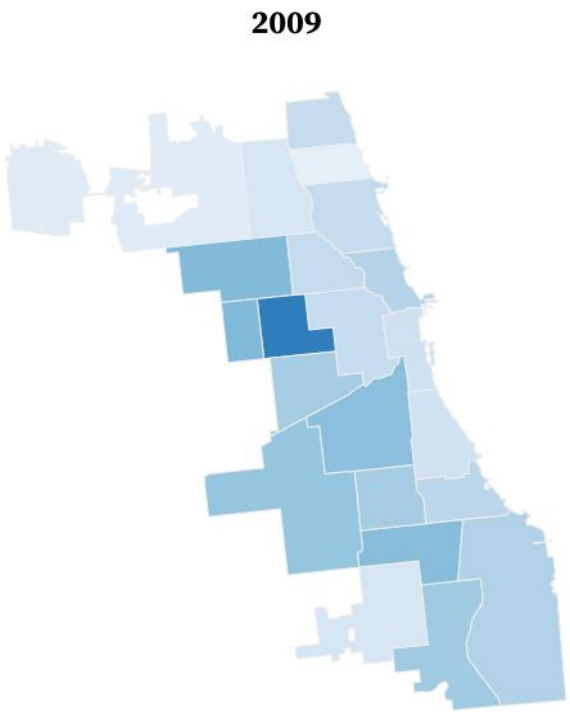
Faraaz Beyabani, Andrew Vagliano
November 3, 2020

Checkpoint 3: Interactive Visualization

For this checkpoint, our goal was to iterate on our analyses from checkpoint 2, which were very civilian focused, by examining police tactical response report (TRR) counts geographically and temporally. Geographically, we looked at the TRR contribution from each police district as a percentage of the total number of TRRs filed across Chicago. Temporally, we examined this metric for every year available to us from the CPDP trr_trr table, from 2004 to 2016. The key questions that we were curious about included:

- Are TRR percentages for each district relatively stable over time?
- How do TRR percentages relate to the median income of each district?
- Are there higher TRR percentages in districts where the population is predominantly a specific race?

While the yearly data may not match up to analyses conducted in previous checkpoints, we juxtaposed the TRR map created with median income and population race maps created in checkpoint 2 to help facilitate our analysis. Our hypotheses for the questions above were as follows: each district should have a similar TRR percentage from year-to-year; TRR percentage and median district income are inversely proportional; and districts with majority Black populations will have higher TRR percentages.



Above is a snippet of the map created for 2009. Districts with a higher TRR percentage are colored a darker shade of blue. Before proceeding, it is important to mention that, similar to our analyses for checkpoint 2, the 31st District has been omitted from our data. This is due to likely erroneous geographic data defining it at both the north and south side of Chicago. No meaningful data was lost from this action.

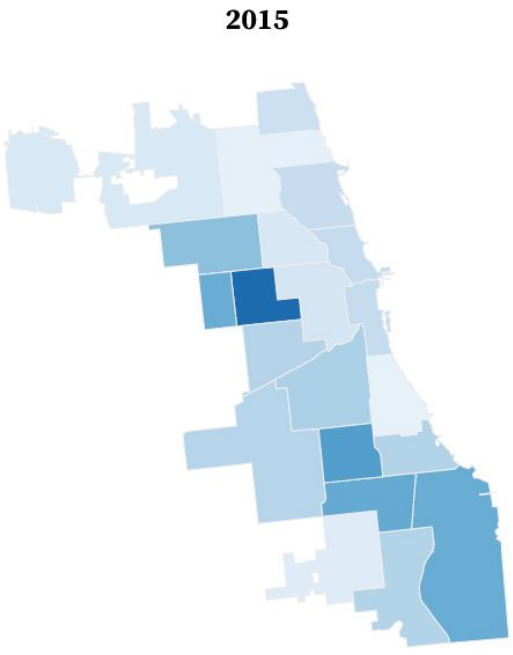
Question 1: Are TRR percentages for each district stable over time?

We found that most districts do stay relatively stable in terms of TRR percent, within some deviations of 1% per district. An example of this would be the 16th District, which has a TRR percent range from 1.71% to 2.25%. Other districts can change as drastically as 2-3% or more. For instance, the 11th District is consistently the district with the highest TRR percentage with a range of 9.1% in 2007 to 12.51% in 2015. Factors that could explain this disparity could include district population, income, department changes, current news (see Laquan McDonald), or even other social phenomena such as white flight.

However, despite the sometimes large variances in TRR percentages from year to year, many districts' patterns remain accurate throughout the entire data. That is to say, districts do not change quartiles over time; if a district is in the first quartile, then it stays in that quartile across the whole dataset. Therefore, the data remains predictable and we can extrapolate from this data for our purposes.

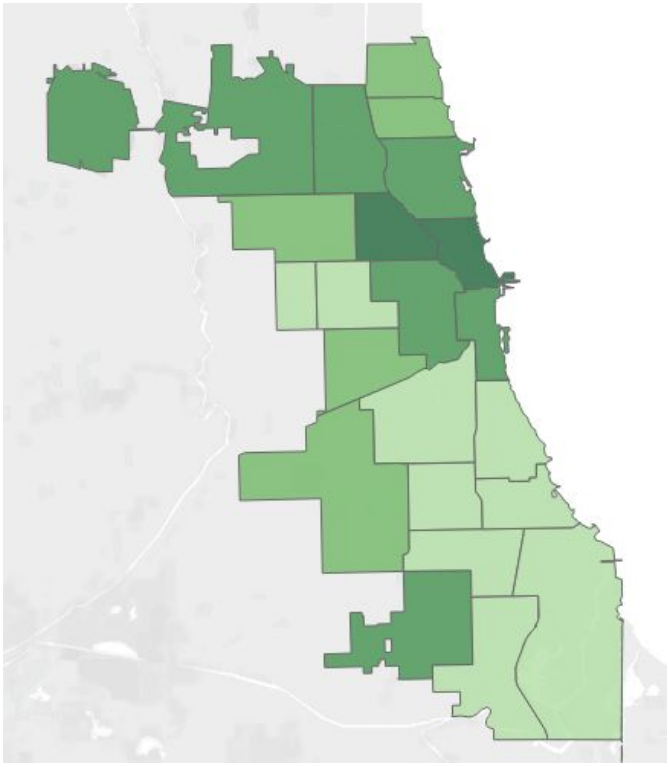
Question 2: How do TRR percentages relate to district median income?

For this analysis, we did not have historic socioeconomic data, so we opted to use the 2015 TRR data, as it felt the most complete and was still somewhat recent. By comparing this



11th District: 11.56% of total Chicago TRRs

data to the median income data in the CPDP database (pictured below), we can see that *roughly*, areas with lower median incomes do in fact have higher percentages of TRRs filed. The inverse is also true, where districts with higher median incomes have a lower percentage of TRRs filed. This supports our hypothesis very well, with all of the districts adhering to this pattern very closely. Furthermore, since this relationship mimics the one between allegation count and median income, it is likely that areas with low median incomes are overpoliced, either through an increased number of officers on patrol or through more aggressive officers escalating situations.

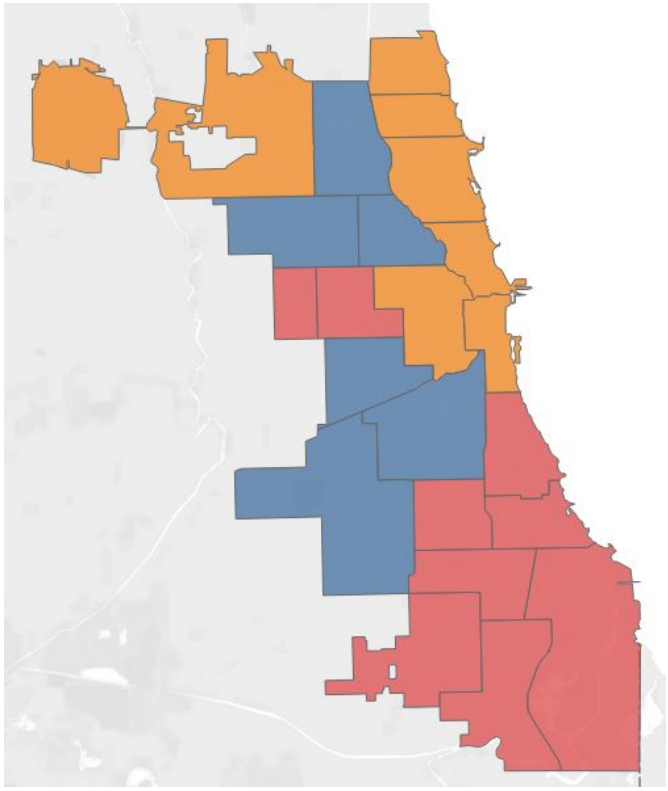


Question 3: Are there higher TRR percentages in districts where the population is predominantly a specific race?

To answer this question, we once again cross-reference our TRR visualization with a previous visualization examining the majority race of the civilian population for each district. Our hypothesis was that districts with large Black and Hispanic populations would also be the districts that contribute a large percent of Chicago's TRRs. For the most part, this conclusion is accurate and matches the data.

The map for population race data is shown below. Districts colored red are majority Black, blue are majority Hispanic, and orange are majority White. For the most part, we can see that districts that are predominantly Black in the midwest and southeast also have very high TRR percentages. While this may simply be the case because Black districts are highly correlated with low median income districts, it is also the case that some Hispanic districts also have a somewhat high TRR percentage. Moreover, these districts do not necessarily have low

median incomes, which further increases the believability that TRR percentage and district race are somewhat correlated.



Overall, while this metric is a little less reliable, we would say that our findings do somewhat support our predictions for this question and for this exercise in general. Specifically, the districts we would like to point out as potentially being overpoliced are Districts 4, 6, 7, 11, 15, and 25, which can be easily sought out on the ObservableHQ map. These districts all have high allegation counts, high TRR percentages, and low median incomes. They are all also predominantly non-White, which further supports the conclusions made in this and previous checkpoints.

However, to truly make sure of our theory, more work can still be done. Moving forward, our plan is to use machine learning to predict the probability of an individual being involved in a police incident for two similar districts, such as Districts 4 and 22. By performing this experiment at an individual level, rather than district level, we can potentially tie in prior data analyses by forming relationships between officer race, victim race, and TRR and/or complaint type. If this process shows that a victim's race has no bearing on the aforementioned probability, we could attribute the varied TRR percentages across Chicago to district population count, rather than district race.