

# Civil rights school-related data

Find your dedicated page and share your analysis process on your chunk of data. Include the following:

Two-three inquiry questions you generated during a survey of your data

A screen shot or two of your pivot table that sheds light on your questions

A screen shot or two of your produced charts

A few samples of your formulas used in your analysis

A bulleted list of conclusions you can (or cannot) draw from the data

Doubts or further questions about the data itself

What would you do next if you were developing this into a project?

Sean Scott

Two-three inquiry questions you generated during a survey of your data

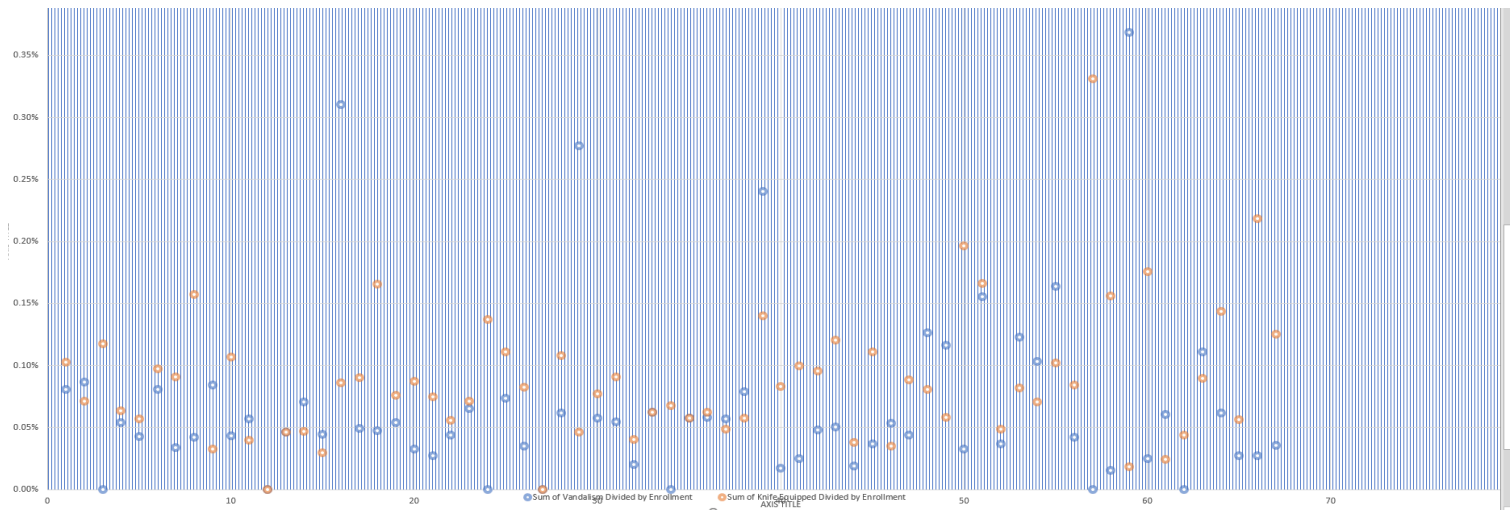
Is there a relationship between frequently of vandalism and frequency of a student equipping a knife when comparing counties?

What are the summary statistics for several violations (Alcohol or drug use, Knife Equipped, Vandalism, Fighting) divided by the number of students enrolled?

A screen shot or two of your pivot table that sheds light on your questions

	Sum of Alcohol and Drugs Divided by Enrollment	Sum of Knife Equipped Divided by Enrollment	Sum of Vandalism Divided by Enrollment	Sum of Fighting Divided by Enrollment
4 Adams	0.19%	0.10%	0.08%	0.21%
5 Allegheny	0.26%	0.07%	0.09%	0.99%
6 Armstrong	0.25%	0.12%	0.00%	0.04%
7 Beaver	0.29%	0.06%	0.05%	0.33%
8 Bedford	0.11%	0.06%	0.04%	0.26%
9 Berks	0.19%	0.10%	0.08%	0.53%
10 Blair	0.15%	0.09%	0.03%	0.25%
11 Bradford	0.19%	0.16%	0.04%	0.51%
12 Bucks	0.13%	0.03%	0.08%	0.20%
13 Butler	0.18%	0.11%	0.04%	0.15%
14 Cambria	0.10%	0.04%	0.06%	0.13%
15 Cameron	0.16%	0.00%	0.00%	0.48%
16 Carbon	0.07%	0.05%	0.05%	0.37%
17 Centre	0.14%	0.05%	0.07%	0.28%
18 Chester	0.20%	0.03%	0.04%	0.36%
19 Clarion	0.12%	0.09%	0.31%	0.14%
20 Clearfield	0.28%	0.09%	0.05%	0.57%
21 Clinton	0.38%	0.17%	0.05%	0.45%
22 Columbia	0.33%	0.08%	0.05%	0.26%
23 Crawford	0.23%	0.09%	0.03%	0.30%
24 Cumberland	0.18%	0.07%	0.03%	0.18%
25 Dauphin	0.18%	0.06%	0.04%	1.13%
26 Delaware	0.16%	0.07%	0.07%	0.98%
27 Elk	0.08%	0.14%	0.00%	0.03%
28 Erie	0.28%	0.11%	0.07%	1.34%
29 Fayette	0.23%	0.08%	0.04%	0.90%
30 Forest	0.19%	0.00%	0.00%	0.19%
31 Franklin	0.20%	0.11%	0.06%	0.18%
32 Fulton	0.05%	0.05%	0.28%	0.65%
33 Greene	0.27%	0.08%	0.06%	0.39%
34 Huntingdon	0.18%	0.09%	0.05%	0.51%
35 Indiana	0.21%	0.04%	0.02%	0.16%
36 Jefferson	0.17%	0.06%	0.06%	0.04%
72 Mean	0.17%	0.09%	0.07%	0.39%
73 Standard Error	9.77442E-05	6.58387E-05	8.48676E-05	0.000347251
74 Median	0.18%	0.08%	0.05%	0.33%
75 Mode	0.00%	0.00%	0.00%	#N/A
76 Standard Deviatio	0.000800071	0.000538913	0.000694671	0.002842371
77 Sample Variance	6.40114E-07	2.90427E-07	4.82568E-07	8.07907E-06
78 Kurtosis	0.322789668	5.417515065	7.598055289	1.454654232
79 Skewness	0.076520194	1.716041422	2.592170853	1.285177472
80 Range	0.38%	0.33%	0.37%	1.34%
81 Minimum	0.00%	0.00%	0.00%	0.00%
82 Maximum	0.38%	0.33%	0.37%	1.34%
83 Sum	11.56%	5.90%	4.46%	25.99%
84 Count	67	67	67	67

A screen shot or two of your produced charts



A bulleted list of conclusions you can (or cannot) draw from the data

There is no relationship between frequency of vandalism and frequency of a student equipping a knife

# Jie

## 1. Two-three inquiry questions you generated during a survey of your data

1) What are the total number of sexual harassments per enrollment and the total number of incidents per enrollment for schools in the Allegheny County, respectively?

2) Which pre-selected counties have the highest and the lowest probability of bullying, fighting, and possession of Knife, respectively?

## 2. A screen shot or two of your pivot table that sheds light on your questions

1) What are the total number of sexual harassments per enrollment and the total number of incidents per enrollment for schools in the Allegheny County, respectively?

County	Sum of Enrollment	Sum of SexualHarassmentPerEnrollement	Sum of IncidentsPerEnrollment
A W Beattie Career Center	655	0	0.019847328
Abraham Lincoln El Sch	328	0	0.009146341
Acadamy at Westinghouse	489	0	0.253578732
Academy CS	188	0.010638298	0.14893617
Acmetonia Primary Sch	301	0	0
Allard El Sch	187	0	0
Allegheny IU 3	504	0	0.047619048
Avalon El Sch	325	0	0.027692308
Avonworth El Sch	779	0	0.002567394
Avonworth HS	417	0.002398082	0.033573141
Avonworth MS	345	0	0.002898551
Baker El Sch	397	0	0
Baldwin SHS	1470	0.000680272	0.019727891
Barrett El Sch	290	0	0.106896552
Bellevue El Sch	363	0	0.03030303
Benjamin Franklin El Sch	338	0	0.00887574
Bethel Memorial El Sch	341	0	0
Bethel Park HS	1561	0	0.062780269
Bon Meade El Sch	454	0	0.002202643
Woodland Hills Academy	483	0	0.00621118
Woodland Hills JHS	485	0	0.094845361
Woodland Hills SHS	1191	0	0.034424853
Wyland El Sch	405	0	0
Young Scholars of Western Pennsylv	230	0	0.004347826
<b>Grand Total</b>	<b>152465</b>	<b>0.228525473</b>	<b>11.65488964</b>

PivotTable Field List

Choose fields to add to report:

- School Name
- County
- Enrollment
- Incidents
- IncidentsPerEnrollment
- SexualHarassmentPerEnrollement

Drag fields between areas below:

<p>Report Filter</p> <p>County</p>	<p>Column Labels</p> <p>Σ Values</p>
<p>Row Labels</p> <p>School Name</p>	<p>Σ Values</p> <p>Sum of Enroll...</p> <p>Sum of Sexua...</p> <p>Sum of Incide...</p>

Defer Layout Update Update

2) Which pre-selected counties have the highest and the lowest probability of bullying, fighting, and possession of Knife, respectively?

Row Labels	Sum of Enrollment	Sum of BullyingPerEnrollment	Sum of FightingPerEnrollment	Sum of PossOfKnifePerEnrollment
Allegheny	142783	0.00107856	0.009889132	0.000714371
Beaver	22136	0.002258764	0.003342971	0.000632454
Clinton	4223	0.000473597	0.004499171	0.001657589
Dauphin	34145	0.001142188	0.011334017	0.00055645
Elk	3642	0.000549149	0.000274574	0.001372872
Franklin	19410	0.000206079	0.001803194	0.001081917
Greene	5188	0.005975328	0.00385505	0.00077101
Huntingdon	5514	0.003445774	0.005077983	0.000906783
Indiana	9833	0.002339062	0.001627174	0.000406793
Jefferson	4797	0.003752345	0.000416927	0.000625391
Lebanon	19026	0.000420477	0.004730369	0.000578156
Mifflin	5273	0.000379291	0.002655035	0.000379291
Northumberland	12007	0.001165987	0.003831099	0.000582993
Philadelphia	137674	0.000239697	0.004263695	0.00166335
Snyder	4889	0.003272653	0.001227245	0.001022704
Tioga	5429	0.001289372	0.004973292	0.000184196
Union	3978	0.001256913	0.001256913	0.001759678
Venango	8239	0.003277097	0.006068698	0.000242748
Washington	27934	0.001288752	0.003257679	0.000894967
York	64633	0.001438893	0.007178995	0.00125323
<b>Grand Total</b>	<b>540753</b>	<b>0.035249977</b>	<b>0.081563213</b>	<b>0.017286943</b>
max of bullying per enrollment		0.005975328		
min of bullying per enrollment		0.000206079		
max of fighting per enrollment		0.011334017		
min of fighting per enrollment		0.000274574		
max of knife possession per enrollment		0.001759678		
min of knife possession per enrollment		0.000184196		

**PivotTable Field List**

Choose fields to add to report:

- Bullying
- BullyingPerEnrollment
- Fighting
- FightingPerEnrollment
- Possession of a Knife
- PossOfKnifePerEnrollment

Drag fields between areas below:

Report Filter:

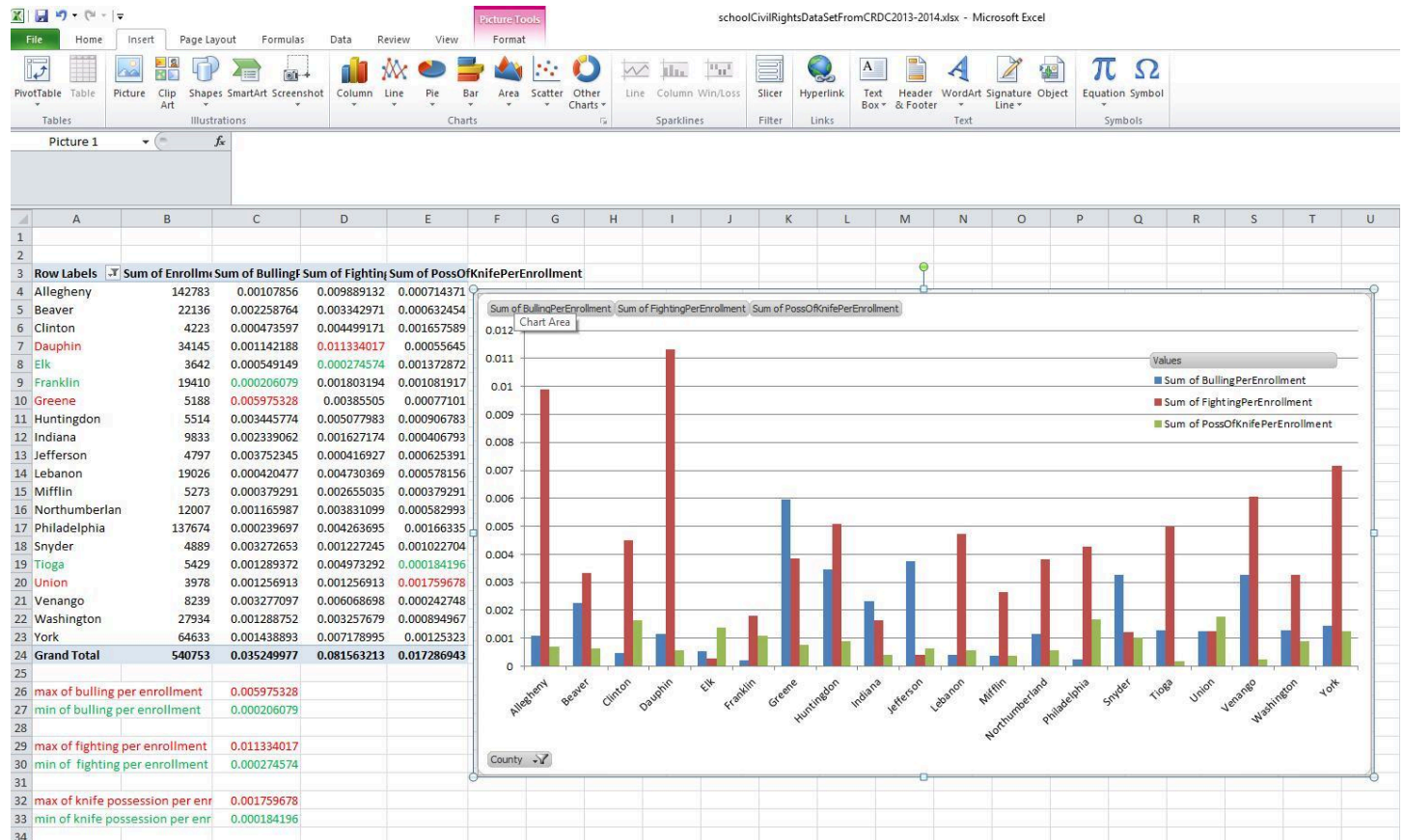
Column Labels:

Row Labels:

Values:

Defer Layout Update Update

### 3. A screen shot or two of your produced charts



#### **4. A few samples of your formulas used in your analysis**

1) probability of bullying (bullying per enrollment) = number of bullying / number of enrollment

2) highest probability of fighting = MAX(D4:D23)

3) lowest probability of knife possession = MIN(E4:E23)

#### **5. A bulleted list of conclusions you can (or cannot) draw from the data**

1) The total number of sexual harassments per enrollment for schools in the Allegheny County is about 0.2285; the total number of incidents per enrollment is about 11.65;

2) Greene County has the highest probability of bullying; Dauphin County has the highest probability of fighting; Union County has the highest probability of knife possession.

3) Franklin County has the lowest probability of bullying; Elk County has the lowest probability of fighting; Tioga County has the lowest probability of knife possession.

#### **6. Doubts or further questions about the data itself**

1) Why are some of the columns left blank? For example, columns named "Assaults on Students", "Sexual Offenses", "Possession of a Firearm", etc.

2) Why does the column "Truancy Rate" shows percentages instead of numbers? How are the percentages calculated?

#### **7. What would you do next if you were developing this into a project?**

1) check the quality of the data. For example, does the spreadsheet mix data, calculation, and reporting?

2) consider adjusting the data set to meet some basic principles. For instance, avoid names, values or fields with blank spaces; avoid using names that contain symbols such as ?, \$, %, ^, &, \*, (,), -, #, ?, ,, <, >, /, |, \, [ , ] , {, and }; make sure that any missing values in the data set are indicated with NA or -9.

# Jeff P.

Find your dedicated page and share your analysis process on your chunk of data. Include the following:

1. Two-three inquiry questions you generated during a survey of your data
  - a. I explored the relationship between the incidence of (reported) bullying and the incidence of a several other classifications of infractions: theft, possession of a knife, possession of tobacco, possession of controlled substances, and disorderly conduct. The comparisons of these incidences were made in terms of a calculated column that divided the total number of each incident for each school district by that school district's total enrollment. I wanted to know if there were any readily identifiable anomalies between the incidence of bullying and these other categories, across the ENTIRE data set. That is, I averaged all districts' enrollment-adjusted metrics for theft, bullying, tobacco possession (etc) and compared the summary statistics of those aggregate averages. I highlighted the bullying column in both the data set and the summary statistics, in green.
  - b. I also wanted to know if there were any noticeable correlations between any of the variables I was tracking, for which I used the CORREL() function in a small table, colored in pink.

2. A screen shot or two of your pivot table that sheds light on your questions

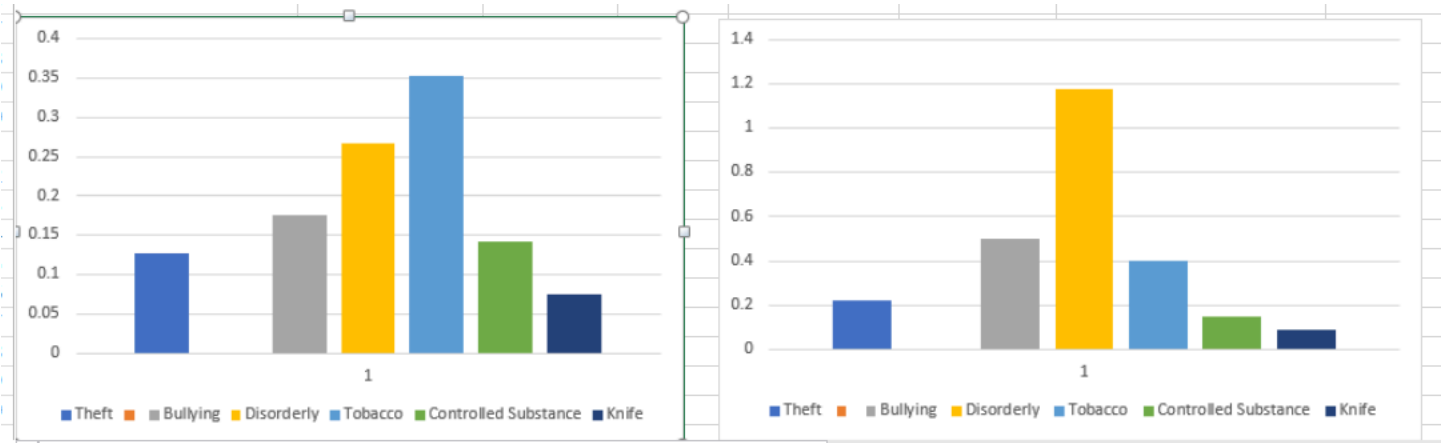
1	LEA Type	District												
2														
3	Row Labels	Sum of Incidents/Enr	Sum of Theft/Er	Sum of Enrol	Sum of Bullying/Enr	Sum of Disorderly	Sum of Possession Tobacco/Enr	Sum of Possession of Sum of Knife/Enrollm					Bullying's	
4	Abington Heights SD	0.302480339	0	3306	0	0	0.030248034	0	0.030248034				CORREL	With...
5	Abington SD	0.287995811	0	7639	0.013090719	0	0.013090719	0.157088624	0.026181437				-0.02001	Enrollment Siz
6	Albert Gallatin Area SD	2.499289974	0.056802045	3521	0.056802045	0.312411247	0.085203067	0.312411247	0.028401022				0.529081	Incidents/Enr
7	Alliquippa SD	1.282051282	0	1170	0	0	0	0.683760684	0				0.102172	Disorderly Cor
8	Allegheny Valley SD	1.210898083	0	991	0.100908174	0.100908174	0.908173562	0	0				0.058402	Drugs
9	Allegheny-Clarion Valley SD	3.698435277	0.426742532	703	0.142247511	0.142247511	0.711237553	0	0				0.08076	Tobacco
10	Allentown City SD	18.71104316	0.687992473	17006	1.522991885	1.664118546	0.2881336	0.111725273	0.270492767				0.037945	Knives
11	Altoona Area SD	2.26342711	0.10230179	7820	0.025575448	0.051150895	0.332480818	0.089514066	0.179028133				0.239202	Theft

=VAR.P(H4:H502)

504	Per Enrollment Categories -->	Incid/Enroll	Theft (Enrollmnt)	Bullying	Disorderly	Tobacco	Controlled Substance	Knife
505	MEAN	2.611314328	0.126600967	3217.01804	0.176177351	0.26629207	0.142843542	0.075313145
506	MEDIAN	1.843075303	0.057234432	2110	0.0241955	0.032310178	0.109349371	0.050994391
507	MIN	0	0	214	0	0	0	0
508	MAX	27.00170358	2.221193892	137674	6.36877247	18.19085487	3.237410072	0.940228341
509	STDEV.P	3.023306611	0.222758213	6626.1007	0.502081914	1.177069911	0.397855705	0.150891758
510	STDEV.S	3.026340537	0.222981754	6632.75008	0.50258576	1.178251116	0.398254958	0.15104318
511	VAR.P	9.140382862	0.049621221	43905210.5	0.252086248	1.385493576	0.158289162	0.022768323

3. A screen shot or two of your produced charts





(on the left: average incidence by category. On the right: stDev by category.)

4. A few samples of your formulas used in your analysis

=STDEV.P(C4:C502)

=MEAN(G4:G502)

=CORREL(E4:E502,B4:B502)

5. A bulleted list of conclusions you can (or cannot) draw from the data

- Just how DOES one define “disorderly conduct?” No one appears to know. Note its aberrantly high standard deviation. But (judging by its stDev, bullying’s definition seems to be at least as well understood as that of “possession of tobacco,” and is as close to being universally understood as possession of drugs or knives or stealing.
- A look at correlations of Bullying:
 

- CORREL	With...
- -0.02001	Enrollment Size
- 0.529081	Incidents/Enrollment
- 0.102172	Disorderly Conduct/Enrollment
- 0.058402	Drugs
- 0.08076	Tobacco
- 0.037945	Knives
- 0.239202	Theft
- 
- There does not appear to be a correlation between enrollment and bullying’s incidence. The positive correlation with Incidents/Enrollment is understandable, since bullying is a subset of this. There is a slightly positive correlation with Disorderly Conduct, and a slightly stronger correlation with Theft. I have no thoughts on the latter correlation, other than curiosity.
- I set out to sniff for inconsistencies in the definition of what bullying is. My choice of theft, knives, drugs, etc were deliberate inasmuch as I thought they were more apt to be clearly articulated offenses (several of them depend on the offender actually possessing something, rather than doing something). My choice to include disorderly conduct was slightly spurious, because I felt it to be similar to bullying inasmuch as both are subjective. I included it to see if their stDev’s across the data set were both high. They are.
- I cannot say from this data that no one knows what bullying is (despite my earlier rhetoric). But I can assert that its definition seems less discrete than many other categories.



6. Doubts or further questions about the data itself

- My project was based upon the doubt about the understanding of a definition (bullying). It is possible that other categories of offenses are diversely understood amongst these many Pennsylvania schools, as well.
- Is smoking or possessing tobacco an offense at all schools? Or most?
- Does the definition of "disorderly conduct" resemble that of law enforcement?

7. What would you do next if you were developing this into a project?

- I'd look for more correlations and include more categories of incidents. That is, I'd broaden my scope of control variables beyond theft, disorderly, drugs, knives, etc.