

Influence of Race and Position?

The Earnest Pirates

Vinit Todai, Shreyas Lele, Tejul Pandit

Background and theme -

Our analysis attempts to determine the relationship between allegations against police officers and the impact that their race or rank has on these allegations being sustained. Using this theme, we want to analyze the possibility of racial prejudice within the CPD and if higher-ranking officers can influence judgements against them. This would also help in identifying those few police officers who misuse their authority or race, allowing them to be appropriately disciplined.

Relational Analytics -

Our initial discovery was made by querying the data to determine the number of allegations with respect to different police units and then computing the per-officer-allegation-count. This can be used to determine the average number of allegations made against a police officer in a given unit.

police_unit	unit_description	count_allegations	total_officers	allegation_ratio_per_unit
97	Community Relations Division	91	3	30.33
173	Gang Enforcement - Area Central	664	27	24.59
175	Gang Enforcement - Area North	452	19	23.78
104	Caps	22	1	22
113	Canine Unit	22	1	22
180	Alternate Response Section	393	19	20.68
174	Gang Enforcement - Area South	536	26	20.61
266	Bureau Of Detectives - Investigative Response ...	525	26	20.19
165	Troubled Building Unit	120	6	20

Table 1: per-officer-allegation-count in a police unit in descending order

Table 1 highlights which police units received the most allegations against their officers. Furthermore, the ‘allegation ratio per unit’ denotes that the average number of allegation counts per officer in the police units listed in Table 1 is between 20 and 30. This analysis calls attention to the police unit mentioned..

To dive deeper and gain insight into officer level information, we mapped each officer with their associated ranks, total number of allegations, and total number of allegations sustained. The results from the CPDP database are shown in Table 2. Furthermore, it can be observed that police officer Jerome Finnigan has the most allegations, 175, against him, with only two of them being sustained.

id	first_name	last_name	rank	total_allegations	sustained_allegations
8562	Jerome	Finnigan	Police Officer	175	2
21837	Joe	Parker	Police Officer	137	12
17816	Edward	May	Police Officer	136	3
8138	Glenn	Evans	Lieutenant of Police	132	12
21468	Kevin	Osborn	Police Officer	127	0
28805	Charles	Toussas	Police Officer	123	5
31631	Adam	Zelitzky	Sergeant of Police	121	9
29033	Jerome	Turbyville	Police Officer	117	3
32166	Emmett	Mc Clendon	Sergeant of Police	110	2
4807	Maurice	Clayton	Police Officer	109	10
13788	Broderick	Jones	Police Officer	107	11
32164	Tamara	Matthews	Sergeant of Police	106	1
31119	Kenneth	Wojtan	Police Officer	106	1
2015	Eugene	Bikulcius	Sergeant of Police	102	3
10890	James	Grubbs	Police Officer	102	4
32265	Raymond	Piwnicki	Sergeant of Police	102	6
3897	Thomas	Carey	Police Officer	100	3
13095	Gregory	Jackson	Sergeant of Police	98	1
11634	Frederick	Hasenfang	Police Officer as Detective	97	3
23265	Robert	Quintero	Police Officer	96	4
12478	Ronald	Holt	Director of CAPS	95	2
8658	Corey	Flagg	Police Officer	95	6
25306	James	Sanchez	Lieutenant of Police	94	2
17041	Paul	Major	Police Officer	93	4

Table 2: count of allegations and sustained allegations for each officer

Moreover, a count of allegations with their sustained count was calculated for each of the ranks too as shown in Table 3 and can also be seen in the visualizations in Fig 2. The ranks such as Chief, Director, Deputy Director and Detective are highly authoritative individuals who have a high number of allegations against them but an abnormally low sustained allegations count. This suggests that some of the officers that belong to these ranks may be misusing their power and authority.

rank	total_allegations	sustained_allegations	percent
Director of CAPS	95	2	2.11
Chief	300	6	2
Lab Tech III	4	0	0
Commander Of General Support	2	0	0
Police Aide	3	0	0
Police Forensic Investigator II	3	0	0
SR Evidence Technician	9	0	0
Explosives Tech II	5	0	0
Police Lab Tech IV	1	0	0
Police Lab Tech II	5	0	0
Commander Of Neighborhood Relations	1	0	0
Principal Stenographer	4	0	0
Director Of Human Resources	6	0	0
Director Of Motor Maintenance	4	0	0
Police Communications Operator I	3	0	0
F/A Id Tech II	3	0	0
Physical Fitness Coordinator	2	0	0
Deputy Director	36	0	0
Deputy Chief Administrator	14	0	0
Director Management / Labor Affairs	2	0	0
Police Forensic Investigator III	4	0	0
Superintendent's Chief Of Staff	1	0	0
Explosives Tech III	12	0	0
Detective	41	0	0

Table 3: Ranks and count of allegations and sustained allegations

Lastly, observations were made on the race of the officers and their corresponding count of allegations and sustained allegations. An interesting insight was identified as seen in Table 4 wherein White officers, who collectively have the most number of allegations against them, have the same count of sustained allegations as their Black counterparts.

race	total_allegations	sustained_allegations	percent
White	137448	8078	5.88
Black	68108	8274	12.15
Hispanic	39534	2324	5.88
Asian/Pacific	3950	211	5.34
Native American/Alaskan Native	660	45	6.82
Unknown	88	13	14.77

Table 4: Race and count of allegations and sustained allegations

In fact, table 4 highlights that allegations against Black officers are sustained the most as compared to officers belonging to any other race. This might indicate a suggested bias against Black officers within the CPD.

Visualization -

Firstly, the following Fig 1 gives an overview of the rankings in the Police Department.

Title	Insignia	Notes
Superintendent of Police	★ ★ ★ ★	Appointed by the Mayor of Chicago. Highest rank in the Chicago Police Department.
First Deputy Superintendent of Police	★ ★ ★	Appointed by the Superintendent of Police. Second highest rank in the Chicago Police Department.
Chief	★ ★	Rank since September 8, 2011. Chiefs are typically in charge of a bureau.
Deputy Chief	★	Rank since September 8, 2011.
Commander	✿	Commanders are typically in charge of a district.
Captain	■	Captains are typically executive officers of districts.
Lieutenant	■	
Sergeant	■ ■	
Field Training Officer	■ ■ FTO	Field training officers wear one chevron over one rocker, with "FTO" in the center of the insignia, but are not considered ranking officers.
Police Officer/assigned as: Detective/Youth Officer/Gang Specialist/Police Agent/Major Accident Investigator/ etc	No Insignia	Chicago detectives are not considered ranking officers, but rather officers assigned to specialized units, e.g. violent crimes, robbery, gang and narcotics (NAGIS), Internal Affairs Division (IAD), Major Accident Investigation Section (MAIS), etc. (Unless they hold the rank of sergeant or above.)
Police Officer	No Insignia	Police officers are the first ranking officers. They are dispatched to radio assignments, conduct patrol, and respond to other emergencies as needed.

Figure 1: Police Department - Officer Ranks

The following visualization (Fig 2) depicts the rank of the police officers in relation to the overall number of complaints registered. The greater the size of the bubble, the greater the number of complaints. The top three ranks of police against whom the most complaints are filed are as follows: 1) Police officers (Total allegation count: 150,977, Sustained allegation count: 12,843). 2) Sergeant of Police (Total allegation count: 40,677, Sustained allegation count: 2,373). 3) Police Officer as Detective (Total allegation count: 28,062, Sustained allegation count: 1,345)

Rank v/s Complaint counts

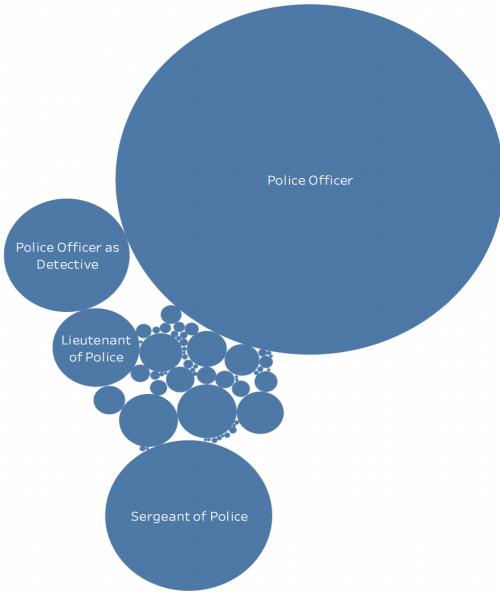


Figure 2: Number of allegations w.r.t Officer Ranks

The second graphic (Fig 3) below depicts the percentage of allegations that remain unsustained for each police officer rank.. This graph supports the hypothesis that higher-ranking officers leverage their position to their advantage. The index indicates that the brighter the color, the higher the percentage of allegations that are not sustained.

Rank v/s %allegations not sustained

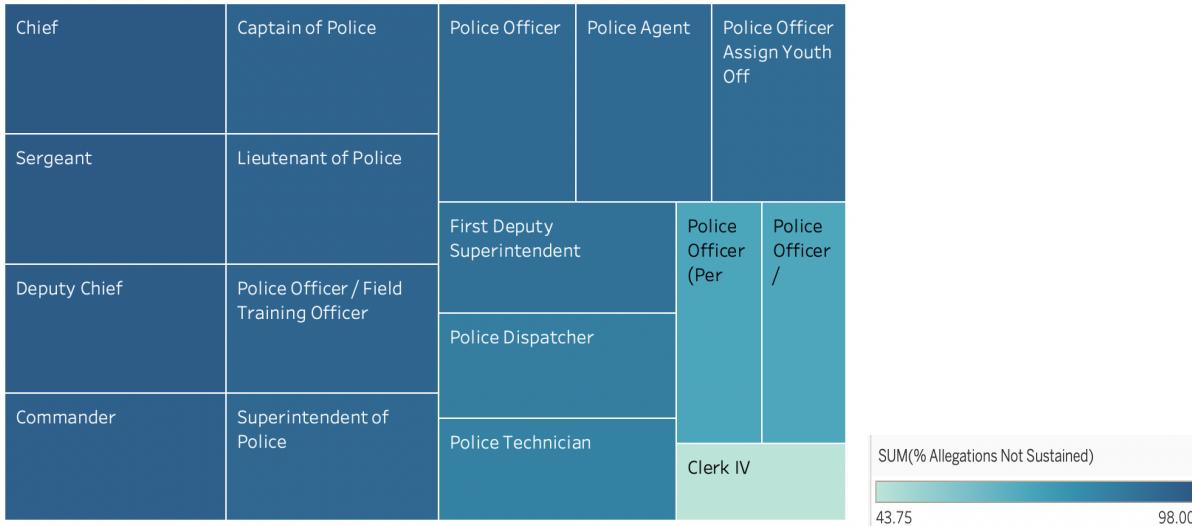


Figure 3: Percentage of allegations not sustained w.r.t Officer Rank

Using the graph below, we try to determine whether race has any effect on the likelihood of an allegation being sustained. The graph below (Fig W) shows that White officers had the most allegations made against them, approaching 140k in total. However, the total number of

allegations sustained is just about 8000. The amount of claims against Black police, on the other hand, is 68K, with 8K sustained allegations. Similarly, the number of claims filed by Hispanics is 40K, with a sustained count of 2.3K. As a result, the ratio of sustained count to allegation count for White Officers is quite low. However, it is double for black officers. As a result, it is possible that a police officer's race could be a factor in determining whether or not an allegation against an officer is sustained.

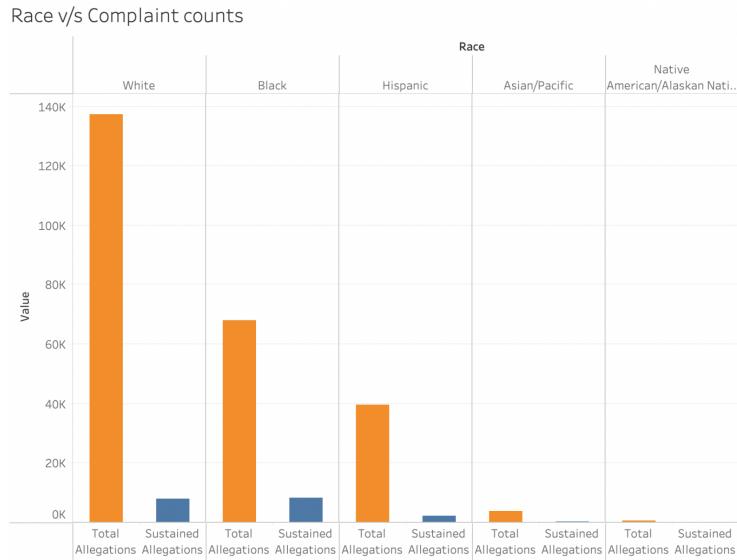


Figure 4 : Number of Allegations and sustained allegations w.r.t Race

Interactive Visualization -

We aimed to investigate the temporal element of the amount of allegations filed against individual police units every year and how it changed over time. In addition, we wanted to visualize the percentage of sustained allegations based on the rank of the police officer.

Bar chart race: Identifying distribution of allegations in a particular police unit with time.

The objective of this visualization is to gain an understanding of the number of allegations associated with specific police units over time. We utilized a bar chart approach¹ to visualize the amount of allegations so that we could see which police units receive more allegations each year.

Zoomable-bar-chart: Identify the relationship between the rank of the police officer and the percentage of allegations sustained.

The objective of this visualization is to examine the relationship between police officer rank and the percentage of allegations sustained. Zooming² the interactive visualization allows for a better understanding of the individual numbers. The insights from the below graph falls in-line with the hypothesis that higher ranking officers leverage their rank.

¹ <https://observablehq.com/@5b0e7a483d797e6e/bar-chart-race>

² <https://observablehq.com/@5b0e7a483d797e6e/zoomable-bar-chart>

Graph Analytics -

We intend to assess and evaluate the impact of police officers with several allegations against them on other officers in the same unit in order to determine whether and to what extent they influence these officers. Graph analytics aids in the identification of the most influential and well-connected police officers. This can give us an early indication of an officer's behavior and how likely they are to influence other cops.

We filtered the officer data based on the ranks; the officers belonging to these ranks are nodes, and the edges represent the relationship between two officers who were co-accused. The subset of ranks selected are as shown in Fig 7. These ranks are obtained by arranging the total number of allegations in descending order and getting top 10 ranks associated.

rank	total_allegations
1 Police Officer	150977
2 Sergeant of Police	40677
3 Police Officer as Detective	28062
4 Lieutenant of Police	10351
5 Police Officer / Field Training Officer	4466
6 Police Officer Assigned Evidence Technician	2319
7 Captain of Police	2253
8 Commander	1481
9 Police Officer / Mounted Patrol Officer	952
10 Police Officer Assign Youth Off	916

Figure 7: Rank of officers used for creating subset

We applied two algorithms, namely triangle count and page rank for our analysis.

count	id	rank	officer_name	allegation_count
29636	3033	Sergeant of Police	Raimondo Brown	17
29633	6315	Police Officer	Terence Davis	38
29597	3744	Police Officer	Derek Campbell	8
25678	18042	Police Officer	Donald Mc Coy	22
25646	441	Police Officer	Fernando Alonzo	16
22131	21530	Police Officer	Michael Overstreet	56
21846	27349	Police Officer	Charles Stanton	11
21836	5180	Sergeant of Police	Stephen Conner	9
21824	5667	Police Officer	Jerry Crawley	30
21814	16747	Police Officer	Evetta Lundin	7
21812	8844	Police Officer	Thomas Flynn	19
21809	23654	Police Officer	Lloyd Reid	4
21809	14750	Police Officer	William Kissane	23
18570	19856	Police Officer	Ronald Muhammad	11
17717	8138	Lieutenant of Police	Glenn Evans	132
17220	29882	Lieutenant of Police	Fred Waller	49
17066	28273	Sergeant of Police	James Taylor	36
17014	5577	Sergeant of Police	Michael Cox	20
17012	28459	Sergeant of Police	Curtis Thomas	36
16982	32074	Police Officer as...	Ronald Jenkins	46

Figure 8: Officers placed in the descending order of triangle count

The triangle count defined as the ‘count’ column in Fig 8 shows police officers who are well connected with other police officers. Thus, it shows that Raimondo Brown, who is Sergeant of Police, is involved in misconduct on a much larger scale with a huge number of other officers and is well connected, which could potentially imply that he could be misusing his authority.

The page rank gives an idea regarding how likely a police officer is involved in a co-accused allegation. Fig 9 highlights that officer Robert Spiegel, who has the highest pagerank, is most influential amongst all the police officers who are co-accused.

id	rank	officer_name	allegation_count	pagerank
32350	Police Officer	Robert Spiegel	20	59.94255110579391
32351	Police Officer	Boonserm Srisuth	25	58.47588273428055
32419	Police Officer	Eric Wier	18	55.54706144219321
32284	Police Officer	Mark Reno	76	55.05553771973534
32384	Police Officer	Edwin Utreras	47	52.46541790094154
32337	Police Officer as...	Louis Silva	21	51.72556067035757
32413	Police Officer	Carl Weatherspoon	69	50.3834356160415
32289	Police Officer	John Rivera	44	49.31297984845296
32074	Police Officer as...	Ronald Jenkins	46	47.94275095359477
32375	Lieutenant of Police	James Triantafillo	31	45.85995467646986
32324	Police Officer	Kathleen Schmidt	15	41.26315879539915
32406	Police Officer	Reginald Ward	17	40.7491314934534
32383	Police Officer	Armando Ugarte	47	39.858307108258586
31906	Police Officer	James Davis	76	38.38390375443792
32378	Police Officer	Mark Trost	29	37.91554514385917
32355	Police Officer	Laurence Stiles	19	37.352703853575974
32341	Police Officer as...	Robert Slechter	12	36.72613447343105
32280	Police Officer	Eddie Randle	32	36.34386614017926
32339	Police Officer	Keith Sinks	20	34.26975822289742
32237	Police Officer	Louis Ortoneda	52	33.881065019348554

only showing top 20 rows

Figure 9: Pagerank distribution in tabular form, sorted by pagerank score.

We filtered the officer data based on the unit he belongs to; the officers who belong to these units as nodes, and the edges represent the relationship between two officers who were co-accused. These police units are those where the ratio of count of total allegations to count of total officers in the police unit is in the top 10.

Similar to previous analysis, we find additional officers with respect to the police unit they belong to who are most well connected and are co-accused in a higher number of allegations. Fig 10 shows that officer John Mc Gee of police unit 174 has the most connections to other co-accused officers. Thus, even if the police officer has fewer allegations against him, he is grouped with a high number of other alleged police officers from the same unit.

Fig 11 provides a pagerank view, which helps in identifying the officers with respect to their police units who are most involved in an allegation in the capacity of a co-accused.

count	id	last_unit_id	officer_name	allegation_count
7	18176	174	John Mc Gee	34
7	3395	173	Adam Burns	16
7	1172	174	Damen Balesteri	58
6	30537	174	Tracy Whitehead	19
6	10558	266	Anthony Granat	21
6	27385	174	Bill Starling	29
5	11641	174	Craig Hatch	7
5	12220	173	Courtney Hill	72
4	27439	173	Robert Stegmiller	62
4	2671	173	Sean Brandon	40
4	10663	174	Andre Green	36
4	23600	173	Brian Reed	23
3	2588	174	Jason Bradford	16
3	6221	174	Herbert Darey	25
3	867	174	Reginald Arrington	23
3	24324	174	Brandon Rodekohr	16
3	25191	173	David Salgado	31
3	7982	173	Kenneth Epich	45
3	29613	173	Marco Villarreal	10
2	24521	175	Carlos Rojas	71

only showing top 20 rows

Figure 10: Officers placed in the descending order of triangle count

	id	last_unit_id	officer_name	allegation_count	pagerank
	31996	175	German Gomez	12	1.305863820393211
	25191	173	David Salgado	31	1.1865471624739725
	27935	165	Matthew Swain	61	1.1847111268651542
	32406	173	Reginald Ward	17	1.1645157563285098
	24521	175	Carlos Rojas	71	1.1580852338143208
	27439	173	Robert Stegmiller	62	1.1577471014734335
	30320	180	Robert Weisskopf	23	1.1524574235387521
	29445	173	Luis Vega	53	1.1440847317012035
	30537	174	Tracy Whitehead	19	1.1299951650932252
	23600	173	Brian Reed	23	1.121360147183769
	30387	173	Mark Wesselhoff	12	1.119362425837314
	18176	174	John Mc Gee	34	1.1001864159642845
	32310	174	Bradley Ruzak	26	1.0931619943906823
	28335	180	Jesse Terrazas	60	1.069512583185627
	32235	174	Kevin Oneill	16	1.0671639477854686
	28734	175	Daniel Torres	9	1.0658958202979916
	27385	174	Bill Starling	29	1.0620978011392457
	10558	266	Anthony Granat	21	1.058339206340742
	19834	174	Todd Mueller	27	1.057729659160206
	32400	175	Adam Wallace	16	1.0402479378562848

Figure 11 : Pagerank distribution in tabular form, sorted by pagerank score.

Natural Language Processing -

Topic modeling is a statistical technique that is used to identify and discover distinct “topics” present in a series of documents or texts. We have tried to identify the different clusters of topics covered from the allegations that are registered against a police officer who belong to a rank that is in the top 10 ranks with the maximum allegation counts. The vision was further narrowed

down to target only those police officers who had a maximum number of complaints against them.vFor our code, we use the Bertopic algorithm. BerTopic is a topic modeling technique that uses transformers (BERT embeddings) and class-based TF-IDF to create dense clusters.

For the final model, the number of topics is selected as 6 which is then fit and transformed onto the allegations text. The updated cluster information is shown in Fig 12.

```
topic_model.get_topic_info()
```

	Topic	Count	Name
0	-1	1419	-1_report_parti_victim_polic
1	0	1248	0_alleg_accus_victim_complain
2	1	1111	1_report_parti_accus_state
3	2	859	2_offic_report_door_parti
4	3	587	3_report_parti_sergeant_arrest
5	4	468	4_vehicl_search_arrest_victim
6	5	465	5_parti_alleg_inventori_plant

Figure 12: Information on updated topics created

Fig 13 provides the words that make up cluster 4 followed by class-based TF-IDF scores. The target words vehicle and search signify the complaints are regarding illegal vehicle search.

```
topic_model.get_topic(4)
```

```
[('vehicl', 0.06328782794183578),
 ('search', 0.05911232458968046),
 ('arrest', 0.05799489396934193),
 ('victim', 0.04596237899730433),
 ('warrant', 0.02818013678229811),
 ('coerc', 0.028153415677698215),
 ('impound', 0.023704699922724427),
 ('permiss', 0.02341500840813096),
 ('detain', 0.023018069050091985),
 ('plaintiff', 0.019480062816899735)]
```

Figure 13: Analysis of topic 4 of model

Topic Word Scores

Figure 14: Topic-wise bar chart visualization

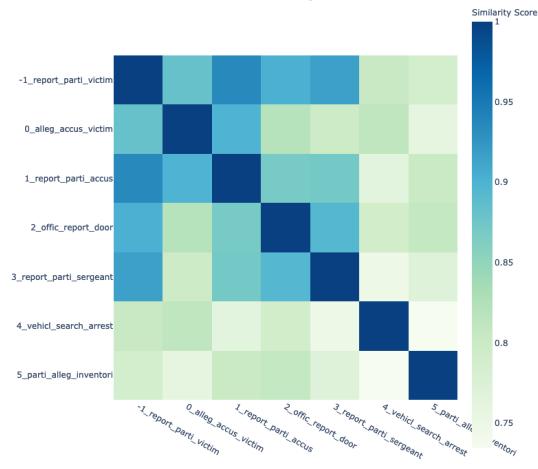
Similarity Matrix

Figure 13: Similarity matrix between different topics

On observing the bar chart in Fig 14, some topics seem similar. To visualize how similar topics are to each other, Fig 15 provides a heat map where darker colored blocks show higher similarity and thus the bright blue color signifies that topics 0 and 1 and topics 2 and 3 are similar to each other.

Conclusion and Future Work -

With the help of various checkpoints and experiments, we were able to view the existing problems within CPD. This includes racial bias when taking into account allegations that are sustained, as well as certain police officers who misuse their authority and the allegations against them are not sustained. We could also identify officers who influenced other officers in the same police unit and are co-accused in an allegation using Graph Analytics. This technique helps to identify those officers that are the main problems in the system and should be disciplined. Additionally, the allegation text is analyzed to create topic models. These topic models provide further insight in the allegation category assigned to each allegation. The topic models can also be helpful for the allegations that could not be assigned to a specific category.

Also, usually it is difficult to determine what could be an individual's next action. With the help of the topic model, a future scope of the project can be used to visualize if any trends are present in the behavior of a police officer based on the allegations received against him/ her to predict and prevent their next action.

Finally, knowing an officer's rank history may provide additional insight into how allegations against police officers were handled. If other officers from a unit had an influence on them, their unit history would also provide extra insight.