

Checkpoint 3: Findings

Yuxin Chen, Song Luo, Jing Jiang

Introduction

For this checkpoint, we will be using D3.js to show some interactive visualizations. Due to our updates in the past checkpoints, we also made some changes on the questions we are trying to solve.

Questions and Findings

Q1: Use Circle Packing to group repeaters in each timespan with age groups(e.g. Age 30-40, Age 40-50) and use color encodings to show the composition of repeaters' races. Detect any trend of changes in terms of age and race of the repeater.

[The full images can be displayed properly with Observable Notebooks, URL under figures]

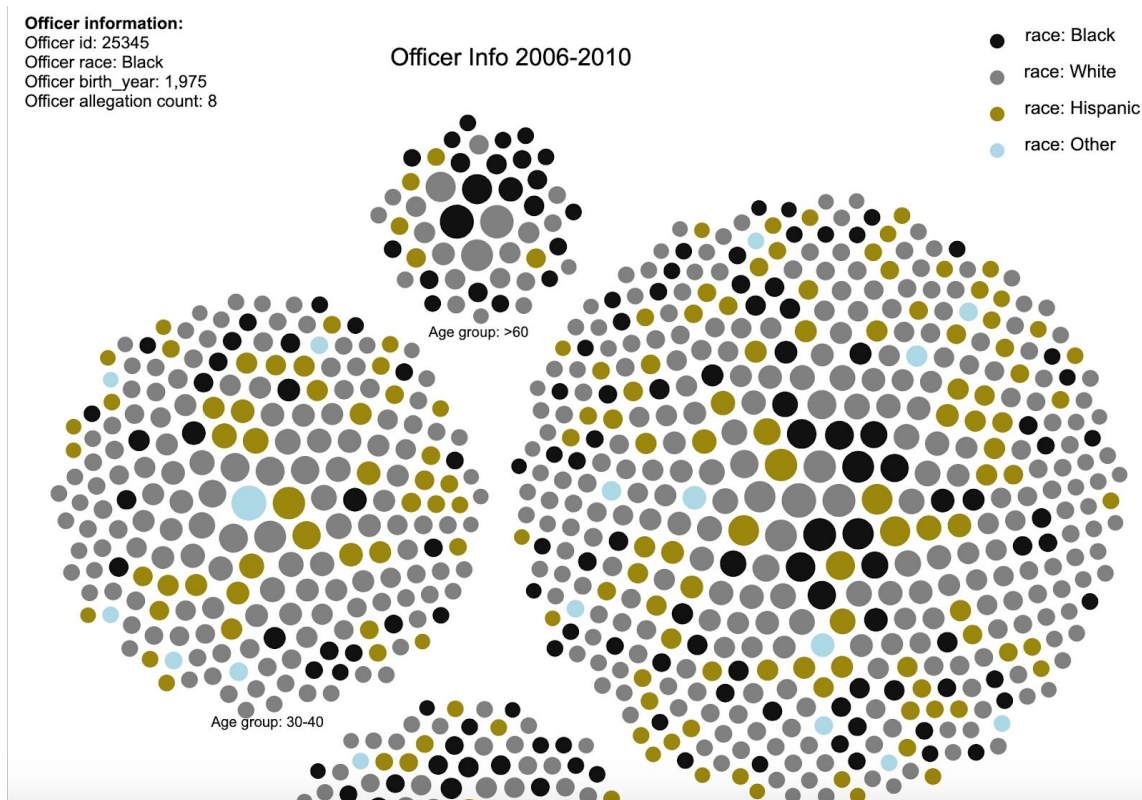


Figure 1: Officer Info 2006-2010

URL: <https://observablehq.com/d/8e6fa26bf1c2167f>

Officer information:
Officer id: 30818
Officer race: Black
Officer birth_year: 1,979
Officer allegation count: 8

Officer Info 2010-2014

- race: Black
- race: White
- race: Hispanic
- race: Other

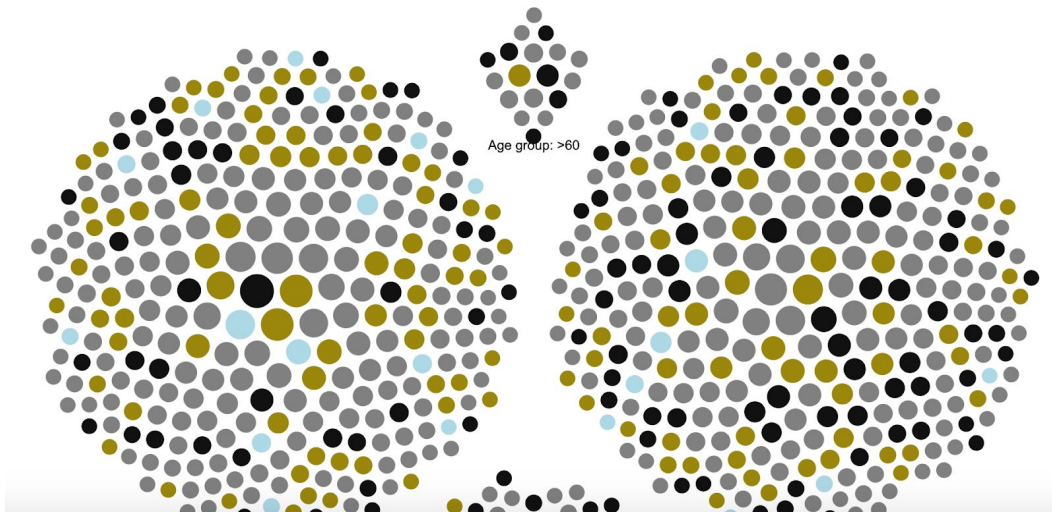


Figure 2: Officer Info 2010-2014

URL: <https://observablehq.com/d/6c2172cb796fb99a>

Officer information:
Officer id: 906
Officer race: Asian/Pacific
Officer birth_year: 1,984
Officer allegation count: 3

Officer Info 2014-2018

- race: Black
- race: White
- race: Hispanic
- race: Other

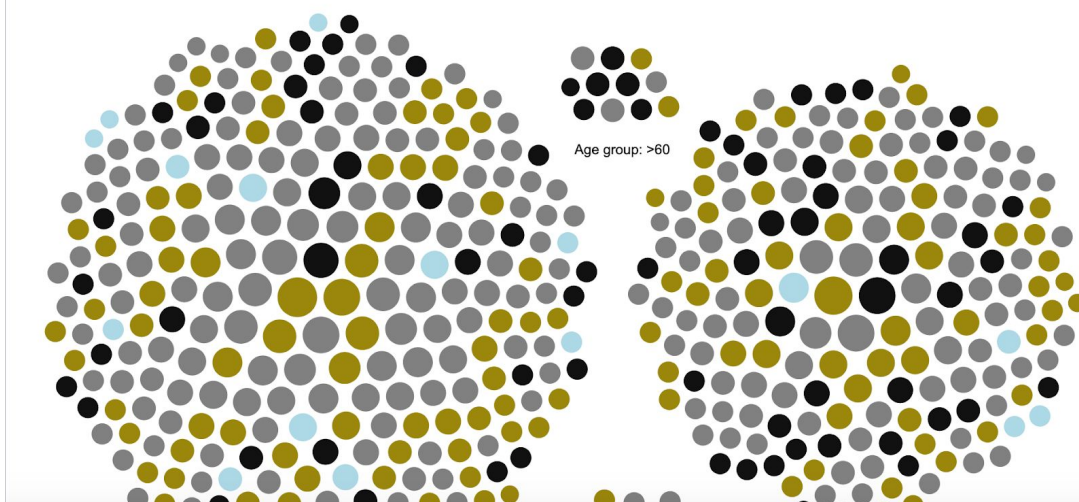


Figure 3: Officer Info 2014-2018

URL: <https://observablehq.com/d/1a35f874d4dc054d>

From the circle packings we have the following key findings:

1. the quick shrinking of two age groups that consist of officers above 50-year-old.
2. The majority of the repeaters are either Black, White or Hispanic. Officers of minority contribute very little to repeaters and this is consistent with the US Social Structure
3. Hispanic is quickly emerging as a dominant part of repeaters, together with White repeaters.

This fits with our theme to see the trait and property of repeaters to help prevent future misconduct. We are not yet sure about if age is an important factor: our current assumption is that when officers get older, they are more likely to assume more senior positions and take more officer jobs, thereby reducing confrontations significantly. In the future we could join the data regarding officer ranks and see if the hypothesis is true. Also regarding the increase of Hispanic repeaters, we want to investigate if there is a certain group of people they are targeting that results in this increase.

Q2: Sequences sunburst is the other Interactive Visualization of us since it can show the categories. This is good for us since it can show things like categories of misconduct or the number of repeaters who receive complaints they are responsible for.



Figure 4: Sunburst visualization

URL: <https://observablehq.com/@jassiy0/sunburst>

From Figure 4 we can easily identify the following patterns for each timespans:

1. For Timespan 1 (2006 - 2010), the top 5 types of allegations for the repeaters are Illegal Search, Use of Force, Operation/Personal Violations, False Arrest, and Lockup Procedures.
2. For Timespan 2 (2010 - 2014), the top 5 types of allegations for the repeaters are the same as that from timespan 1: Illegal Search, Use of Force, Operation/Personal Violations, False Arrest, and Lockup Procedures.
3. For Timespan 3 (2014 - 2018), the top 5 types of allegations for the repeaters are Operation/Personal Violations, Illegal Search, Use of Force, False Alert, and Lockup Procedures. The Operation/Personal Violations become the most frequent category. However, the ordering of other categories remains the same.

If we then look at the overall traits of the visualization, we can find that:

1. Different colors denote data from different time spans. We can identify that the number of allegation cases from timespan 1 to timespan 3 is gradually decreasing.
2. The number of cases from timespan 1 contributed to over half the total cases.
3. The number of cases from timespan 3 is less than half of that from timespan 2.

To conclude, from the sunburst visualization, we can see that the number of cases follows a descending trend from timespan 1 to 3. But the top categories of misconduct of repeaters are stable. They are Illegal Search, Use of Force, Operation/Personal Violations, False Alert, Lockup Procedures. The ordering of misconduct frequency is quite stable for all three timespans, but in Timespan 3, Operation/Personal Violations become the top category. It would be interesting to research why the number of cases of repeaters decreases over time. Also, we would like to investigate if the change of ordering of the misconduct types is related to the drop in the number of allegations across timespans.