Project Proposal

The Hungry Pitbulls
Vittorio Iocco
October 20, 2021

Theme

The theme of our project is to investigate the discrepancies in filing and disciplining in the CPDP. Specifically, we are interested in why these discrepancies exist as well as any correlations between these discrepancies, an officer's future adverse encounters and biased policing.

The discrepancies we will be looking into are:

1. The recommended findings/outcomes and the final findings/outcomes do not match.
2. Explicit or implicit missing final findings/outcomes or recommended findings/outcomes.
    a. Explicit missing data means the data is physically missing in the database. Implicit missing data means the data is present but ambiguous. An example of this is the 'Not Sustained' category which implies that there was not sufficient evidence to sustain. This implicitly creates ambiguity in the data because every tuple equivalently labeled with a 'Not Sustained' finding might be labeled as so for different reasons. These reasons are the missing information.
3. An allegation for use of force/excessive force is filed by a civilian but there is no Tactical Response Report (TRR) filed by the officer.

To this end, we posed and SQL-queried the following set of questions:

<u>Data Analytics</u>

Question 1: What are the frequencies of recommended and final outcomes and findings?

1. What are the unique categories for a final finding? When grouped by final finding, what are the officer allegation counts? When grouped by final finding, what is the percent of officer allegations in each category?
2. What are the unique categories for a final outcome? When grouped by final outcome, what are the officer allegation counts? When grouped by final outcome, what is the percent of officer allegations in each category?
3. What are the unique categories for a recommended finding? When grouped by recommended finding, what are the officer allegation counts? When grouped by recommended finding, what is the percent of officer allegations in each category?
4. What are the unique categories for a recommended outcome? When grouped by recommended outcome, what are the officer allegation counts? When grouped by recommended outcome, what is the percent of officer allegations in each category?
5. When grouped by final finding, what is the percentage of each final outcome as a percent of the total outcomes for that final finding? What are the counts?

Question 2: What are the frequencies of mismatches in TRR and complaint report filings?

1. What are the categories for which an officer filed a TRR? What are the counts and percentages (as a percentage of total officers allegations with a matching TRR) for these TRR categories?
2. How many instances are there when an allegation for force/excessive force is filed by a civilian and there is a Tactical Response Report (TRR) filed by the officer? What is this as a percentage of allegations of force/excessive force?

Question 3: Are certain groups of categories more likely to result in a less/more severe final findings or outcome than recommended for officers?

1. When grouped by allegation category, what percent of complaints in each allegation category have a less/more severe recommended finding than final finding (as a percent of? What percent of complaints in each allegation category have recommended findings or final findings that are null/unknown?
2. When grouped by allegation category, what percent of complaints in each allegation category have a less/more severe recommended outcome than final outcome? What percent of complaints in each allegation category have recommended outcomes or final outcomes are null/unknown?

Question 4: Is there a correlation between the race of the victim and the rate at which the recommended finding and outcome are finalized?

1. When grouped by the race of the victim, what is the distribution of recommended findings as counts and percentages of the race's total allegations?
2. When grouped by the race of the victim, what is the distribution of final findings as counts and percentages of the race's total allegations?
3. When grouped by the race of the victim, what is the distribution of recommended outcomes as counts and percentages of the race's total allegations?
4. When grouped by the race of the victim, what is the distribution of final outcomes as counts and percentages of the race's total allegations?
5. When grouped by race of victim, what percent of complaints in each race have a less/more severe recommended finding than final finding (as a percent of? What percent of complaints in each allegation category have recommended findings or final findings that are null/unknown?
6. When grouped by race of victim, what percent of complaints in each race have a less/more severe recommended outcome than final outcome? What percent of complaints in each allegation category have recommended outcomes or final outcomes are null/unknown?

<u>Data Exploration & Interactive Visualization</u>

1. Do discrepancies between the final findings and recommended findings reveal any racial bias? Are there any trends between severity, race, and category of the allegation?
    a. A set of pie charts (one for each race) detailing recommended findings
    b. A set of pie charts (one for each race) detailing final findings
    c. A set of heat maps (one for each race) of the relationship between recommended findings and the category of a complaint.
    d. A set of heat maps (one for each race) of the relationship between final findings and the category of a complaint.

    e. A parallel plot of the ratio of final findings that are more severe than recommended findings to final findings that are less severe than recommended findings, plotted over the categories of allegations. There will be one line for each race.
2. Do discrepancies between the final outcomes and recommended outcomes reveal any racial bias? Are there any trends between severity, race, and category of the allegation?
    a. A set of pie charts (one for each race) detailing recommended outcomes
    b. A set of pie charts (one for each race) detailing final outcomes
    c. A set of heat maps (one for each race) of the relationship between recommended outcomes and the category of a complaint.
    d. A set of heat maps (one for each race) of the relationship between final outcomes and the category of a complaint.
    e. A parallel plot of the ratio of final outcomes that are more severe than recommended outcomes to final outcomes that are less severe than recommended outcomes, plotted over the categories of allegations. There will be one line for each race.
3. Do discrepancies between the final findings and recommended findings reveal any community bias? Are there any trends between severity, community, and category of the allegation?
    a. A choropleth of the city of Chicago broken down into beats with each beat being colored according to the percentage of officer allegations with equal recommended and final finding
    b. A choropleth of the city of Chicago broken down into beats with each beat being colored according to its ratio of more severe final findings than recommended findings to less severe final findings than recommended findings.
    c. A choropleth of the city of Chicago broken down into beats with each beat being colored according to its ratio of more severe final findings than recommended findings to less severe final findings than recommended findings.
4. Do discrepancies between the final outcomes and recommended outcomes reveal any community bias? Are there any trends between severity, community, beat and category of the allegation?
    a. A choropleth of the city of Chicago broken down into beats with each beat being colored according to the percentage of officer allegations with equal recommended and final finding
    b. A choropleth of the city of Chicago broken down into beats with each beat being colored according to its ratio of more severe final findings than recommended findings to less severe final findings than recommended findings.
    c. A choropleth of the city of Chicago broken down into beats with each beat being colored according to its ratio of more severe final findings than recommended findings to less severe final findings than recommended findings.
5. Does receiving less severe punishments embolden officers to commit future offences? Does receiving more severe punishments disembolden officers from committing future offences?
    a. A histogram plotting the number of allegations in an officer's last 20% of their time on the force against the percent of allegations in the first 80% of their career that received less severe outcomes or findings than recommended.
    b. A histogram plotting the number of allegations in an officer's last 20% of their time on the force against the percent of allegations in the first 80% of their career that received more severe outcomes or findings than recommended.

6. Is there a correlation between how well an officer follows TRR policy and their likelihood to be a repeat offender?
    a. Heat map of the percent of use of force allegations against an officer that have a TRR and the total number of allegations against that officer?
7. Is there a correlation between the race of the victim and whether or not officers decide to write TRRs?
    a. A pie chart showing the racial breakdown of use of force allegations.
    b. A pie chart showing the racial breakdown of use of force allegations that receive a TRR.
8. Is there a correlation between the beat where an incident occurs and whether or not officers decide to write TRRs?
    a. A choropleth of Chicago divided into beats with a coloring scheme that indicates the percent of use of force allegations that received a TRR.

<u>Graph Analytics</u>

1. Is there a correlation between how well an officer follows TRR policy and how likely their partners are to follow TRR policy?
    a. Create a graph where each node is an officer that has broken TRR policy and links are branches between officers that have been in the same unit. The size of the node is scaled according to how many times TRR policy is broken.
2. Is there a correlation between how often officers get less severe punishments than recommended and how often their partners get less severe punishments than recommended?
    a. Create a graph where each node is an officer that has received a less severe punishment than recommended and branches are links between officers that have been in the same unit. The size of the node is scaled according to the number of less severe punishments an officer has received.
3. Is there a correlation between how often officers get less severe findings than recommended and how often their partners get less severe findings than recommended?
    a. Create a graph where each node is an officer that has received a less severe finding than recommended and branches are links between officers that have been in the same unit. The size of the node is scaled according to the number of less severe findings an officer has received.
4. Is there a correlation between the number of missing recommended findings/outcomes an officer receives and the number their partners receive?
    a. Create a graph where each node is an officer that has received a null recommended finding/outcome and branches are links between officers that have been in the same unit. The size of the node is scaled according to the number of null findings/outcomes an officer has received.

<u>Natural Language Processing</u>

1. Use association analysis on the summaries of an allegation to determine if there is some explanation for the large number of recommended finding/outcomes that are null?
2. Use association analysis on the summaries of an allegation to determine if there are certain types of inappropriate actions an officer can commit that are harder to “prove” and thus become ‘Not Sustained’?