Project Report

OUR INVESTIGATION

The goal of our project is to investigate how the racial demographics of a neighborhood influence police misconduct within that neighborhood. Communities of racial minorities, particularly black communities in the southern regions of Chicago, are notorious for having high incidence of crime and consequently, being highly policed. Not only does this offer more opportunity for police misconduct in general, but it may influence base proportions and type of misconduct that officers engage in. We find it prudent to investigate these relationships in order establish a higher level of accountability for officers operating within these zones.

CHECKPOINT 1

Our investigation concerns the relationship between racial demographics and the nature of police misconduct, so our primary goal for this checkpoint was to establish some preliminary answers to questions that are more foundational to the investigation of the topic. Before we can dive deeper into the intricacies of this relationship, we need to establish some base level understanding of the statistics, such as how heavily policed different areas are, what kind of misconduct they commit, and how the racial demographics break down geographically. We chose to investigate these questions with the following queries:

1) What is the number of police officers per capita that are assigned to each police district?
2) What is the number of complaints per capita per police district?
3) What is the proportion of black residents in each police district?
4) Which type of complaint is most likely in each police beat?

Our analysis begins with question 3 where we broke down the proportion of black residents in each district. District ids 1537 (97.5% black), 1534 (96.8% black), 1535 (94.4% black), 1540 (94.2% black), 1547 (93.3% black), and 1546 (84.7% black) had the highest proportions of black residents in all of Chicago. On the other side, district ids 1542 (1.0% black), 1527 (3.3% black), 1531 (6.6% black), 1533 (6.8% black), and 1544 (9.1% black) had the lowest proportions of black residents in Chicago. In our analysis we will focus on these districts to find correlations between police misconduct and the proportion of black residents in each district.

Our first question is aimed at determining how heavily policed each area is. Each police district is broken into several beats which are typically patrolled by 8 or 9 officers at a given time. However, beat and district size vary greatly meaning that in different police districts, the number

of police officers assigned and the density of officers in a district is not consistent across districts. By looking specifically at the number of officers assigned to each district, we can see which geographic locations have higher concentrations of police officers. The highest policed districts were district ids 1546 (86.56 officers/100k residents), 1534 (61.91 officers/100k residents), 1535 (53.17 officers/100k residents), and 1549 (50.97 officers/100k residents). Of the four highest policed districts, 2 of them correspond to the districts with the highest proportions of black residents. Contrastingly, the least policed districts were district ids 1541 (17.73 officers/100k residents), 1542 (18.55 officers/100k residents), 1527 (18.74 officers/100k residents) 1543 (19 officers/ 100k residents), and 1531 (19.92 officers/100k residents). Of the 5 least policed districts, 3 of them also have the lowest proportions of black residents in Chicago. Thus, a high correlation exists between the proportion of black residents in a district and the number of officers assigned to that district, indicating the possibility of over policing in black neighborhoods.

Question 2 is an expansion of question 1, looking at how likely police are to get into an altercation providing a more nuanced perspective on how aggressive the officers are in their policing. The district ids with the highest allegations per capita were 1534 (0.289), 1546 (0.256), 1535 (0.182), 1537 (0.180), 1540 (0.157), and 1547 (0.152). These districts correspond to the six districts with the highest proportions of black residents in Chicago meaning there is a very high correlation between the number of complaints from police misconduct in a neighborhood and the proportion of black residents. This suggests that police are more likely to commit misconduct against black residents. The district ids with the fewest allegations per capita are 1531 (0.030), 1541 (0.037), 1542 (0.042), 1527 (0.047), and 1528 (0.054). Of these five districts, three of them have the fewest proportions of black residents in Chicago, reinforcing the conjecture that police misconduct is directly correlated to the proportion of black residents in the district. Districts with the highest proportions of black residents had 4.5 times the amount of complaints against police officers than districts with the lowest proportions of black residents on average.

For our fourth question we saw very little correlation between the demographics of a beat and the most frequent type of complaint. Across all beats except for three, operation/personnel violation was by far the most common complaints civilians had against officers. Although the number of complaints varied by beat, for example officers in beats with higher percentages of black residents tended to have more complaints, the actual complaints against the officer were extremely comparable to those in beats with low percentages of black residents. This meant that although officers had more allegations against them in black

communities, the type of misconduct in the Chicago police department is highly similar across the whole city. It would be interesting to see if police misconduct across the country followed similar patterns, or if this was unique to Chicago. If unique, CPD could reevaluate their training methods to possibly reduce their most common misconduct in the future.

CHECKPOINT 2

For our second checkpoint we investigated the relationship between the proportion of Black residents and the

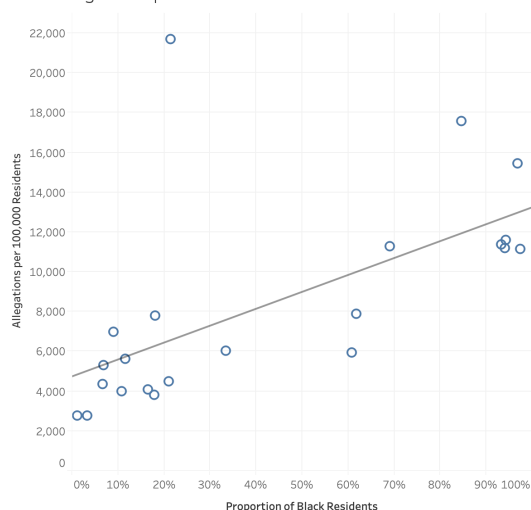Do areas with a higher proportion of black residents have more allegations per resident?

Fig 1. Scatter Plot of the relationship between proportion of black residents in a district and the number of complaints per capita

number of allegations in a neighborhood. The existence of a disparity between the number of complaints per person in black vs white neighborhoods is not necessarily surprising, but the magnitude of that disparity absolutely is. The result of this visualization provides foundational evidence to support the conclusion that racial demographics influence the way an area is

policed, and more specifically, the propensity for an officer to commit misconduct in black and white neighborhoods. This visualization not only legitimizes the investigation, but opens the door to new, deeper inquiries about the nature of police interactions with minorities. Some examples may include investigating disparities in what kinds of allegations are submitted in black vs white neighborhoods (are complaints in black neighborhoods more likely to regard violence?), how the rates of complaint submission by minorities relates to the proportion of minorities in a locality, or how the racial demographics of a locality affect the rate at which complaints are sustained.

For our second visualization, we produced a set of histograms that showed the breakdown of complaint type per race. These histograms clearly demonstrate that
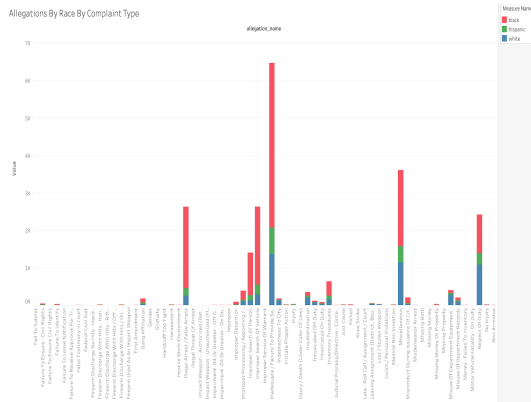
black residents make up the vast majority of

Fig 2. Grouped histogram showing breakdown of proportion of complaints submitted by black vs non-black residents for each type of complaint

allegations like improper search of a vehicle which have an extremely disproportionate number of allegations from black residents. This is in comparison to the already disproportionate amount of complaints we see across all allegations. In contrast, association with a felon has the most complaints from white residents. The previous example demonstrates that black residents seem much more likely to be the targets of vehicle searches than non-black residents. Similarly, black residents are almost 10 times more likely to be the victims of a false arrest than white or hispanic residents.

complaint types in nearly all categories. Interestingly, the distribution of complaints is not always the same. There are some

CHECKPOINT 3

In checkpoint 3 we looked at two visualizations of the data to answer questions about the types of complaints black residents are submitting. The stacked bar graphs compare the types of complaints balck residents are submitting in each police district to the types of complaints all residents are submitting. What is important to look at are not the numbers because many of the allegations have missing race data. Instead the proportions of complaints types in the two graphs tell a larger story. We assume that race data is missing randomly across all races so the proportions of complaints submitted by each race is consistent.
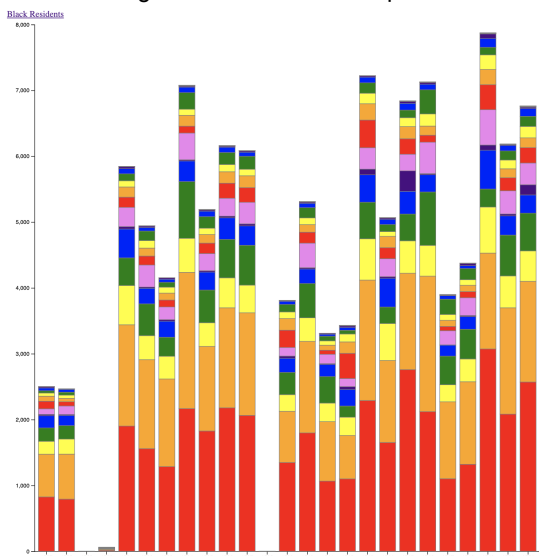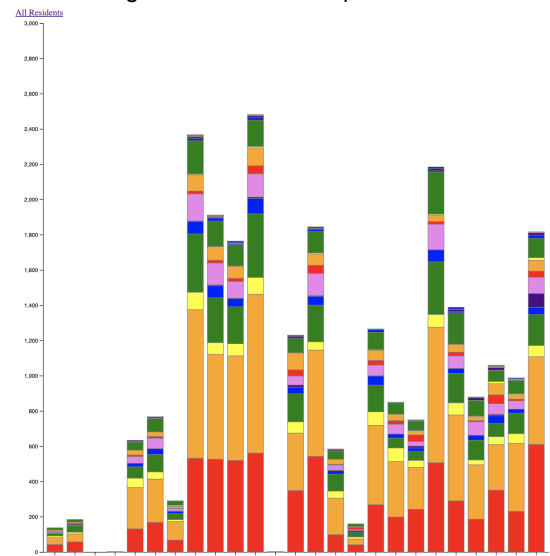
Fig. 3 Black resident complaints

Fig 4. All resident complaints

In the data for all residents, the top six complaint types on average are, Operational/Personnel Violations, Use of Force, Verbal Abuse, Traffic, and Lockup Procedures. The order of verbal abuse, traffic, illegal search, and lockup procedures are not always the same, but Operational/Personnel Violations is the most common complaint type in every beat except for beat 1533 which has 46 more Use of Force complaints and beat 1530 which has 1 more Use of Force complaint. These are the only instances in which Use of Force is first and Operational/Personnel Violations is second. Looking at the data for black residents, the top five complaint types on average are Use of Force, Operational/Personnel Violations, Illegal Search, False Arrest, and Lockup Procedures. Unlike for all residents, the complaints are highly concentrated in these 5 categories and the order of complaint type frequency is almost always in the order above. The data for Black residents is indicative of two phenomena. First, the higher complaint frequency of Use of Force compared to all residents suggest Black residents are more likely to be victims of physical force from police, whereas all residents are more likely to be the victims of verbal abuse. Second, illegal search complaints are much more frequent than for Black residents that the average of all residents. Because illegal searches complaints are typically filed when the officer searches an individual without enough prior reason, it seems as if Black residents are being targeted by the police disproportionately. Moreover, the high concentration of false arrest and lockup procedure complaints also reinforces the idea that Black residents are being targeted. Among all residents, false arrest complaints are not even within the top six most common, but among Black residents false arrest complaints are the fourth most common complaint type, suggesting that Black residents are more likely to be the victims of this specific type of police misconduct.

Similarly, by looking at the heat map we were able to come to a similar conclusion. The fewer the proportion of Black residents in a police district, on average the larger the gap between Use of Force and Operational/Personnel Violations complaints. We can thus again conclude that Black residents are more likely to submit complaints about Use of Force than all other residents. Because the heatmap works with numbers and not proportions of complaints, it is most helpful to compare complaint types within the same police district. By doing this we see a familiar pattern to the stacked bar graphs unfold. Districts with higher proportions of black residents are more likely to have proportionately more illegal search and false arrest complaints. Our two data visualizations paint a picture of police misconduct based on the racial demographics of a police district. Black residents are often the subject of different types of police misconduct than other residents are, and as a result neighborhoods with higher proportions of Black residents are more likely to have higher counts of certain misconduct complaints. We can see that Black residents are being policed differently than all other racial

demographics in Chicago. However, it raises the question about cause and effect. Are Black residents more likely to be on the receiving ends of use of force, illegal arrest, and illegal searches because they act differently as a result of their past mistreatment by police, or are they simply consistently targeted by police for these specific types of misconduct. Either way Black residents of Chicago are not being treated the same as all other residents by CPD.

CHECKPOINT 4

For checkpoint 4, the first graph is constructed with officers represented as nodes and officers mentioned in the same allegation represented as edges. By looking at the average clustering in different categories we are able to extract racial disparities from the data. For the sake of simplicity I will be focusing on Black and White complainants as they make up the majority of the data. Looking at the data by police district, the districts with the top three highest clustering averages are 10th (0.48), 11th (0.44), and 15th (0.42). All three of these districts are majority Black with Black populations of 57%, 74%, and 85% respectively–all Black majority districts. On average, the higher the proportion of Black residents in the police district, the higher the clustering average. Overall, Black residents have the highest clustering average (0.346). White residents have the second lowest clustering average (0.316) only just higher than the Hispanic clustering average (0.312). High clustering averages suggest officers are more likely to commit misconduct in groups against Black residents than they are against White residents. Officers committing misconduct in groups against a demographic that sees statistically higher misconduct is consistent with our findings from previous investigations and suggests an even more systemic presence of misconduct within the Chicago police department.

Our second question was aimed at further uncovering differences in Black vs non-Black complaints against the police. Each node in this graph represents a complaint type and each edge represents an officer involved in both types of complaints. The graph looks specifically at officers who have offended in more than one category. By constructing the graph in this way, we are able to determine if officers involved in multiple different complaint types generally have the same groups of complaints filed against them. For example, do officers with high use of force complaints have a higher propensity to also have illegal search complaints filed against them? After querying the data, we determine White complainants have 104 edges between 19 allegations with an average clustering of 0.84, Black complainants have 132 edges between 19 allegations with an average clustering of 0.90, Hispanic complainants have 97 edges between 17 allegations with an average clustering of 0.82, and Asian/Pacific Islander complainants have 6 edges between 13 allegations with an average clustering of 0.84. For simplicity I will be

focusing mainly on the discrepancies between the White complainants and the Black complainants as they make up the majority of the data.

White complainants have 104 edges, meaning there exists 104 connections between two complaint types from the same officer. Black complainants have 132 connections between two complaints from the same officer. This data suggests that not only are officers more likely to commit misconduct against Black residents, but each officer is more likely to commit a wider range of misconduct against Black residents compared to White residents. However, the average clustering for White complainants is 0.84 whereas the average clustering for Black complainants is 0.90. This means the types of allegations coming from Black residents are more concentrated in the same allegations than they are for White complainants. Thus, although officers are more likely to commit a wider range of police misconduct against Black residents, the majority of the type of police misconduct is still concentrated within a smaller number of allegation types than those against White residents. Because both the number of connections and the average clustering is higher for Black complainants than White complainants, Black complainants must experience higher levels of police misconduct across the board and experience higher levels of police misconduct concentrated in specific categories like use of force and false arrest (from checkpoint 3) than their White counterparts.

This graph suggests that Black residents are targeted in specific ways (use of force etc.) as the clustering rates for officers with multiple allegations are higher for Black residents than any other racial demographic. The graph also suggests that Black residents are victims of higher counts of police misconduct when compared to other demographics because of the higher count of edges meaning either 1) officers are committing misconduct in a larger number of categories against Black residents than any other racial demographic group or 2) more officers are committing misconduct in more than one category against Black residents than any other racial demographics group.

The two graphs tell a similar story to one another and are consistent with the data we have analyzed so far. Black residents are more likely than any other racial demographic to be on the receiving end of police misconduct, and are likely to have higher rates of misconduct in specific categories. The higher rates of officers committing misconduct in multiple categories against Black residents raises the question of why. Are there specific reasons officers are more inclined to commit a broader array of misconduct against Black residents or is it simply a function of higher complaints rates for Black residents? Either way, our data continues to portray a pattern of police misconduct specifically against Black residents.

CHECKPOINT 5

For our checkpoint 5 we wanted to analyze the sentiment of complaints narratives specifically in relation to the proportion of Black residents in the neighborhood. As we have seen throughout our project, statistically significant discrepancies exist in the way different racial groups are policed–in particular we have observed Black residents are far more likely to be victims of police misconduct. With this checkpoint we wanted to investigate how officers perceive interactions with the residents who have submitted complaints about them through their complaint narratives. After running the sentiment analysis for White, Black, Hispanic and Asian residents, we have the following four graphs with the proportion of residents that are the selected race on the x-axis and the mean sentiment for the selected race on the y-axis:
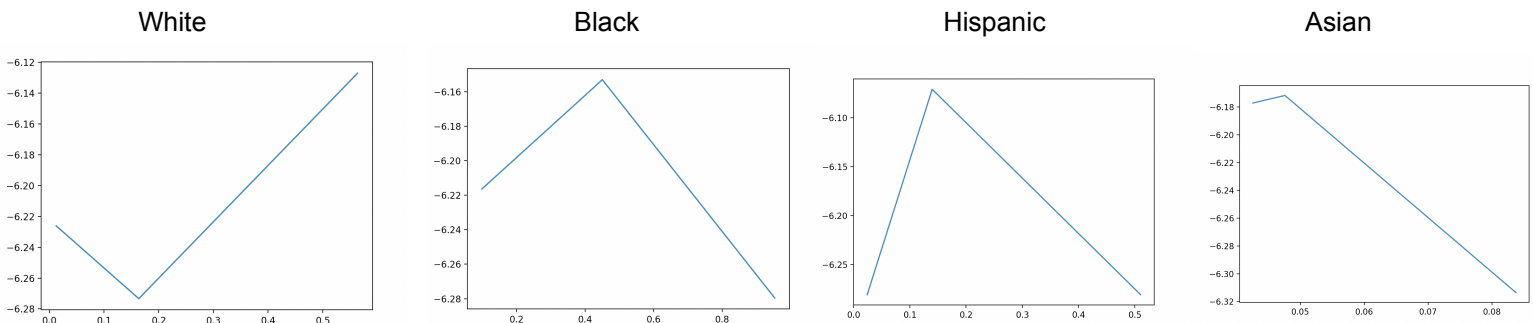

Fig 5. Plot of proportion of racial representation against average allegation sentiment. Each data point represents a group of police districts. Groups were determined such that each one shared a roughly equal number of allegations.

We focus primarily on the White, Black, and Hispanic data as they make up the majority of Chicago's population. Neighborhoods in Chicago only range from 0-8% Asian, so without sufficient data no correlative conclusions can be drawn. Overall, complaints have negative narratives with none reaching about -6, but this is expected as they are complaints being filed using supposedly strict, unbiased language. Thus we would expect the sentiments to be relatively negative to begin with. In the White data, the average sentiment of the complaints starts low reaching about -6.275 around 17% white and then trends upward from there. Thus, we observe neighborhoods with higher proportions of White residents have more positive average sentiments from the police and sentiment increases positively as the proportion fo White residents increases. The slight decrease in sentiment from 0%-17% White neighborhoods can most likely be explained by the officers perception of the neighborhood as a whole. Although the individual is White, the officer's perception of the neighborhood most likely influences the complaint narrative written by the officer.

In the Black and Hispanic data we see similar trend lines. Both start with relatively negative average sentiments, then have a peak and a sharp decline. For Hispanic residents,

the sentiment value peaks around -6.7 at a Hispanic proportion of 12%. For Black residents, the sentiment peaks around -6.15 at a Black proportion of 43%. After the peak both Hispanic and Black residents have declines in their average sentiment value to around -6.28. However this happens for Hispanic residents around 0.5 on the x-axis whereas it happens to Black residents around 0.8 on the x-axis. Although it happens later for Black residents, in both the Black and Hispanic graphs, have the lowest average sentiment at the highest represented x value indicating a strong negative correlation between sentiment and the complainant race.

Again we see a targeting of racial minorities by the Chicago police department. Interestingly this targeting is not only against Black residents, but also Hispanic residents suggesting a stronger White vs non-White correlation instead of the Black vs non-Black relationships we have analyzed previously. Additionally, it is possible their complaints are not deliberately more negative towards non-white residents. Instead, officers subconsciously use more negative language towards non-White residents. If this were the case it would reveal a larger picture of systemic prejudices held by the officers as the prejudices they project are subconscious. Another interesting observation is the decreasing portion of the White trendline from around 0.0 to 0.7. Because Chicago remains highly segregated, it is possible the sample size in neighborhoods with less than 10% of residents identifying as White is very small leading to this oddity. However, no clear answer exists.

DISCUSSION

Through our analysis we confirmed some of our original hypotheses. We found that the demographics of a neighborhood strongly influence the police misconduct within that neighborhood. However, the results and reasons are not as straightforward as they may appear. We found high disparity in the types of allegations submitted across races, namely that Black residents were far more likely to be subjected to Use of Force, False Arrest, and Illegal Search, demonstrating a propensity for physical allegations against Black residents. Moreover, in Checkpoint 4 we found officers are more likely to commit misconduct in groups against Black residents than they are against other races. Combined with our strong results from Checkpoint 5, the misconduct within the Chicago Police Department seems not only to be targeted, but also systemic. We have no way of telling whether or not officers are deliberately committing misconduct against Black residents or if they simply hold implicit biases, but either way the pattern of misconduct portrays deep rooted issues that need complete overhauls to reduce this disparity in the way CPD officers police Chicago neighborhoods.