Data Science Seminar - MSAI 339

Checkpoint 1

October 29, 2021

Professor
Jennie Rogers

Students
Aleksandr Simonyan
Dimitrios Mavrofridis
Donald Baracskay

Theme & Background

By looking at the Tactical Response Reports, we intend to examine the areas and officers involved in violent incidents. We would then hypothesize about the crime in said areas and the officers tangled in such altercations. We will achieve this by filtering the incidents that involve certain actions associated with violence while grouping by location. For example when the police officer's taser or firearm is used against the subject. Additionally, the subjects and officer's physical state will also be taken into consideration along with the number of shots fired during the incident. Moreover, we would like to know whether officers listed in TRRs receive more awards than those who deal with less violent incidents or vice versa. One important aspect that we are going to be covering is understanding the relationship between the number of trr-s and the number of officers in a certain location. We are planning on adding additional dimensions to the data by taking into consideration attributes such as the race. This will allow us to understand whether there is over policing in certain areas of Chicago and whether the TRRs connected in said areas are likely to be violent in nature.

Introduction

For our first checkpoint, we intended to link together the data found in the tactical response reports with community analytics. We will compare the number of TRRs in each community, in order to determine over policing and note the areas where this occurs. We will also use crime data and several action response categories in order to dive into the reasons why over policing may occur based on the crime rates and the socioeconomic status of said areas.

Results

Question 1: Do officers over police areas of lower socioeconomic status more than higher status?

Number of TRRs	Median Income	Community (Name)	Total Population	Policing %
254	$14,916.00	Riverdale	7090	3.58%
458	$19,589.00	Fuller Park	2457	18.64%
1858	$19,816.00	Englewood	26121	7.11%
2365	$21,307.00	East Garfield Park	20665	11.44%
606	$21,869.00	Washington Park	12081	5.02%
2733	$22,132.00	North Lawndale	35276	7.75%
87	$23,457.00	Burnside	2601	3.34%
2093	$23,781.00	West Garfield Park	17733	11.80%
205	$24,140.00	Armour Square	14007	1.46%

Figure 1. Policing percentage on areas with low socioeconomic status

Number of TRRs	Median Income	Community (Name)	Total Population	Policing %
803	$77,639.00	Near South Side	22401	3.58%
2876	$83,382.00	Near North Side	85711	3.36%
40	$84,331.00	Edison Park	11356	0.35%
101	$87,696.00	Mount Greenwood	18783	0.54%
98	$89,038.00	Beverly	20815	0.47%
336	$91,197.00	North Center	34623	0.97%
980	$91,851.00	Loop	33442	2.93%
549	$92,870.00	Lincoln Park	66959	0.82%
59	$101,237.00	Forest Glen	18437	0.32%

Figure 2. Policing percentage on areas with high socioeconomic status

For this question and Question 3, we attempt to determine over policing by calculating the number of filed Tactical Response Reports per community divided by the total population of the community. This way we have a ratio between the number of people living in an area and the number of police responses in the said area. From the data we can clearly see that areas with the lowest median income have the highest rates of over policing and vice versa. We can conclude that the correlation between two variables is very high. In addition, from the generated tables we can see that with a slight increase in the median income number, the policing rate decreased significantly. There is of course variation in both the high and low socioeconomic statuses, and this is likely due to some other factors, potentially crime rates or proportion of ethnicities, which will be covered later in this paper.

Question 2: Do officers respond more violently in areas of lower socioeconomic status?

Total Population	Incident Count	Median Income	Community (Name)	Violent TRRs %
7090	369	$14,916.00	Riverdale	32.4
2457	554	$19,589.00	Fuller Park	30.7
26121	2430	$19,816.00	Englewood	32.1
20665	3000	$21,307.00	East Garfield Park	31.9
12081	789	$21,869.00	Washington Park	33.1
35276	3410	$22,132.00	North Lawndale	30.4
2601	111	$23,457.00	Burnside	31.4
17733	2788	$23,781.00	West Garfield Park	33.3
14007	243	$24,140.00	Armour Square	28.8

Figure 3. Violent Responses in areas with low socioeconomic status

Total Population	Incident Count	Median Income	Community (Name)	Violent TRRs %
22401	900	$77,639.00	Near South Side	29.6
85711	3569	$83,382.00	Near North Side	30.4
11356	40	$84,331.00	Edison Park	24.5
18783	110	$87,696.00	Mount Greenwood	30.1
20815	108	$89,038.00	Beverly	30.0
34623	315	$91,197.00	North Center	23.2
33442	1088	$91,851.00	Loop	28.1
66959	660	$92,870.00	Lincoln Park	29.1
18437	64	$101,237.00	Forest Glen	26.9

Figure 4. Violent Responses in areas with high socioeconomic status

For this question and Question 4, we attempted to determine the percentage of police responses that were violent. For this, we used the action response sub category column from the database. We specified a cutoff of 4.0 (chemical weapon discharged) for violent crimes. For more information on the categories, please see Appendix 1. We have broken down the police action responses by severity (ranging from 0.0 up to 6.0), which is the most violent. For each community, we divided the number of TRRs with a subcategory greater than or equal to 4.0 (violent ones) by the total number of TRRs for that community. From Figures 3 and 4, we can clearly see that areas with lower socioeconomic status experienced considerably more violent police responses than their higher status counterparts.

Question 3: Do officers over police in areas with different proportions of ethnicities (ie. areas with more African Americans)?

Community (Name)	Policing %	Race: African Americans	Race: Hispanic	Race: Asian	Race: White	Race: Other
Fuller Park	18.64%	2268	152	0	37	0
West Garfield Park	11.80%	16954	353	61	315	50
East Garfield Park	11.44%	18862	728	57	878	140
Pullman	8.34%	5380	551	34	464	72
North Lawndale	7.75%	31397	2470	99	811	499
West Englewood	7.48%	29955	1563	19	235	384
Englewood	7.11%	24810	682	76	199	354
Austin	5.99%	81066	11132	481	4353	611
Greater Grand Crossing	5.58%	31077	358	22	466	423
Auburn Gresham	5.23%	44402	689	126	231	394
Humboldt Park	5.03%	22481	28612	271	2958	689
Washington Park	5.02%	11595	158	9	48	271
Woodlawn	4.57%	22555	695	814	1885	497
New City	4.16%	9832	24990	815	5094	352
Roseland	4.07%	40846	450	185	557	467
South Chicago	4.00%	20818	6083	116	782	296
South Shore	3.92%	45990	784	115	1084	1182
Chatham	3.80%	30297	287	129	388	258
Near South Side	3.58%	5382	1393	4450	10404	772

Figure 5. Policing Rates based on Race Distribution per Area

For this question, we extended Question 1, but instead of using median income, we focused on comparing against the proportion of ethnicities in each community. As a result we have generated a table with all the available ethnicities from the CPDB. From here we can note things like the areas with the highest policing have comparatively high proportions of African American/Hispanic populations. From this table, it is clear that Fuller Park, which has a considerably higher number of African Americans compared to other races, has the highest percentage of policing.

Question 4: Do officers respond more violently in areas with different proportions of ethnicities (ie. areas with more African Americans)?

Area ID	count	Community (Name)	Violent TRRs %	African American %	White %	Hispanic %	Asian %	Other %
452	2788	West Garfield Park	33.3	95.6	1.8	2.0	0.3	0.3
432	789	Washington Park	33.1	96.0	0.4	1.3	0.1	2.2
457	1385	South Lawndale	32.4	11.1	3.1	85.2	0.2	0.4
476	369	Riverdale	32.4	94.2	1.1	3.9	0.4	0.5
491	2430	Englewood	32.1	95.0	0.8	2.6	0.3	1.4
490	3079	West Englewood	32.1	93.2	0.7	4.9	0.1	1.2
475	1296	West Pullman	32.0	93.4	1.0	4.8	0.2	0.5
453	3000	East Garfield Park	31.9	91.3	4.2	3.5	0.3	0.7
429	770	Grand Boulevard	31.6	90.7	3.2	2.5	0.5	3.1
477	95	Hegewisch	31.5	6.3	45.6	47.4	0.0	0.7
470	2206	Roseland	31.4	96.1	1.3	1.1	0.4	1.1
468	111	Burnside	31.4	100.0	0.0	0.0	0.0	0.0

449	3455	Humboldt Park	31.3	40.9	5.4	52.0	0.5	1.3
464	2385	South Shore	31.3	93.6	2.2	1.6	0.2	2.4
434	1464	Woodlawn	31.0	85.3	7.1	2.6	3.1	1.9
497	781	Washington Heights	31.0	95.8	1.3	1.1	0.0	1.9
486	502	Gage Park	30.9	3.3	3.5	92.3	0.8	0.1
426	561	Douglas	30.7	70.7	10.7	2.4	14.4	1.9
428	554	Fuller Park	30.7	92.3	1.5	6.2	0.0	0.0

Figure 6. Violent Responses based on Race Distribution per Area

For Question 4, we extended Question 2 but instead determined the percentage of each ethnicity's population when compared to total population. From this we can identify the ethnic breakdowns of areas with comparatively high levels of violent police responses. For example, we can see that the area with the highest violent response rate is also an area with a population consisting almost entirely of African Americans in West Garfield Park.

Question 5: Do officers over police / react more violently in areas with higher crime rates?

Violent TRRs %	Community (Name)	Crime / Population Ratio
20.50%	Fuller Park	928.57%
103.16%	West Garfield Park	730.10%
89.91%	Englewood	688.84%
111.00%	East Garfield Park	622.09%
113.92%	West Englewood	615.98%
29.19%	Washington Park	599.15%
126.17%	North Lawndale	565.68%
82.73%	Greater Grand Crossing	523.12%
40.26%	Loop	496.23%

Figure 7. Policing Rates based on High Crime Rate Communities

Violent TRRs %	Community (Name)	Crime / Population Ratio
24.20%	Albany Park	116.81%
11.66%	North Center	116.44%
8.07%	Clearing	107.92%
10.10%	Jefferson Park	99.33%
9.44%	Dunning	97.42%
4.37%	Norwood Park	81.53%
4.07%	Mount Greenwood	81.39%
2.37%	Forest Glen	67.60%
1.48%	Edison Park	58.87%

Figure 8. Policing Rates based on Low Crime Rate Communities

In order to answer this question we had to import additional data into the CPDB to have access to population and crime statistics for the city of Chicago. We collected the data from the Chicago Data Portal and imported the csv file as a table into the public CPDB. As a result, we were able to map the population and the crimes that occur within areas of Chicago with the tactical response reports in order to identify whether officers

tend to over-police and react more violently in areas with higher crime rates. From the two tables that were generated, we can clearly see a big difference in the crime rates between certain communities. The difference in crime rates is striking. There exist communities whose crime rates are extremely high compared to other communities. Additionally, the crime rate ratio for each one of the communities is also correlated with the violent response ratio. Hence, police officers tend to respond significantly more violently in communities where the crime rate is higher than those with lower ones.

Analysis & Conclusions

Based on our analysis, we can see a strong correlation between crime rates and violent police responses. Police responses also tend to be more forceful in certain communities with different demographic compositions than in others. In addition, we can see how significant a difference is in crime rates in various communities. Besides, police tend to react differently to crimes that take place in communities with higher median income and over-policing those communities that have a relatively small median income.

Future Research

This project could potentially be expanded in a number of ways. First, for Questions 2 and 4, a binary criterion was applied - violent or non-violent. While this works well to simplify the analysis, it does miss out on some granularity in the data. For example, in testing, we noted that if the threshold for violent responses was set to 6.0, some higher socioeconomic status communities actually had much higher percentages than their lower status counterparts. Perhaps this implies that the criminals operating in these communities have access to better weaponry and thus the police are, in turn, forced to escalate their response. Alternatively, there are fewer crimes in these communities and thus these crimes represent a larger part of the TRR dataset. Another avenue would be to find additional breakdowns beyond income and race. For example gender or homelessness proportions could be an interesting breakdown to look into.

Perhaps the most impactful and potentially most difficult breakdown to get would be mental health. It would be interesting to see if the police are more aggressive in areas with higher proportions of mental illness as the people in these areas likely need a less aggressive, more sympathetic approach.

Appendix - Use of Force Metric Table

This table represents the police officers actions severity on a scale from 0 all the way to 6, with 6 being the highest and most aggressive reaction to an event.

Use of Force Metric	Description of Actions Taken
0.0	**Officer Present**
1.0	**Verbal Commands**
2.0	**Canine Response**
3.1	**Taser pointed at subject/spark shown**
3.2	**Chemical Weapon w/ approval**
3.3	**Subject restrained**
4.0	**Chemical Weapon discharged**
4.1	**Open hand strike (likely a slap)**
4.2	**Takedown performed/Handcuffed**
5.1	**Melee strikes (eg. punching)**
5.2	**Forceful impact with a weapon**
5.3	**Taser discharged at subject**
6.0	**Firearm(s) discharged**