

Checkpoint 1: SQL Analytics

Jake Rogers, Kelly Jiang, Alex Reneau

1) *Can relationships be identified between geography, and likelihood of arrest?*

```
SELECT CAST(arrest_true_count AS float) /  
(arrest_true_count+arrest_false_count) AS proportion_arrest,  
(arrest_true_count+arrest_false_count) AS total_warrants , arrest_f.district  
FROM  
(SELECT Count(arrest) AS arrest_true_count, district FROM data_searchwarrant  
WHERE arrest = true GROUP BY district ORDER BY district) arrest_t  
INNER JOIN (SELECT Count(arrest) AS arrest_false_count, district FROM  
data_searchwarrant WHERE arrest = false GROUP BY district ORDER BY district)  
arrest_f  
ON arrest_f.district = arrest_t.district  
ORDER BY proportion_arrest;
```

	proportion_arrest ÷	total_warrants ÷	district ÷
1	0.0975609756097561	123	41
2	0.10932475884244373	622	19
3	0.17391304347826086	322	1
4	0.2554890219560878	501	16
5	0.38015166501813386	3033	11
6	0.3924050632911392	79	18
7	0.49767441860465117	860	31
8	0.5135593220338983	1180	4
9	0.5145631067961165	206	0
10	0.5244215938303342	389	12
11	0.5257861635220126	795	25
12	0.5278654048370137	951	8
13	0.5282608695652173	460	2
14	0.5395299145299145	936	3
15	0.5456919060052219	1149	9
16	0.5714285714285714	1134	10
17	0.5794310722100656	2285	7
18	0.5855855855855856	444	22
19	0.5897435897435898	1053	15
20	0.6119257086999023	1023	6
21	0.6184538653366584	1203	5
22	0.6194690265486725	226	14
23	0.6526315789473685	95	20
24	0.6625766871165644	163	17
25	0.6794871794871795	156	24
26	0.68	25	13

To determine how location is related to the likelihood of an arrest resulting from a search warrant we queried the search warrant table by grouping it by district and then calculating the proportion of warrants that resulted in an arrest. We found that there are very

large differences in these proportions with three districts having a result of less than 0.2 and 6 with over 0.6. It was also interesting how this statistic seemed pretty independent of how many warrants were executed in each district. For example, district 11 has had 3033 warrants executed in it, and the proportion of warrants that resulted in an arrest is 0.38. Whereas district 5 has had 1203 warrants executed in it, but an arrest to warrant proportion of 0.62. Since each judge only presides over 1 district these results could imply that certain judges are more lenient in issuing warrants. This could be investigated with a similar query. The results could also imply some level of bias against populations of people within certain districts either by the judge or by police units requesting a warrant.

These results bring up some interesting questions. One would be how do the relative positions of these districts relate to the proportion of warrants resulting in an arrest. This could be done by linking the districts to longitude and latitude and then calculating the correlation coefficients on those. Another would be to investigate further into the human aspect. Like mentioned above this could involve looking into how these proportions would appear for specific judges instead of districts, or looking into the police unit associated with the warrants to view which proportion of those result in an arrest.

2) Does racial bias play a role in the outcomes of search warrants?

```
SELECT a.name AS district_name, SUM("count") population, black_pop, white_pop
FROM data_racepopulation dr
JOIN data_area a ON a.id=dr.area_id
JOIN (
    SELECT a2.name, SUM("count") black_pop
    FROM data_racepopulation dr2 JOIN data_area a2 ON a2.id=dr2.area_id
    WHERE area_type = 'police-districts' AND race='Black'
    GROUP BY a2.name
) subtable ON a.name = subtable.name
JOIN (
    SELECT a2.name, SUM("count") white_pop
    FROM data_racepopulation dr2 JOIN data_area a2 ON a2.id=dr2.area_id
    WHERE area_type = 'police-districts' AND race='White'
    GROUP BY a2.name
) subtable2 ON a.name = subtable2.name
WHERE area_type='police-districts'
GROUP BY district_name, black_pop, white_pop
ORDER BY district_name;
```

	district_name	population	black_pop	white_pop
1	10th	118093	39440	4300
2	11th	70474	59671	1664
3	12th	127869	23039	51775
4	14th	117738	8027	49809
5	15th	59458	55468	959
6	16th	199482	2022	138045
7	17th	144096	4782	55743
8	18th	117041	10681	88418
9	19th	200786	13305	150551
10	1st	62781	13452	32952
11	20th	91279	9909	49420
12	22nd	101941	62094	34863
13	24th	141038	25322	60488
14	25th	200391	33033	29371
15	2nd	95439	65993	17747
16	3rd	75235	71010	1472
17	4th	123575	76399	9925
18	5th	74396	70064	843
19	6th	90841	88525	312
20	7th	71071	68787	262
21	8th	247373	52219	51491
22	9th	165201	19044	24307

Our findings on this front are inconclusive. Of the districts with the top proportion of arrests, the 24th, 17th, 20th, and 14th are predominantly white, while the remainder of the top

10 (districts 5, 6, 15, 22, 7, and 10) are located within the predominantly black south and west sides of Chicago. We note, however, that the predominantly black districts drive an overwhelmingly higher number of search warrants overall.

```
SELECT a.name AS district_name, COUNT(DISTINCT s.id) AS search_warrant_cnt, population, white_pop, black_pop
FROM data_area a
JOIN data_searchwarrant s ON ST_INTERSECTS( geog1: a.polygon, geog2: s.point)
JOIN data_racepopulation dr ON a.id = dr.area_id
JOIN (
    SELECT a2.name, SUM("count") black_pop
    FROM data_racepopulation dr2 JOIN data_area a2 ON a2.id=dr2.area_id
    WHERE area_type = 'police-districts' AND race='Black'
    GROUP BY a2.name
) subtable ON a.name = subtable.name
JOIN (
    SELECT a2.name, SUM("count") white_pop
    FROM data_racepopulation dr2 JOIN data_area a2 ON a2.id=dr2.area_id
    WHERE area_type = 'police-districts' AND race='White'
    GROUP BY a2.name
) subtable2 ON a.name = subtable2.name
JOIN (
    SELECT a2.name, SUM("count") AS population
    FROM data_racepopulation dr2 JOIN data_area a2 ON a2.id=dr2.area_id
    WHERE area_type = 'police-districts'
    GROUP BY a2.name
) subtable3 ON subtable3.name = a.name
WHERE area_type='police-districts'
GROUP BY district_name, population, white_pop, black_pop
ORDER BY search_warrant_cnt DESC;
```

district_name	search_warrant_cnt	population	white_pop	black_pop
11th	3092	70474	1664	59671
7th	2297	71071	262	68787
9th	1280	165201	24307	19044
4th	1183	123575	9925	76399
10th	1140	118093	4300	39440
5th	1109	74396	843	70064
6th	1040	90841	312	88525
15th	990	59458	959	55468
8th	976	247373	51491	52219
3rd	955	75235	1472	71010
25th	803	200391	29371	33033
19th	624	200786	150551	13305
16th	532	199482	138045	2022
22nd	512	101941	34863	62094
12th	432	127869	51775	23039
1st	378	62781	32952	13452
2nd	369	95439	17747	65993
14th	237	117738	49809	8027
17th	168	144096	55743	4782
24th	156	141038	60488	25322
20th	99	91279	49420	9909
18th	88	117041	88418	10681

This is suggestive of racial bias in CPD's macro level decision making. If police are more aggressive in pursuing search warrants in primarily black districts despite no evidence that these searches are more likely to result in arrests, that should be investigated.

3) What is the geographic extent of search warrants executed by each police unit?

```
SELECT avg(x.num_districts), max(x.num_districts), min(x.num_districts)
FROM (
    SELECT Count(DISTINCT district) as num_districts, policeunit_id FROM
data_searchwarrant GROUP BY policeunit_id
) x;
```

	avg	max	min
1	11.5087719298245614	28	1

To determine the geographic extents of search warrants executed by specific police units we queried the search warrant table to determine how many districts each police unit has executed a warrant in. We found that the average number of districts a police unit

operates in is 11.51, the min is 1, and the max is 28. This means that most units are involved with several districts and brings up the question of how those districts are impacted by the specific units. It would be interesting to plot the locations of the executed warrants on a map to see how much different police units overlap, and what percentage of warrants executed in certain areas are by the same unit. Interestingly, from looking at the results from these two queries we can see that many of the highest per officer warrant execution rates also execute warrants in more districts. These results can also be tied back to the first question. There was a large difference between the proportion of search warrants resulting in an arrest between districts, so it would be interesting to see if there is a similar relationship with police units.

4) *What proportion of lawsuits list unlawful search/seizure as the primary cause?*

```
SELECT COUNT(*), primary_cause
FROM lawsuit_lawsuit
GROUP BY primary_cause
HAVING COUNT(*) > 1;
```

50	Excessive force		
5	Monell		
2			
196	Illegal search/seizure		
29	Illegal search & seizure		
11	Malicious prosecution		
288	Excessive force/minor		
37	Other police misconduct		
2	Illinois whistleblower act		
13	Extended detention/malicious prosecution		
471	False arrest		
67	Extended detention		
3	Pursuit/offender accident		
23	Dui stop		
2	Illinois domestic violence act		
29	Failure to provide medical care		
174	Excessive force/serious		
34	Reversed conviction		

$$(196 + 29)/1454 = 15.47\%$$

Note adding the HAVING COUNT(*) > 1 clause to this query halves the number of rows returned, smoothing over some data entry irregularities such as typographical errors. Also note that not all lawsuits resulting from the execution of a search warrant list illegal search as a primary cause. From matching against summaries, it is possible to find lawsuits from search warrants under the false arrest and excessive force categories. Lawsuits will be a valuable tool for investigating the execution of search warrants, including patterns of misconduct like serving warrants at incorrect addresses and unlawful searches with no warrant at all.

5) *Of the complaints categorized as improper search, what proportion alleged search of premise without a warrant?*

```
SELECT category, allegation_name, COUNT(DISTINCT allegation_id) cr_count
FROM data_allegationcategory c JOIN data_officer allegation d
ON c.id = d.allegation_category_id
WHERE category_code ILIKE '03_'
GROUP BY category, allegation_name;
```

	category	allegation_name	cr_count
1	False Arrest	Illegal Arrest / False Arrest	3601
2	False Arrest	Improper Detention	58
3	False Arrest	Planting Controlled Substance In Premise	2
4	False Arrest	Planting Controlled Substance In Vehicle	4
5	False Arrest	Planting Controlled Substance On Person	17
6	False Arrest	Unlawful / Excessive Investigative Detention (Witness)	3
7	First Amendment	Failure To Ensure Civil Rights	7
8	First Amendment	First Amendment	30
9	Illegal Search	Improper Search Of Person	2910
10	Illegal Search	Improper Search Of Vehicle	2447
11	Illegal Search	Miscellaneous	27
12	Illegal Search	Search Of Person - Property / Use Taken	23
13	Illegal Search	Search Of Premise - Property / Use Taken	6
14	Illegal Search	Search Of Premise Without Warrant	8252
15	Illegal Search	Search Of Vehicle - Property / Use Taken	6
16	Racial Profiling	Racial Profiling	27
17	Use Of Force	Illegal Threat Of Arrest	1
18	Use Of Force	Injury / Death (Under Color Of Law)	8

Narrowing the search of the complaint report table categorically was helpful for determining which complaints were likely to be associated with either a search warrant or a lawsuit. This query shows search of premises without a warrant was the most common category within the improper search group, accounting for nearly half of the rows returned. The high volume of unlawful searches reported is troubling given the small number of lawsuits returned by the previous query.

```
SELECT DISTINCT d.allegation_id, category_code, cr_text
FROM data_allegationcategory c
JOIN data_officer allegation d ON c.id = d.allegation_category_id
JOIN data_allegation da ON d.allegation_id = da.crid
WHERE category_code = '03C' AND cr_text IS NOT NULL;
```

As demonstrated by the intake texts generated by this query, these offenses are extremely serious, as they are often accompanied by other forms of misconduct, including stalking,

property damage, and theft. Better understanding of how Chicago PD chooses to search residences, with or without warrants, could help to improve the transparency of their investigative processes.