

Plot 1: I want to understand the relationship between the racial distribution of officers in a district with the racial distribution of people that live in the district. Is there a discrepancy between the folks that police a district versus those who live in that district? I am going to do this with a choropleth map that is colored by officers per 10k residents. It will then have a popup for each district and show the racial distribution of officers. This will show the relationship between officers per capita and the racial make-up of officers in a district.

The process of answering this question included many steps. First, I was able to build off of the SQL code from Checkpoint 1 to get the officers per 10,000 residents rate (abbreviated as officers per 10k residents). The next part I needed was the racial distribution of officers for each district across the currently active officers. I got this by combining different tables such as data_officer and data_policeunit, which contain the officer races and their units. There was some trickiness in the way that we had to not only find this from the table, but also ensure that each district represented all 6 of the racial categories even if none of the officers belonged to that group. We added rows for which counts were equal to zero if no officers were in that group. From here, I started working with observable hq to figure out how to create a choropleth plot. The best resource I found was [this one](#) from an old classmate's project, which contained a choropleth with pop up values when you hover over. Along the way, I realized that the best thing I could do was match my data in the format they had used in their notebook. To this end, I went back into DataGrip to convert my table into a similar format. I pivoted the table and made sure to transform the polygons into appropriate geojson geometries. This allowed me to input them into the notebook nicely. In the end, I created a plot to show the ratio of officers per 10,000 residents by the Chicago Police district. If you hover over an area, then you can view the racial distribution of officers in a district. Each racial group has a percentage with the combination adding up to 100%. The district number and the actual count of officers per 10k residents is included in the popup.

In relation to the initial question I asked, the plot shows that there is some relationship between the location of districts, officers per 10k residents, and the racial distribution of officers. At first look, there is not a super clear relationship between officers per 10k residents and location. What I mean by this is that similar rates of officers per 10k residents are not extremely obvious when comparing the north and south sides of Chicago. My examination of the split between north and south is based on the history of Chicago redlining and the stories told about differences in life between north and south. In this case, I am using districts 1, 11, 12, and 15 as the boundary between the north and south side of Chicago. Upon second look, there is a slight pattern. Both airport areas have the largest rates of officers per capita and are the lightest shade of blue, since they have fewer than 20 officers per 10,000 residents within each district. Outside of those two districts, rates between 20 to 70 are more common on the south side, while rates from 10 to 50 are more common on the south side. Interpreting the officers per 10k residents rate is to say we divide $10,000 / (\text{number of civilians} / \text{number of officers})$ for each district. This scales the officer per capita rate such that as the rate increases, then the more officers there are for each 10k set of residents. Overall, what I learned from this is that there is a slight pattern in the number of officers per 10k resident and location. The north side of Chicago has fewer officers per resident, while the south side has more officers per resident.

Another layer to this plot is the racial distribution of officers. Given the slight pattern I mentioned above, there is a similarly associated pattern with the percentage of white or black officers and the north and south side of Chicago. Districts on the south side have higher percentages of black officers although the percentage of white officers is almost always the largest. The only districts where the percentage of white officers is not the largest are the 2nd, 3rd, 5th, 6th, and 7th, which are all on the south side. This means that districts on the south side not only have more officers per resident, but those officers are not primarily white (across most of the south side). This is all to say that the officers per 10k residents rate is related to the racial makeup of a district. A potential outlier or extreme case of this is the district that includes O'hare airport, which has the largest number of officers per capita of which 93.8% are white.

Plot 2: To supplement plot 1, I will create a bar chart that has race on the x-axis and percentage of people on the y-axis. It will have two bars for each race category: one for officers and one for residents. It will have a selection drop down for the districts.

The process I did to complete this question included pulling from the past checkpoint and using elements from plot 1 in this checkpoint. Here, I had to find the racial distribution of the district residents which I calculated for specific districts in checkpoint 1. I generalized this for all districts then I had to do a similar calculation to part one to ensure that all racial categories were accounted for in each district. I tried to combine this with the racial distribution of officers in each district using what I calculated in part 1, but the category names were spelled differently. I cleaned the data to ensure similar names before joining the two tables and exporting that version. This gave me a table where I had the district, racial category, percentage of officers for that group, and percentage of residents for that group. This was only the SQL part. I struggled a lot with the d3.js code for this section, but I was able to get it to work and be interactive. I created a bar plot which uses racial group on the x-axis and the percentage of the group on the y axis. It compares the racial distributions between officers and residents in one specific district. To specify a district, select one of the options on the drop down menu above the graph. In the plot, the blue bar is for the percentage of officers in that racial group and the orange bar is for the percentage of residents.

The goal of this plot is to compare at the district level the racial distribution of officers with the racial distribution of residents to see if the officers are representative of the districts they serve. As I noted in part 1, the districts on the southside with a larger percentage of black officers are districts: 2nd, 3rd, 5th, 6th, and 7th. Looking at these districts, the 2nd district had very similar distributions between officers and residents. The 3rd district had slightly less similar distributions since there were more white officers than white residents in the district. The 5th, 6th, and 7th districts had more significant differences in the number of white officers compared to the residents in the districts. Other districts on the south side such as district 8 have larger (57%) percentages of hispanic residents, without corresponding rates of hispanic officers to serve the neighborhood, while white officers make up 86%. In comparison, on the north side districts 20 and 24 had similar trends in terms of white officers making up the largest group as well as white residents. However, in these districts there were significant sizes in the percentage of all racial groups without the corresponding representation in officers. Notably, district 25 on the northside has a large group (66%) of hispanic residents with primarily (85%) white officers, which parallels district 8 on the south side.

Themes: Think about the broader picture, how this fits into the theme of your project, what you learned, and any future lines of inquiry it reveals.

In relation to my theme of the impact of police presence on community members, I was able to center my questions around that story. I tried to address issues around the degree of police presence in districts and the racial representation of that presence. What I saw was that the rate of officers per 10,000 residents was lower in northern Chicago districts than in southside districts, which may indicate a lesser degree of police presence on the northside of the city. This is to say that there is a larger police presence on the southside of Chicago. Of these officers, it seems there were more diverse racial distributions of officers on the south side than on the north side with respect to black and white groups. However, that does not mean the racial distribution of officers was representative of the residents in the district. Rather the distributions between these two groups were only very similar (seemingly representative) in the 2nd district, while most other districts had large percentage differences between groups. As a whole, the visualizations I created help me understand the ways in which officers exist across districts and how they are distributed in terms of race.