CS 396/496 Data Science Seminar
Quick Turkeys - Lucas Yuan, Jiren Li, Renzhi Hao

Final Report

Introduction:

For our project, we are investigating how neighborhood demographics play a part in what types of force police officers use, and in order to accomplish this we were looking at the Tactical Response Report (TRR) data. This data contains a variety of tables with fields such as firearm_used, taser_used, and details about the officer and the subject that we can take a look at to see where and how police officers use force. We also want to take a look at which types of officers have the most TRRs and whether there is any relation between these officers.

Some examples of questions we looked into would be , is an officer more likely to use a weapon against people of different races? We were also looking into what types of officers commit TRRs the most. We want to look into this in order to look further into how police officers use violence, and what situations/locations they are more or less likely to use violence in.

Checkpoint 1 - Relational Analytics:

We answered five questions using SQL Analytics:

Which race is most commonly the subject of a TRR as compared to the racial composition of that neighborhood?

We compare the predominant race of composition of every area that exists in data_racepopulation(which is 99 in total) as well as its predominant subject race of TRRs that happened inside of that area.The result reveals that most of the TRRs have the predominant race of “Black” while most of the predominant race of the related areas vary from“Black, White and Hispanic” with few “Asian”. For predominant neighborhoods that have “Black” as predominant subject, there is not a race of areas that has a critical amount ahead between “Black, White and Hispanic” related to them. For neighborhoods that have “Hispanic” as predominant subject, the neighborhoods are with high possibility to have “Hispanic” as the predominant race. For neighborhoods that have “White” subject, the neighborhoods most likely also have “White” as the predominant race. There is no TRR that has “Asian” as the predominant race. In all, the Hispanic and White people gathered areas are usually to have the most TRR of the same predominant race while the likelihood of Black people as the subject of TRRs doesn’t relate to the predominant race of areas.

How do TRRs break down at the intersection between the racial demographics of officers and the racial demographics of the subjects of the TRRs?

From our analysis here we found that the most likely combination that created a TRR was white officers against black subjects at 42.23%, followed by hispanic officers against black subjects which was 15.94% and black officers against black subjects at 14.37%. However, this likely occurs because there are the most white officers. However, based on the probability matrix we constructed using the officer demographics and the Chicago population demographics we find that the most likely occurrences should be White officer on white subject at 19.29%, white officer on black subject at 18.5%, and White officer on Hispanic subject at 17.4%. The second and third most common occurrences, hispanic officers against black subjects should be at 4.53%, and black officers on black subjects should be at 7.01%.

How often firearms/tasers are used against subjects of different races?

We found that both are used much more often against people of color as compared to white people, with Asians having the highest percentage for both at 2.78% for firearms and 7.19% for tasers followed by Hispanics for firearm use at 2.40% and Blacks for taser use at 7.06%. However, this could be a result of small sample size as for Asians there were only 431 total TRRs and 12 and 31 uses of firearm and taser respectively. On the other hand, blacks have the most TRRs against them and thus cannot fall under this same. White people were the lowest at 1.42% for firearm use and 6.04% for taser use.

What is the number of excessive force allegations as compared to the number of trrs?

The number of trrs/The number of excessive force allegations=2.1, which means one complaint is made for every 2.1 trrs. In fact, considering that some police officers may conceal their use of excessive force, the real number of TRRs may be larger than 67,019. However, according to the survey results of Invisible Institute, many subjects do not decide to make a complaint even if they are treated with excessive force. Therefore, the real ratio is probably smaller than 2, which means more than a half of used force is inappropriate and excessive.

Which demographics of officers have the most TRRs?

White police officers have a similar proportion in all officers compared with the trrs raised by them. Asain/Pacific, Native American/Alaskan Native and Hispanic police officers raised a larger proportion of trrs than the proportions of their population. Black police officers seem to use less force than other officers.

82% of police officers are male, and they account for 88.6% trrs, which indicates that male officers tend to use more force than female officers.

Police officers of most ranks raised a similar or smaller number of trrs compared with their population. However, the proportion of trrs raised by sergeants of police is nearly twice of the proportion of their population. It is possible that police officers with higher rank than sergeants seldom do field work, so maybe senior officers tend to use more force than junior officers.

Overall throughout the exploration of the data, we've found that there is consistently more violence used against both communities and people of color, with black people in Chicago being subject to more TRRs than any other race by a wide margin. We can see this in which communities are being policed, and how often tasers and firearms are used against officers. We also found that more than a half force use is complained by the subject and some demographics of officers tend to use force more often.

Checkpoint 2 - Data Exploration:

Which race is most commonly the subject of TRRs with the extent of the force used?

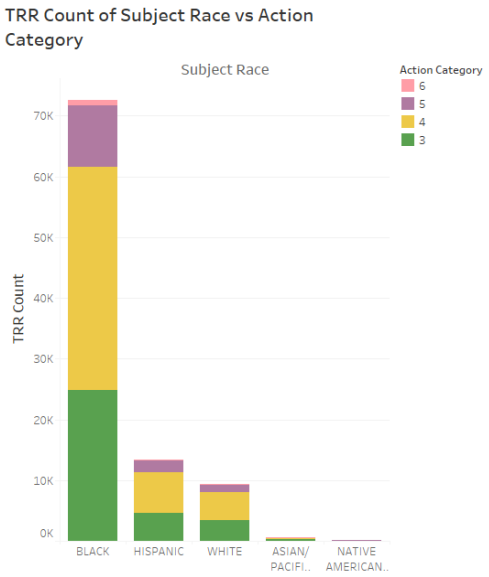

Figure 1

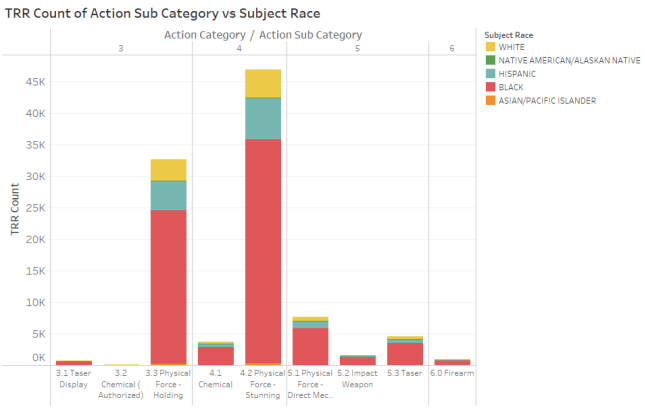

Figure 2

It shows that the black race is the most commonly the subject of a TRR in general which reaches over 70K, and also is the predominantly most commonly the subject of a TRR in all 4 action categories that involve forces (the 0 and 1 don't involve forces and 2 is not classified descriptions).

To be specific, we created a stack bar graph for subject race vs action sub category. As shown from the graph below, the 4.2 physical force-stunning is the most commonly the extent of member action for all the races and reaches over 45K with the black race taking the most part of it.The second is 3.3 physical force-holding which reaches over 30K with the black race taking the most part.

What is the number of excessive force allegations as compared to the number of trrs?

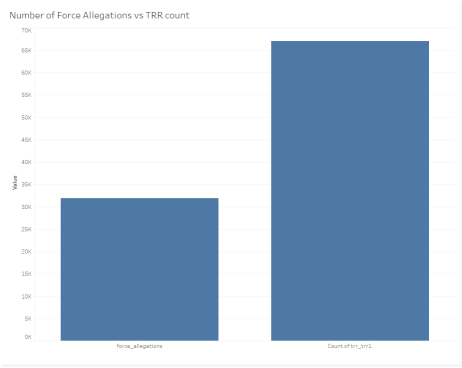

Figure 3

Figure 3 is a graph that visualizes the amount of allegations as compared to the number of TRRs, as described in the fourth question in Checkpoint 1. However, this data doesn't tell the whole story since some of these use of force allegations may not be legitimate complaints. On the other hand, people who have legitimately had force used against them may be unwilling to step forward due to a variety of societal factors. Based on this, we are unsure as to whether or not use of force allegations are more or less likely than the data would indicate.

What locations are TRRS most likely to occur in?

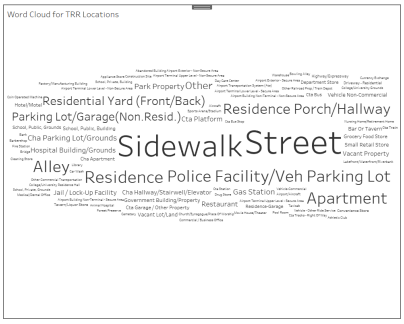

Figure 4

Figure 4 is a word cloud that shows which general locations TRRs are most likely to occur in. We don't have anything conclusive to draw from this since most of these track with where we think police activity is likely to occur.

Create a chart for which demographics of officers are most likely to have TRRs.

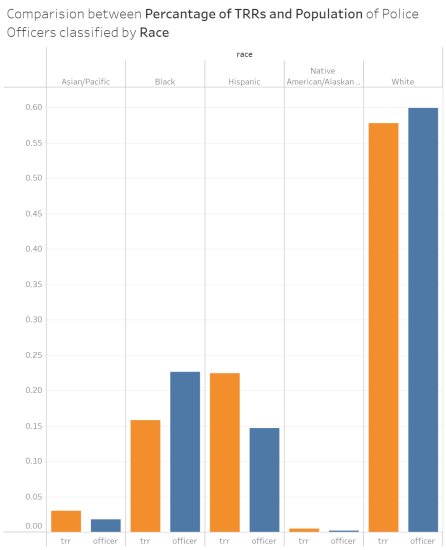

Figure 5

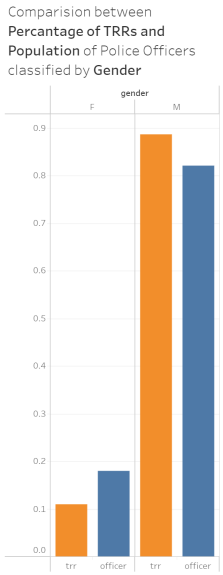

Figure 6

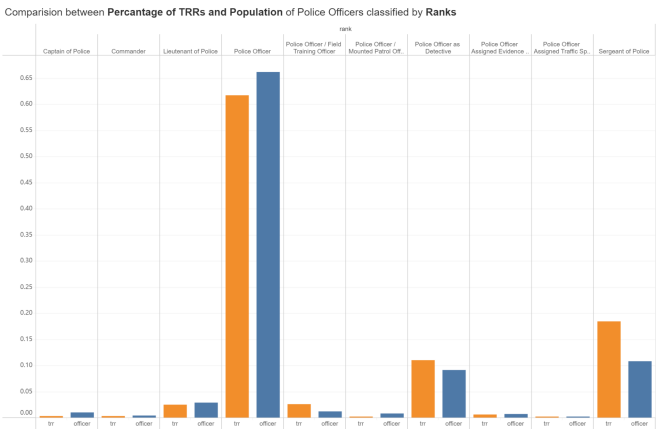

Figure 7

From Figure 5 we know that although white officers raised most TRRs, the proportion is smaller than the proportion of their population, while Hispanic and Asain/Pacific officers raised more TRRs than they should have raised, and Black officers seem to be the least likely to use force.

From Figure 6, we can see that male officers are more likely to use force.

Figure 7 shows that ordinary officers are less likely to use force than other officers, while sergeants and detectives have higher levels of TRRs rates. For detectives, it may be understandable because they may work in severe conditions. However, for sergeants, maybe it means senior officers are more likely to use force. Although captains, lieutenants and commanders are less likely to have TRRs, it is probable that these police officers have fewer chances of field work than other officers.

Hispanic and Asain/Pacific officers raised more TRRs than other officers while black officers had fewer TRRs. Male officers are more likely to use force than female officers. Senior officers may tend to use force more often. Black race tend to receive more forces related misconducts in any kind of subcategory of forces.

Checkpoint 3 - Interactive Visualization:

How do TRRs break down at the intersection between the racial demographics of officers and the racial demographics of the subjects of the TRRs?

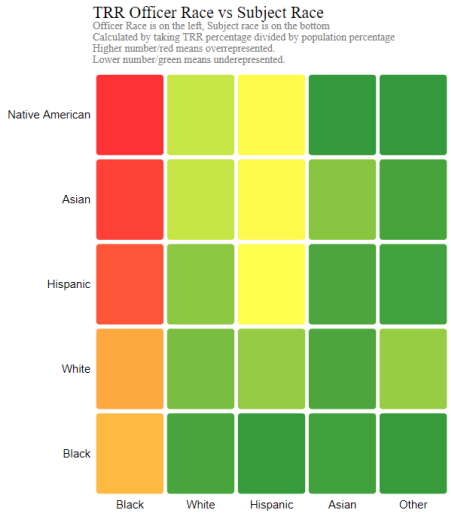

Figure 8

Figure 8 is a heatmap based on the data that we discussed in the second question of checkpoint 1, and green represents underrepresented while red represents overrepresented with officer race on the left and subject race on the bottom. Hispanic subjects being accurately represented when faced with Native American, Asian, or Hispanic police officers with values of 1.09x, 1.12x, and 1.03x the amount of TRRs you would expect based on demographic data. However, we can also see that on the left black citizens are heavily overrepresented compared to their population, with similar values between black officers at 2.05x and white officers at 2.28x, but spiking with Asian, Native American, and Hispanic officers with multipliers of 3.51x,

3.78x, and 5.51x. However, these values may be caused by a smaller sample size with 71 and 614 police officers for Native American and Asian cops respectively.

Create a heatmap of TRRs by neighborhood by year (which neighborhoods have the most TRRs per year).

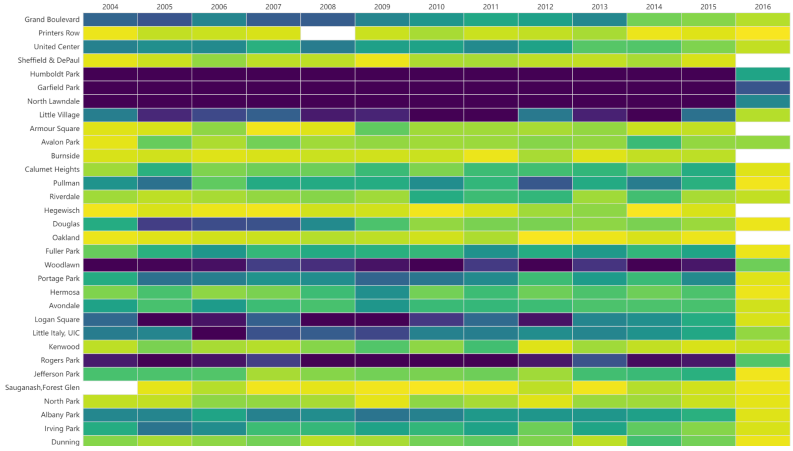

Figure 9

From Figure 9, we know that the number of TRRs in a certain neighborhood is relatively stable during the decade: neighborhoods with large numbers of TRRs continue to have it, while neighborhoods with few TRRs in 2004 did not change a lot in 2015.

It is noticeable that 2016 has a small number of TRRs compared with other years. It is most likely because TRR data from CPDP does not cover the whole 2016. Therefore, data of 2016 is not of much reference value.

Furthermore, for clear discrimination of most grids, we cannot make more subdivisions for numbers larger than 100, so “all-purple” neighborhoods look the same. However, in fact some purple grids represent numbers larger than 500, and these “all-purple” neighborhoods always reached a similar number in the decade. It is incredible that some neighborhoods have a hundred times more TRRs than others.

From Figure 9, we know that variance in the number of TRRs between different neighborhoods is huge. From the color of grids in the heatmap, we can see that most neighborhoods had less than 50 TRRs per year, and some even did not have a single TRR in a whole year. However, some “all-purple” neighborhoods always had more than 100 TRRs a year.

Checkpoint 4 - Graph Analytics:

Are groups of officers working together a lot for TRRs?

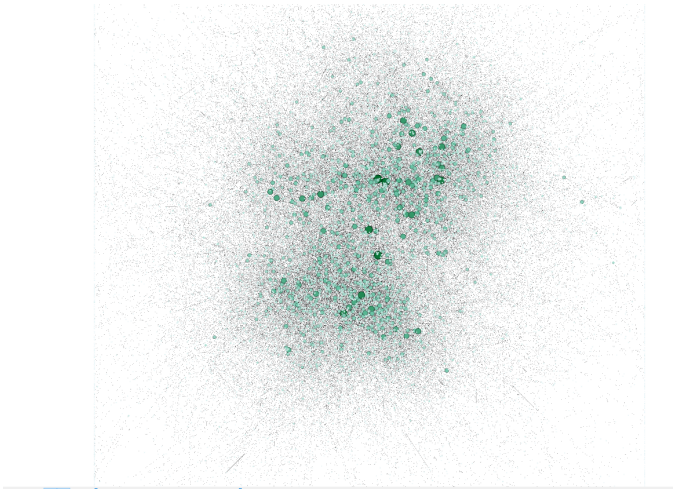

Figure 11

Figure 11 is a graph that has officers as vertices with the size based on number of TRRs, and the edges as TRRs which is done using the shared event_ids in the TRR data. The graph shows a lot of officers with few TRRs spread throughout the graph, but officers with many TRRs are clustered together in the middle in two semi-connected clusters. Through this we can see that there are not officers that commit large numbers of TRRs that are isolated from the other officers.

Are there differences in the properties of the graphs of officers who are co-offenders, as compared to the expected officer deployment patterns?

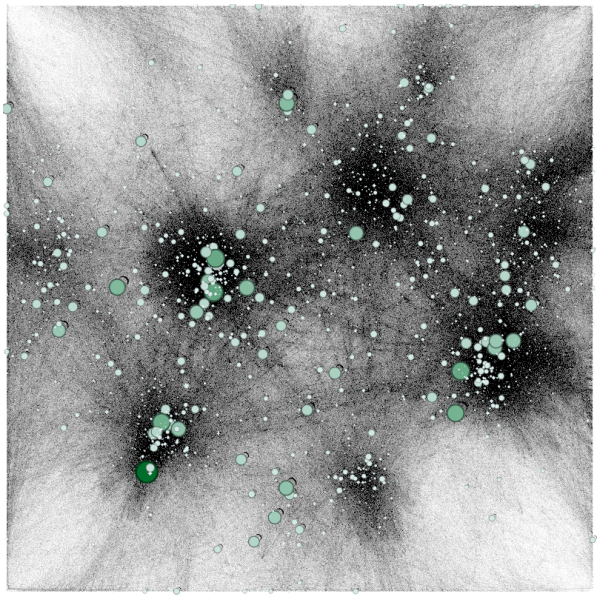

Figure 12

Figure 12 has 13k vertices and 563k edges. Vertices represent officers and color and size of each vertex represent the number of trrs the officer involved in. Edges are drawn between officers who were appointed to the same beats, in other words, officers who share the same value in the officer_assigned_beat column of table trr_trr.

It is obvious that there are several “black areas” in the graph, which means officers in these areas are more often assigned to the same beats in trr events than other officers. It is also apparent that most large vertices, which represent officers with more trrs, are located in these “black areas”.

According to the conclusion drawn from Figure 11, officers who caused more trrs form “co-offender” clusters. From the second graph, we can infer that these officers are probably assigned to the same beat at the beginning. Therefore, the deployment of officers is related to the “co-offender” relationship.

There are two additional conclusions for Figure 12. First, if we assume that officers are appointed evenly (which means the reason for “black areas" is not that aggressive officers were appointed to the same locations), then we can conclude that the use of force may “transmit” between officers, which means one aggressive officers may cause more officers to have a tendency of using force. Second, we have to consider the possibility that police officers assigned to beats with worse public order may have to use force more often than other officers, so the “black areas” may not all result from the officers’ willingness.

Checkpoint 5 - Natural Language Processing:

What is the average sentiment of the subject of allegations against police officers toward specific races?

```
AVERAGE NATIVE AMERICAN/ALASKAN NATIVE SENTIMENT FROM 31 ALLEGATIONS: -4.45
AVERAGE BLACK SENTIMENT FROM 8969 ALLEGATIONS: -4.23
AVERAGE HISPANIC SENTIMENT FROM 1471 ALLEGATIONS: -3.98
AVERAGE ASIAN/PACIFIC ISLANDER SENTIMENT FROM 188 ALLEGATIONS: -3.84
AVERAGE WHITE SENTIMENT FROM 3132 ALLEGATIONS: -3.63
```

We had a lot of trouble organizing the OCR from the TRR Narrative data so we decided to look into the sentiment of the allegation data to see if there was a disparity between different races. The sentiment of all of them is quite negative, which is to be expected when looking at allegation data, but there is a notable disparity between the sentiment of black individuals and the others. However, one can see that native american/alaskan native sentiment is even more negative than that of black, but this may just be due to low sample size.

Conclusion:

Throughout our project, we have found that there are many more TRRs against black people than any other race. They are consistently the subjects of police violence more than any other, and officers are much more likely to patrol neighborhoods that are majority black, and much less likely to patrol neighborhoods that are majority white, which may be a cause of many of the trends we have found throughout our project. It could also be the case that black communities are often more socio-economically disadvantaged due to societal factors, and this leads to increased crime which leads to increased policing, but these factors are beyond the scope of our project.

We have also found that police officers often commit TRRs in clusters based on their beat, and that there are no circumstances where we find that police officers are committing TRRs and violence on their own. Furthermore these beats are typically in neighborhoods that are majority black. Furthermore these officers often have some greater biases against black people as compared to those of other races as we found during Checkpoint 5. Due to all of these factors we've found that black people are the most often to be the subject of police violence.