Luke Salamone
Simon Benigeri
Renpin Luo

Checkpoint 2: Visualization

Our group produced visualizations for the following 2 points:

1. Comparison of police officer doppelgangers to understand the relationship between allegations and awards.
2. Do police officers frequently accused of misconduct get rewarded more than their less offending peers? Do factors like whether an allegation is sustained or use of force affect this result?

For each question, we wrote SQL scripts to produce tables, exported the data as csv, and inputted it to D3 to make visualizations.

In the README.md and in this report, we link to the D3 Visualization notebooks, where you can see the visualization, the input data, and the code.

1. Comparison of police officer doppelgangers to understand the relationship between allegations and awards.

Visualization: Calendar View
Script: timelines.sql
Data: timelines.csv

We wanted to study the awards and complaints records for officers that we can consider to be doppelgangers. We chose Jason Van Dyke as our starting point. He is a white male, born in 1978, joined the force in 2001, and after his initial assignment to unit 44, the Recruit Training Station, he went to unit 9, also known as District 9, to serve as a police officer.

We chose Jason Van Dyke because he is well-known for murdering Laquan McDonald by shooting him 16 times. Also, he joined unit 9, which our research showed to be one of the two units with the highest complaint records between 2007 and 2015, the years for which the cpdb dataset seems to be the most complete.

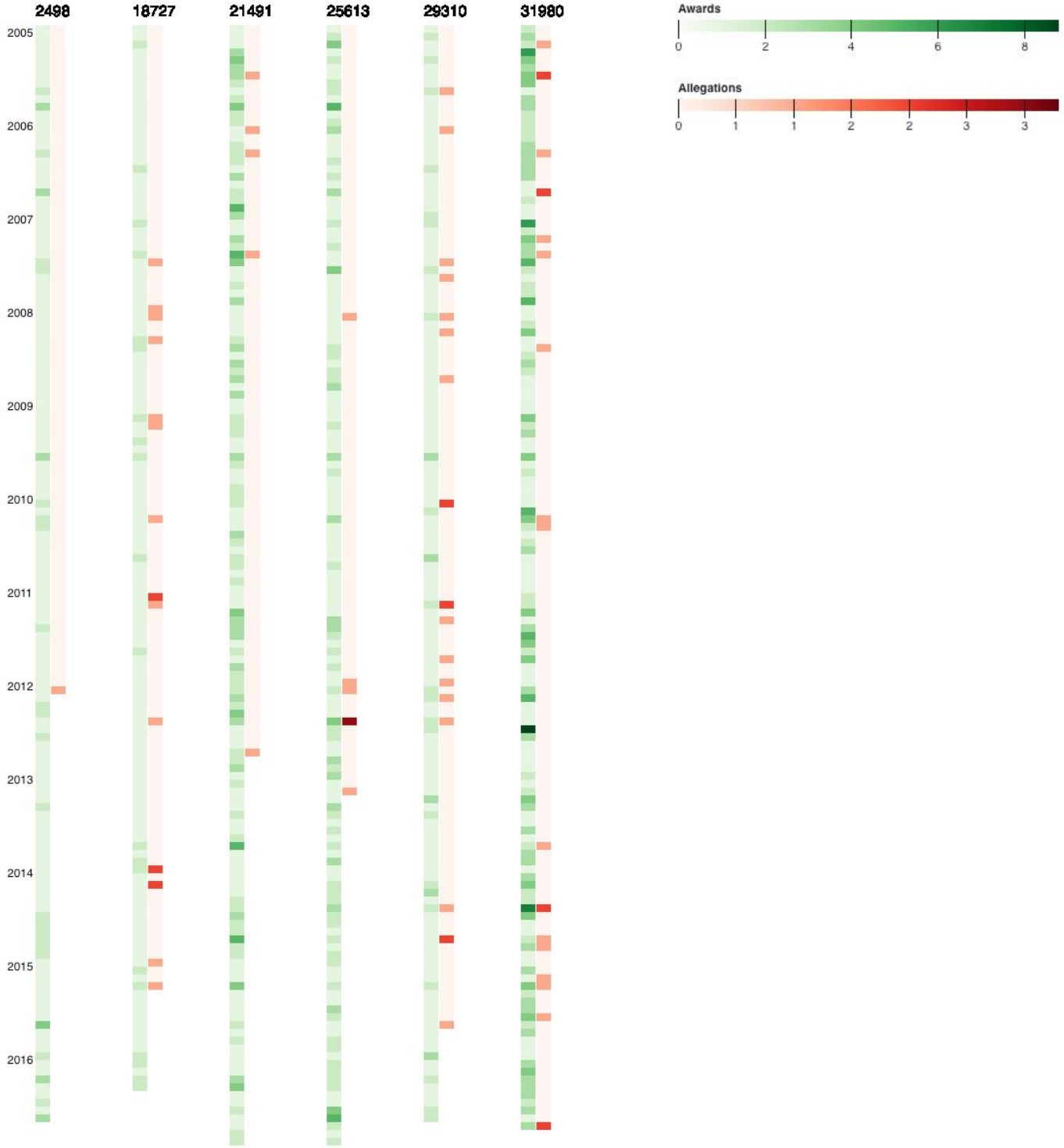
We interpreted doppelganger to mean the following: two officers are doppelgangers if they share similar demographic and career characteristics. We looked for white, male, police officers born between 1972 and 1978, who joined the force in 2001 and unit 9 in 2002. We then mapped their complaints and awards on a calendar visualization to produce a timeline of their careers from 2007 to 2015.

The doppelganger chart shows that even within a unit there is quite a bit of diversity even within a unit. Officers 31980 and 29310 seem to have received quite a few complaints, while officer 25613 received relatively few complaints. Officer 31980 seems to receive many complaints but also many awards. Officer 18727 is rather interesting. He seems to get little attention for his first couple years in terms of awards and complaints, then in mid-2007 begins to receive his first few complaints. This more aggressive pattern continues and appears to even accelerate towards the end of his timeline. Officer 18727 is Jason Van Dyke.

It is difficult to assess the differences between Van Dyke and Officer 31980, who seems to be even more prolific in terms of complaints. But 31980 also receives quite a few awards, and appears to be awarded at a relatively constant rate independent of the number of complaints he receives. He went from June 2010 to October 2013 without receiving a single complaint, but continued to receive awards during that time. In contrast, Van Dyke (18727) was never heavily awarded at any point in his career, receiving only sporadic positive attention from his superiors. Van Dyke's tenure with the police abruptly ended after a troubling increase in complaints.

We did not incorporate timelines for sustained complaints and use of force into the visualization. We can add these dimensions at a later date. Future research could include analysis of these factors.

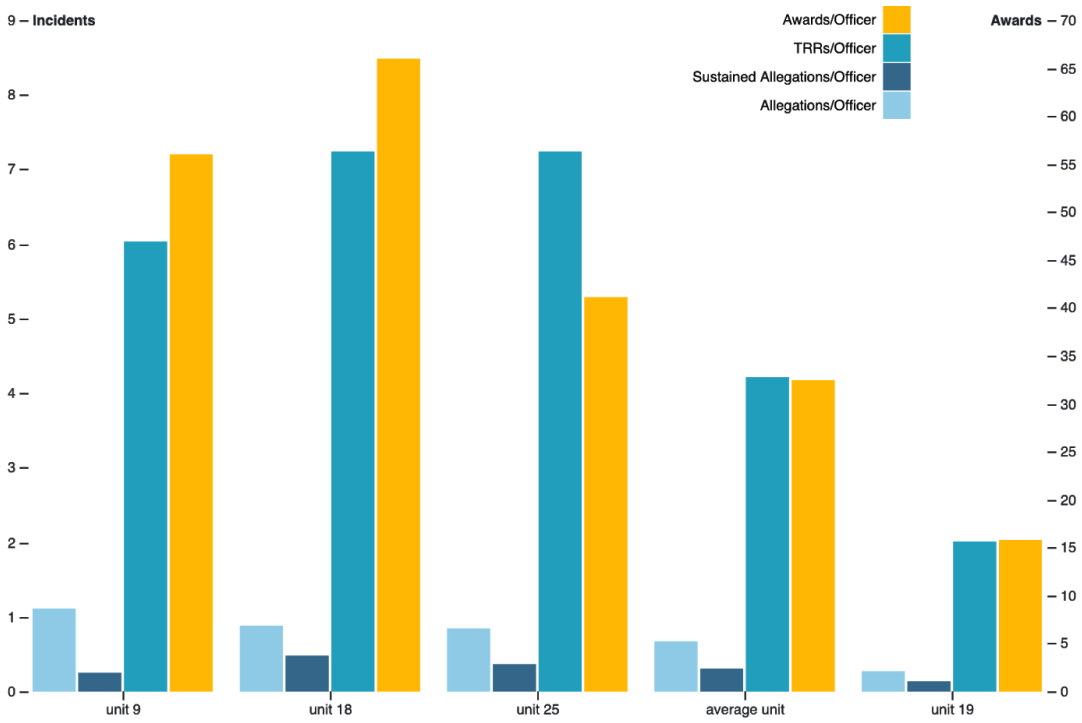
Timelines of Awards and Allegations for Police Officers



2. Do police officers frequently accused of misconduct get rewarded more than their less offending peers? Do factors like whether an allegation is sustained or use of force affect this result?

Visualization: Officer Allegations and Awards
Script: aggregation_2007_to_2015.sql
Data: aggregation.csv

Aggregate Police Officer Unit Misconduct, Use of Force, and Awards per capita from 2007 to 2015



We aggregated data from police units, complaints, allegations, officer history, trrs, and awards to map incidents of misconduct, use of force, or awards to police officers in units 1-25 over the period 2007 to 2015. Recall that we chose this period because it is the most complete for these datasets.

For each unit, we measured the TRR count per capita, allegation count per capita, sustained allegation count per capita, and awards count per capita, from 2007 to 2015. These figures are aggregates. That is, for a given unit, for each type of event, we count the occurrences from 2007 to 2015. We then divide this aggregate count by the aggregate of the number of officers in that unit over the years 2007 to 2015.

We displayed our results in a grouped bar chart, selecting units 9, 18, and 25 for their high figures in terms of misconduct and use of force occurrences per capita. We also displayed unit 19's numbers because it had the lowest figures in terms of misconduct and use of force, as well as the average counts per capita for all the units.

These results show that higher counts of misconduct and use of force do not preclude units from receiving awards. Quite the opposite, the two seem to go hand in hand. Surprisingly, this result still holds for sustained allegations. Units with higher counts of sustained allegations per capita tend to get more awards per capita.

It seems to us that awards are very common. Our current analysis ignores the significance of each award. In the future we think we should research types of awards in more detail, for example, honorable mentions.

Future research

In subsequent studies, we hope to expand the existing database to obtain a more representative and generalizable population, especially when we are focusing on analysing the career period of doppelgangers. The current database can only support our research and analysis of data from the Chicago Police Department for the period 2005 to 2018. The results we obtained provide a picture of how the Chicago Police Department has changed in recent years. Therefore, they are limited to that region.

In order to have a more authoritative analysis that is generally representative of the current state of police departments in the US, the scope of the database must be expanded. For example, we could include NY and New Jersey data. With a more comprehensive database we can conduct a more in-depth analysis. We can use datasets on multiple police departments to identify patterns that apply to police departments more generally. With datasets from police departments from different regions, we may have the means to identify underlying social tensions over police abuse. Since political, demographic, and economic context vary across different regions we can use their differences to control certain variables in.