

## CP1 Findings

CS496

Question 1: What are the most popular categories of complaints when it comes to civilian complaints? What about officer complaints? Are there any similarities to the trends between the two?

```
--most popular civilian complaints
SELECT dac.category, dac.allegation_name, COUNT(*)
FROM data_allegation da JOIN data_allegationcategory dac on da.most_common_category_id
= dac.id
WHERE da.is_officer_complaint = false
GROUP BY dac.category, dac.allegation_name
ORDER BY COUNT(*) DESC
LIMIT 10;
```

category	allegation_name	count
Use Of Force	Excessive Force / On Duty - Injury	16633
Operation/Personnel Violations	Inadequate / Failure To Provide Service	11997
Operation/Personnel Violations	Neglect Of Duty	8622
Illegal Search	Search Of Premise Without Warrant	8152
Operation/Personnel Violations	Miscellaneous	7097
Verbal Abuse	Use Of Profanity	6560
Use Of Force	Excessive Force / Off Duty - No Injury	4342
Operation/Personnel Violations	Reports	3835
Lockup Procedures	Prisoners Property	3641
False Arrest	Illegal Arrest / False Arrest	3537

Based on the result from our queries, the most popular allegations when it comes to civilian complaints belong in the categories of Use of Force, Operation/Personnel Violations, and other miscellaneous categories such as Illegal Search, False Arrests, and Verbal Abuse. Specific allegations include Excessive force for on/off-duty cops that may or may not have resulted in injury, failure to provide service/neglect of duty, and search of premise/belongings without a warrant. These results present an interesting finding: the most popular complaints are at opposite sides of the spectrum, where excessive force and search of premise without warrant are *overextensions* of power, whereas the second and third most common complaints are due to simply not doing

their job. One can reason that since both are present and commonly reported, police officers are simply not doing their job *correctly* at very significant rates.

```
--most popular officer complaints
SELECT dac.category, dac.allegation_name, COUNT(*)
FROM data_allegation da JOIN data_allegationcategory dac on da.most_common_category_id
= dac.id
WHERE da.is_officer_complaint = true
GROUP BY dac.category, dac.allegation_name
ORDER BY COUNT(*) DESC
LIMIT 10;
```

category	allegation_name	count
Operation/Personnel Violations	Neglect Of Duty	1685
Operation/Personnel Violations	Miscellaneous	822
Conduct Unbecoming (Off-Duty)	Association With Felon	547
Operation/Personnel Violations	Seat Belts	392
Operation/Personnel Violations	Insubordination	390
Operation/Personnel Violations	Misuse Of Department Equipment / Supplies	349
Domestic	Domestic Altercation - Physical Abuse	317
Lockup Procedures	Escape	259
Operation/Personnel Violations	Weapon / Ammunition	259
Operation/Personnel Violations	Inadequate / Failure To Provide Service	242

For officer complaints, most complaints belong in the Operation/Personnel Violations category for things such as neglect of duty, associating with felons, not wearing seatbelts, misuse of department equipment/supplies/weapons/ammunition, and insubordination, with other categories such as domestic altercation/assault. One thing that hasn't really been discussed in our readings is this "Association with Felon" complaint; the fact this is the third most common complaint warrants further analysis. Does this prescribe certain feelings towards repeat offenders, and contribute to a police culture that reinforces the prison industrial complex? On another note, we see neglect of duty at the top; we think this is telling of a general culture of police simply being ineffective and we will see if the other data supports this. Lastly, it should be noted that police across the country have significantly higher reported cases of domestic abuse than the rest of the population; this query result supports that as well.

Question 2: What is the average years of service for police officers over a certain threshold of complaints? What positions do these officers generally hold i.e. field cop or office cop, high rank or low rank etc.?

```
--average allegation counts and civilian compliment counts per officer
SELECT AVG(allegation count), AVG(civilian compliment count)
FROM data_officer
WHERE data_officer.active = 'Yes';
```

```
+-----+-----+
|avg      |avg      |
+-----+-----+
|10.9696391394864677|3.3318875780707842|
+-----+-----+
```

The average allegation count of all active officers is about 11 allegations per officer. The average compliment count from civilians is about 3. We think 11 allegations per officer on average for an active duty police officer is extremely alarming! The ratio of allegation to compliment also shows that in general, those who are willing to write on police interactions will likely be disapproving of their interaction.

```
--average age of all police officers
SELECT AVG(2022 - birth_year)
FROM data_officer
WHERE data_officer.active = 'Yes';
```

```
46.2989643964018444
```

```
--average age of police officers with above average allegation counts
SELECT AVG(2022 - birth_year)
FROM data_officer
WHERE active = 'Yes'
AND allegation_count > (SELECT AVG(allegation count) FROM data_officer);
```

```
52.6951432564640112
```

The average age of all active duty officers is about 46 years of age, and the average age of active duty officers with an allegation count over the threshold of the average complaint count is 53 years old. We wanted to do this to gauge how many complaints

per year an officer may receive on average; if we consider that an officer on average begins in the mid 20s, this is around 1 allegation every 3 years. Of course, an average doesn't indicate distribution; we are interested in the distribution of allegations across officers as well, and that will come at a further point. We should note that in the future, we'd like to do this against appointment date instead of age (to implicit biases against ageism, for example), but that requires strings to be stripped from the appointment\_date column. We are also slightly alarmed that police officers can maintain their jobs for so long on average when they have received 10+ allegations on average over their careers.

We should note that our proposal indicated we wanted to analyze rank. We chose to ignore this part of the question, as during our proposal critique we were informed that rank was mostly flat (as in most of the officers were of the same rank) so we did not attempt to answer that question.

Question 3: Which area of patrol has the highest amount of complaints per capita?

```
--area of patrol with highest allegations per capita
SELECT daa.area_id, CAST(COUNT(*) as float) / CAST(tp.total_pop as float) as
complaints_per_capita
FROM data_allegation_areas daa
JOIN (
  SELECT SUM(count) as total_pop, area_id
  FROM data_racepopulation
  GROUP BY area_id) as tp on daa.area_id = tp.area_id
GROUP BY daa.area_id, tp.total_pop
ORDER BY CAST(COUNT(*) as float) / CAST(tp.total_pop as float) DESC
LIMIT 10;
```

area_id	name	area_type	total_pop	complaints_per_capita
428	Fuller Park	community	2457	1.0162800162800163
472	Pullman	community	6501	0.3977849561605907
463	Loop	community	33442	0.27746546259195026
453	East Garfield Park	community	20665	0.22013065569804016
1549	1st	police-districts	62781	0.21689683184402925
490	West Englewood	community	32156	0.19660405523075009
426	Douglas	community	20323	0.18619298331939182

1546	11th	police-districts	70474	0.1757669495132957
491	Englewood	community	26121	0.1619769534091344
452	West Garfield Park	community	17733	0.15716460835729995

```
--area of patrol with highest allegations per capita per police district
SELECT daa.area_id, da.name, da.area_type, tp.total_pop, CAST(COUNT(*) as
float) / CAST(tp.total_pop as float) as complaints_per_capita
FROM data_allegation_areas daa
JOIN (
  SELECT SUM(count) as total_pop, area_id
  FROM data_racepopulation
  GROUP BY area_id) as tp on daa.area_id = tp.area_id
JOIN data_area da on daa.area_id = da.id
WHERE da.area_type = 'police-districts'
GROUP BY daa.area_id, tp.total_pop, da.name, da.area_type
ORDER BY CAST(COUNT(*) as float) / CAST(tp.total_pop as float) DESC
LIMIT 10;
```

area_id	name	area_type	total_pop	complaints_per_capita
1549	1st	police-districts	62781	0.21689683184402925
1546	11th	police-districts	70474	0.1757669495132957
1534	7th	police-districts	71071	0.15454967567643624
1535	3rd	police-districts	75235	0.11573071044061939
1547	15th	police-districts	59458	0.11386188570083083
1551	2nd	police-districts	95439	0.11252213455715168
1540	5th	police-districts	74396	0.11195494381418356
1537	6th	police-districts	90841	0.11148049889367136
1536	4th	police-districts	123575	0.07898037628970261
1545	12th	police-districts	127869	0.07808773041159311

For this question, we see that there are certain communities that have police allegations that have been against them at extremely high rates; for Fuller Park, for example, the number of allegations is higher than the entire community's population, which is highly alarming. When looking at a cursory google search of Fuller Park, Pullman, and East Garfield Park, we find that these communities are predominantly Black and impoverished, which may lead to a conclusion that Black communities are being heavily overpoliced and that overextensions of force can perhaps be found at exceedingly high levels in these communities, inviting the possible (and to be frank, extremely likely) conclusion that police are heavily biased when policing these types of communities.

We wanted to also look solely at police districts, because the data\_racepopulation table that we have queried the total population from is first and foremost comprehensive for police districts and communities; but also, because we wanted to see the bigger picture, where total\_population is the largest. We find that the community and police-district data largely support each other; for example, the 1st district covers the Loop and surrounding areas, and 11th district and 7th districts cover Garfield Park and Englewood respectively, which we can see is are communities that have a very high number of allegations per capita as seen in the first query. We move on to question 4 to answer further questions about allegations made in these police districts.

Question 4: What is the demographic profile of the districts with the highest number of allegations per capita? We changed this from our original proposal question, though both are concerned with demographic and race breakdowns of the complaint data. The original question of comparing race of officers and areas of patrol for each given area was unanswerable by the tables provided in the database.

```
--Demographic breakdown for above query
SELECT da.name as police_district,
       CAST(CAST(COUNT(*) as float) / CAST(tp.total_pop as float) as decimal(16,2))
as complaints_pc,
       CAST(CAST(bp.Black_population as float) / CAST(tp.total_pop as float) as
decimal(16,2)) as p_Black,
       CAST(CAST(hp.Hispanic_population as float) / CAST(tp.total_pop as float) as
decimal(16,2)) as p_Hispanic,
       CAST(CAST(ap.Asian_population as float) / CAST(tp.total_pop as float) as
decimal(16,2)) as p_Asian,
       CAST(CAST(wp.White_population as float) / CAST(tp.total_pop as float) as
decimal(16,2)) as p_White,
       CAST(CAST(op.Other_population as float) / CAST(tp.total_pop as float) as
decimal(16,2)) as p_Other
FROM data_allegation_areas daa
JOIN (
```

```

SELECT SUM(count) as total_pop, area_id
FROM data_racepopulation
GROUP BY area_id) as tp on daa.area_id = tp.area_id
JOIN (
SELECT count as Black_population, area_id
FROM data_racepopulation
WHERE race = 'Black') as bp on daa.area_id = bp.area_id
JOIN (
SELECT count as Hispanic_population, area_id
FROM data_racepopulation
WHERE race = 'Hispanic') as hp on daa.area_id = hp.area_id
JOIN (
SELECT count as White_population, area_id
FROM data_racepopulation
WHERE race = 'White') as wp on daa.area_id = wp.area_id
JOIN (
SELECT count as Asian_population, area_id
FROM data_racepopulation
WHERE race = 'Asian') as ap on daa.area_id = ap.area_id
JOIN (
SELECT count as Other_population, area_id
FROM data_racepopulation
WHERE race = 'Asian') as op on daa.area_id = op.area_id
JOIN data_area da on daa.area_id = da.id
WHERE da.area_type = 'police-districts'
GROUP BY daa.area_id, tp.total_pop, da.name, da.area_type, bp.Black_population,
hp.Hispanic_population, wp.White_population, ap.Asian_population,
op.Other_population
ORDER BY CAST(COUNT(*) as float) / CAST(tp.total_pop as float) DESC
LIMIT 10;

```

police_district	complaints_pc	p_black	p_hispanic	p_asian	p_white	p_other
1st	0.22	0.21	0.06	0.17	0.52	0.17
11th	0.18	0.85	0.12	0.00	0.02	0.00
7th	0.15	0.97	0.02	0.00	0.00	0.00
3rd	0.12	0.94	0.01	0.00	0.02	0.00
15th	0.11	0.93	0.04	0.00	0.02	0.00
2nd	0.11	0.69	0.03	0.06	0.19	0.06
5th	0.11	0.94	0.03	0.00	0.01	0.00
6th	0.11	0.97	0.01	0.00	0.00	0.00
4th	0.08	0.62	0.29	0.00	0.08	0.00
12th	0.08	0.18	0.33	0.06	0.40	0.06

The result from our query indicates the police districts with the highest number of complaints per capita are predominantly Black, indicating again the likelihood of overpolicing and overuse of force in predominantly Black populations. However, the first and 12th districts are outliers, where the 1st district is majority White and the 12th

district has a white plurality. When conducting a cursory search of the number of crime complaints between the 1st and 11th districts, we find that they experience a comparable number of crime complaints (2012 for 1st district vs 2748 for 11th district in 2022). Further analysis needs to be done to see why complaints are so high in the 1st district, as it is a significant outlier; there could be a multitude of explanations. One suspicion is that despite the population in the 1st district (containing Loop and surrounding areas) is that this is not representative of the people that are present in the loop for these allegations, as the 1st district is Downtown Chicago and this is a central meeting point for differing demographics of people to travel to spend time in (and could also could make the population in the 1st district highly variable at any given time). Another suspicion is that perhaps Chicago Police has overinvested resources in protecting the “image” of a predominantly wealthy and White neighborhood, and the number of complaints may reflect an increased police presence in this district. However, these just suspicions, to be further explored.