01-29 16:39:00	2024-01-29	11:39:00	11	4600 BLOCK PAUL ST	

```
11  SELECT * FROM new_geo_schools_df
12
13  SELECT * FROM "PoliceLocation"
14
```

Data Output | Messages | Notifications

	ZIP_CODE character varying	PHONE_NUMBER character varying	GRADE_LEVEL character varying	GRADE_ORG character varying	TYPE integer	TYPE_SEPCIFIC character varying	geometry character varying	District integer	District_PSA character varying
1	19106	(215) 400-7850	HIGH SCHOOL	12-Sep	1	DISTRICT	POINT (-75.1523359999574 39.9502520000276)	9	95
2	19134	(215) 400-7210	MIDDLE SCHOOL	8-May	1	DISTRICT	POINT (-75.1148200223767 39.9944190479345)	24	242
3	19141	(215) 400-8330	ELEMENTARY/MIDDLE	K-8	1	DISTRICT	POINT (-75.1460740001465 40.0255260001875)	35	352
4	19128	(215) 400-3430	ELEMENTARY/MIDDLE	K-8	1	DISTRICT	POINT (-75.2068019997144 40.0224720000888)	5	51
5	19140	(215) 400-3860	ELEMENTARY SCHOOL	K-5	1	DISTRICT	POINT (-75.12993625036739 40.0026291206135)	25	253
6	19144	(267) 270-4665	MIDDLE/HIGH	12-Jun	3	PRIVATE	POINT (-75.1764454724389 40.0336765476524)	14	143
7	19140	(215) 400-3890	ELEMENTARY/MIDDLE	K-8	1	DISTRICT	POINT (-75.1481843568952 40.0036406776715)	25	254
8	19121	(215) 400-7330	ELEMENTARY/MIDDLE	K-8	1	DISTRICT	POINT (-75.1831230017859 39.9853002642558)	22	222
9	19131	(215) 400-7280	ELEMENTARY/MIDDLE	K-8	1	DISTRICT	POINT (-75.2151434499055 39.9726290989846)	16	162
10	19123	(215) 400-7630	HIGH SCHOOL	12-Sep	1	DISTRICT	POINT (-75.1431779114975 39.9680504922372)	26	262
11	19145	(215) 400-8210	ELEMENTARY/MIDDLE	K-8	1	DISTRICT	POINT (-75.17670508385341 39.9145523452285)	1	12
12	19137	(215) 400-7240	ELEMENTARY/MIDDLE	K-8	1	DISTRICT	POINT (-75.06666516407 40.0002731480946)	15	151
13	19125	(215) 400-7490	ELEMENTARY/MIDDLE	K-8	1	DISTRICT	POINT (-75.12735288860161 39.9866007655954)	26	263
14	19136	(215) 400-3320	ELEMENTARY/MIDDLE	K-8	1	DISTRICT	POINT (-75.02462520159639 40.0434625873175)	8	81
15	19143	(215) 400-7910	ELEMENTARY/MIDDLE	K-8	1	DISTRICT	POINT (-75.2435629997548 39.9527911813642)	18	181



CRIME INCIDENTS AND SCHOOL METRICS

Query | Query History

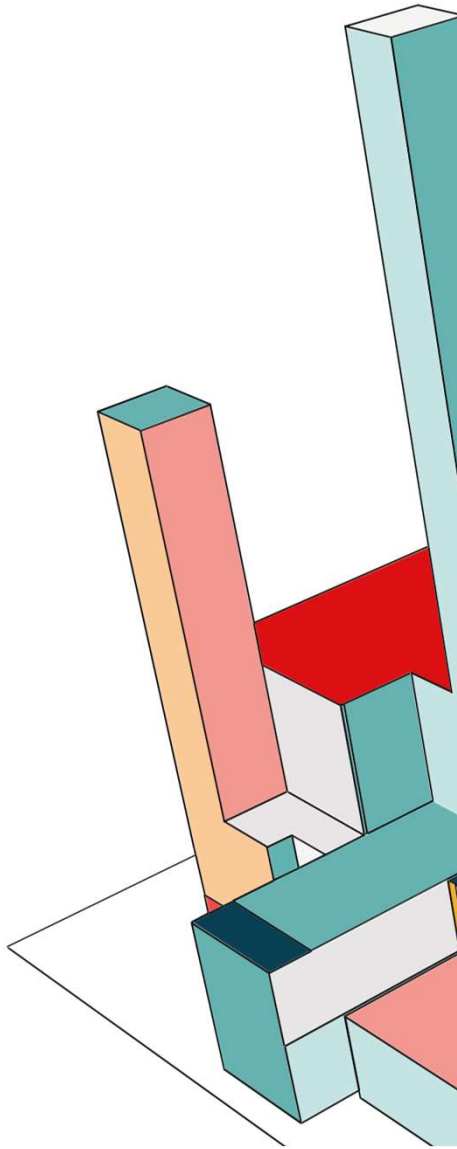
```
---15. Total Crimes and Average of elementary School metrics

SELECT
    ng."District_PSA",
    AVG(es."PSSA_ELA_Percent_Students_Proficient_or_Advanced_Grades_3_to_8") AS "Average_ELA_Proficiency_3_to_8
    AVG(es."PSSA_Math_Percent_Students_Proficient_or_Advanced_Grades_3_to_8") AS "Average_Math_Proficiency_3_to
    AVG(es."Percent_Students_Attending_at_Least_90_Percent") AS "Average_Percent_Attending_90",
    b."Total_Crimes"
FROM
    "new_geo_schools_df" ng
JOIN
    "elementary_school_metric_scores" es ON ng."school_id" = es."school_id"
JOIN
    (
        SELECT
            itp."District_PSA",
            COUNT(*) AS "Total_Crimes"
        FROM
            "IncidentsTimePlace" itp
        GROUP BY
            itp."District_PSA"
    ) b ON ng."District_PSA" = b."District_PSA"

GROUP BY
    ng."District_PSA", b."Total_Crimes"
ORDER BY
    ng."District_PSA";
```

Data Output | Messages | Notifications

	District_PSA character varying	Average_ELA_Proficiency_3_to_8 double precision	Average_Math_Proficiency_3_to_8 double precision	Average_Percent_Attending_90 double precision	Total_Crimes bigint
1	11	34.7	28	54.2	1029
2	121	22.8	14.2	57.6	1211
3	122	22.35	12.600000000000001	59.7	809
4	123	24.6	8.3	59	1816
5	141	27.05	14.2	59.7	1717
6	142	23.1	14.3	55	1207
7	151	22.1	10.65	52.449999999999996	2175

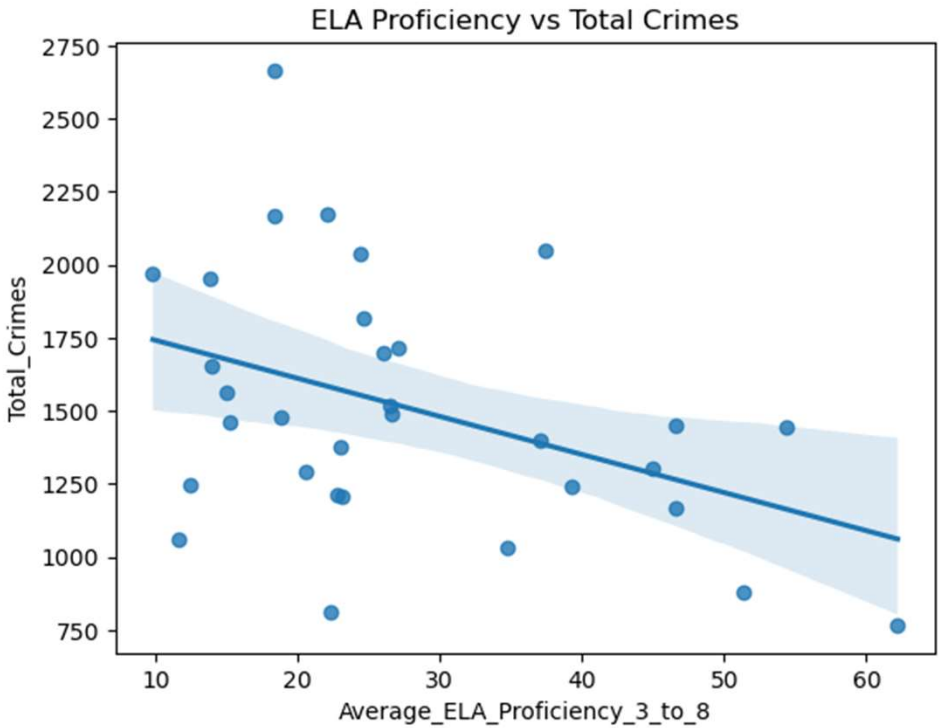


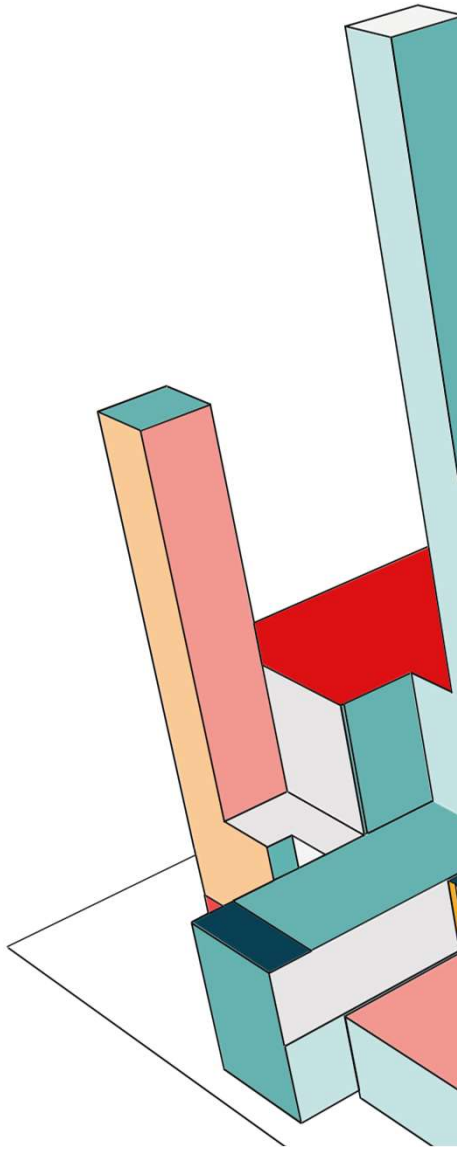
ANALYSIS

When this Crime Incidents and School Metrics for Elementary Schools was exported from SQL and added as a Panda's DataFrame, it was found that there was a correlation between Crime and Schools Metrics.

	District_PSA	Average_ELA_Proficiency_3_to_8	Average_Math_Proficiency_3_to_8	Average_Percent_Attending_90	Total_Crimes
District_PSA	1.000000	-0.405626	-0.443798	-0.324194	0.026480
Average_ELA_Proficiency_3_to_8	-0.405626	1.000000	0.954116	0.785444	-0.410879
Average_Math_Proficiency_3_to_8	-0.443798	0.954116	1.000000	0.760567	-0.388501
Average_Percent_Attending_90	-0.324194	0.785444	0.760567	1.000000	-0.466778
Total_Crimes	0.026480	-0.410879	-0.388501	-0.466778	1.000000

ANALYSIS- WITH PANDAS DATAFRAME





AND NOW WE WILL ANALYZE WITH JAVA SCRIPT VISUALIZATIONS

https://elizabethvandergrift.github.io/Project-3/

THANK YOU

Elizabeth Vandergrift

Juan Camilo Bohorquez Rozo

Eric Croston