

## Checkpoint 2: Data Visualization

### The Brave Ducks

#### Introduction

Before we start our analysis, the questions that we are investigating are in checkpoint1 are as following:

1. Using our definition of types of "income neighborhood", what is the total number of officer allegations for all **low, middle, high income** neighbors?
2. What is the rate of increase for officer complaints for **low, middle, high** neighborhoods between 2002-2007 and 2007-2012 timeframes, 2007-2012 and 2012-2017 timeframes?
3. What is the percentage of misconduct allegations (illegal search, use of force, etc) out of all allegations for these **low, middle, high** neighborhoods?
4. Among the officer allegations with complaints filed in the **low, middle, high** neighborhood, what percentage of the cases are dismissed?

In this checkpoint, we are interested in answering questions 3 and 4 using packed bubbles and treemaps. These questions we intend to investigate further are:

- What is the percentage of misconduct allegations (illegal search, use of force, etc) out of all allegations for these **low, middle, high** neighborhoods?
- Among the officer allegations with complaints filed in the **low, middle, high** neighborhood, what percentage of the cases are dismissed?

#### Analysis

What are the top misconduct allegations (illegal search, use of force, etc) out of all allegations for these **low, middle, high** neighborhoods?

What is the proportion of allegations (illegal search, use of force) out of all allegations for low income neighborhoods?

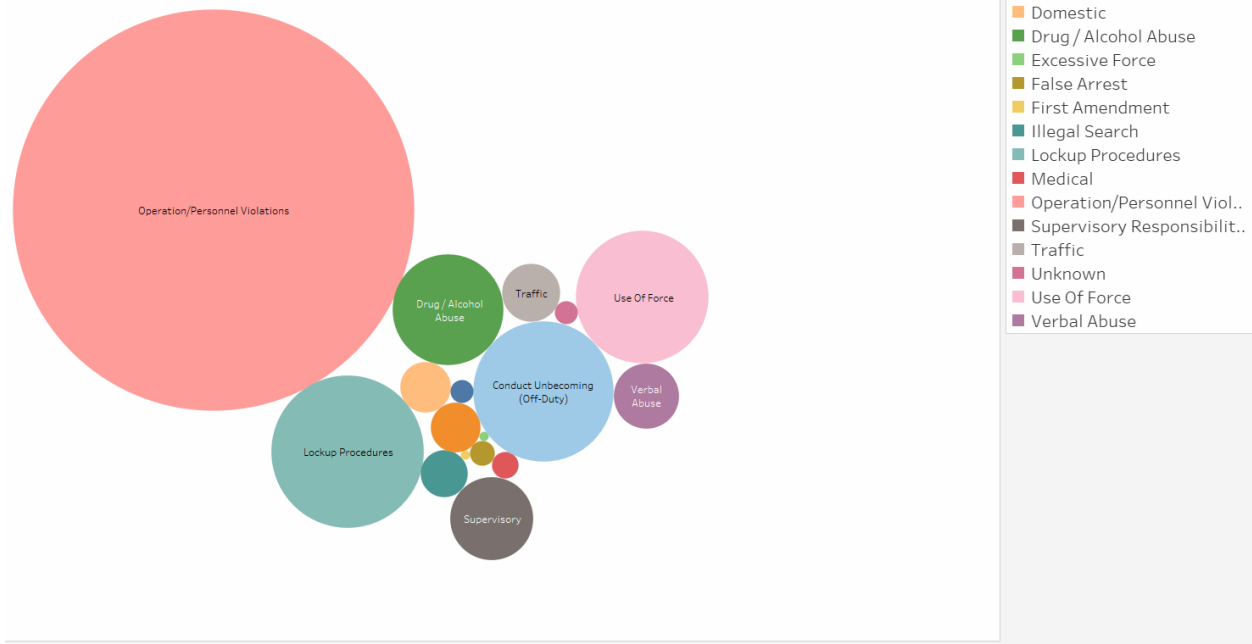


Figure 1: number of cases for each allegation category in low income neighborhoods

As we can see from the packed bubble chart, the major categories of misconduct allegation are operation/personnel violations, lockup procedures, conduct unbecoming, and use of force in low income neighborhoods.

What is the proportion of allegations (illegal search, use of force) out of all allegations for middle income neighborhoods?

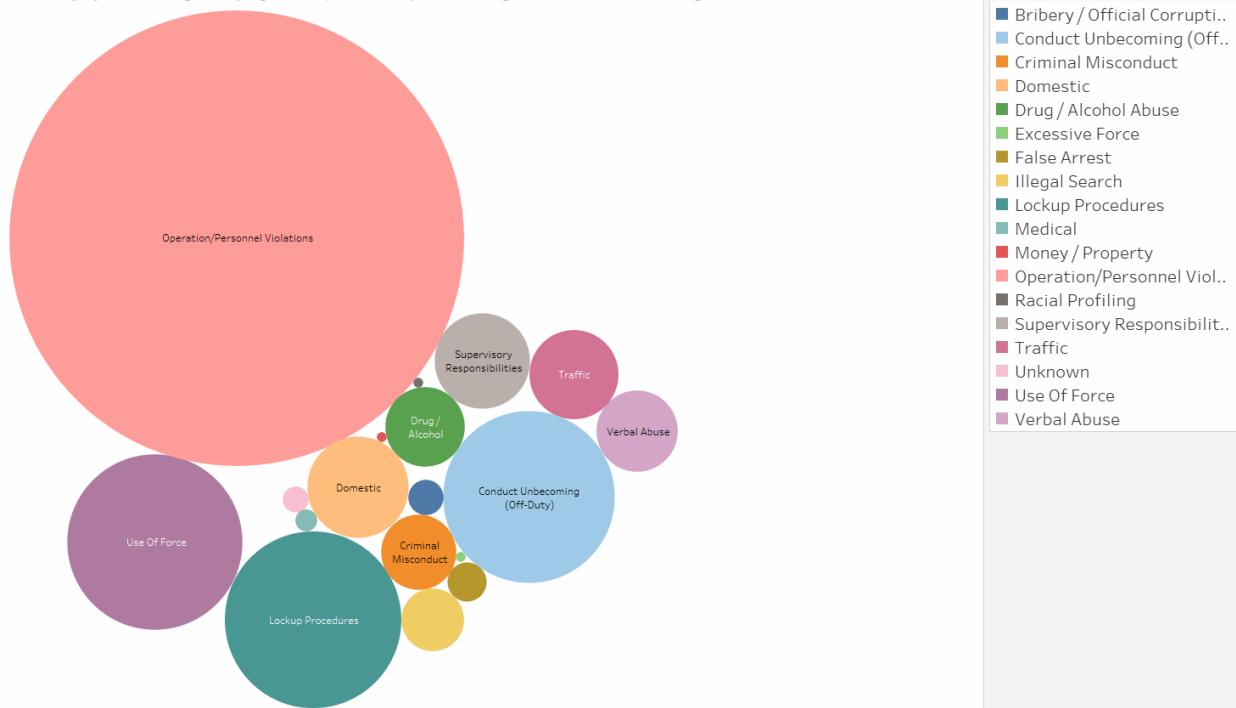


Figure 2: number of cases for each allegation category in middle income neighborhoods

As we can see from the packed bubble chart, the major categories of misconduct allegation are operation/personnel violations, lockup procedures, conduct unbecoming, and use of force in low middle neighborhoods as well. Additionally, other categories which were previous "eclipsed" in the low income neighborhood have increased.

What is the proportion of allegations (illegal search, use of force) out of all allegations for middle income neighborhoods?

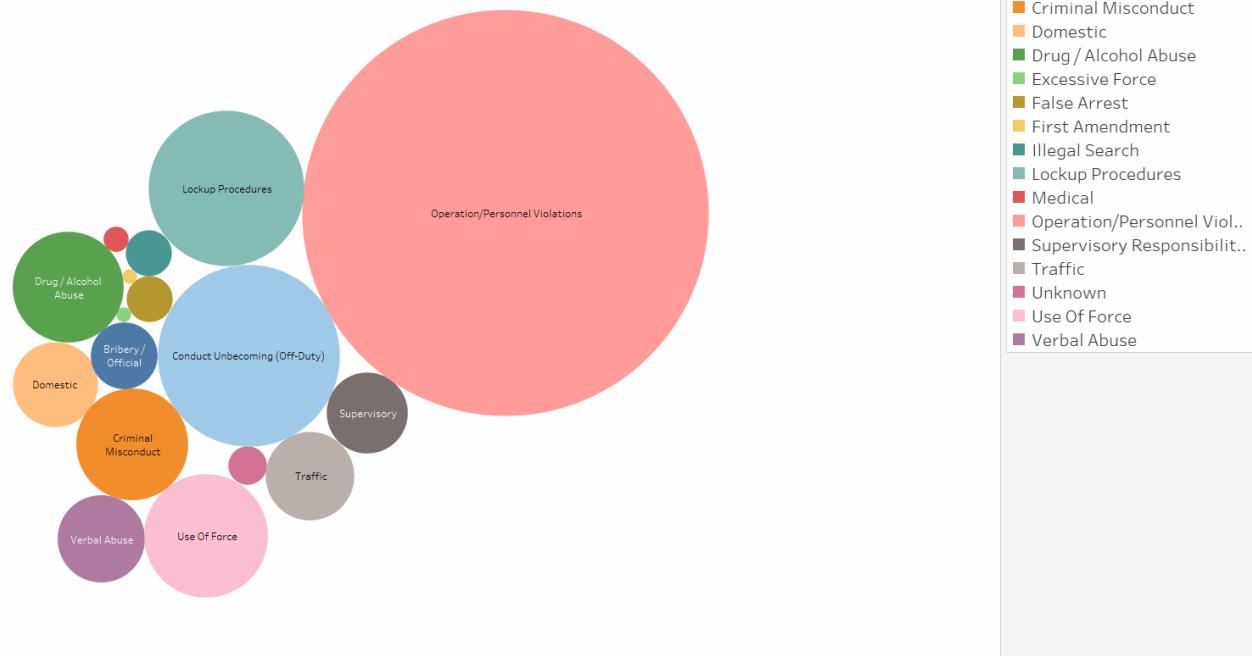


Figure 3: number of cases for each allegation category in high income neighborhoods

Similar to figure 5,6, we see the same trend as well for high income neighborhoods. And the gaps between the most significant category - operation/personnel violations and other allegations get narrowed.

Among the officer allegations with complaints filed in the **low, middle, high** neighborhood, what are the final outcomes?

Among the officer allegations with complaints filed by low income people, what is the allegation outcome?



Figure 4: histogram of case resolutions for low income neighborhoods

As we can see from the top graph, for low income neighborhoods, we have seen most of the outcomes for the cases are "no action taken", followed by "Reprimand" and "Unknown".

Among the officer allegations with complaints filed by middle income people, what is the allegation outcome?

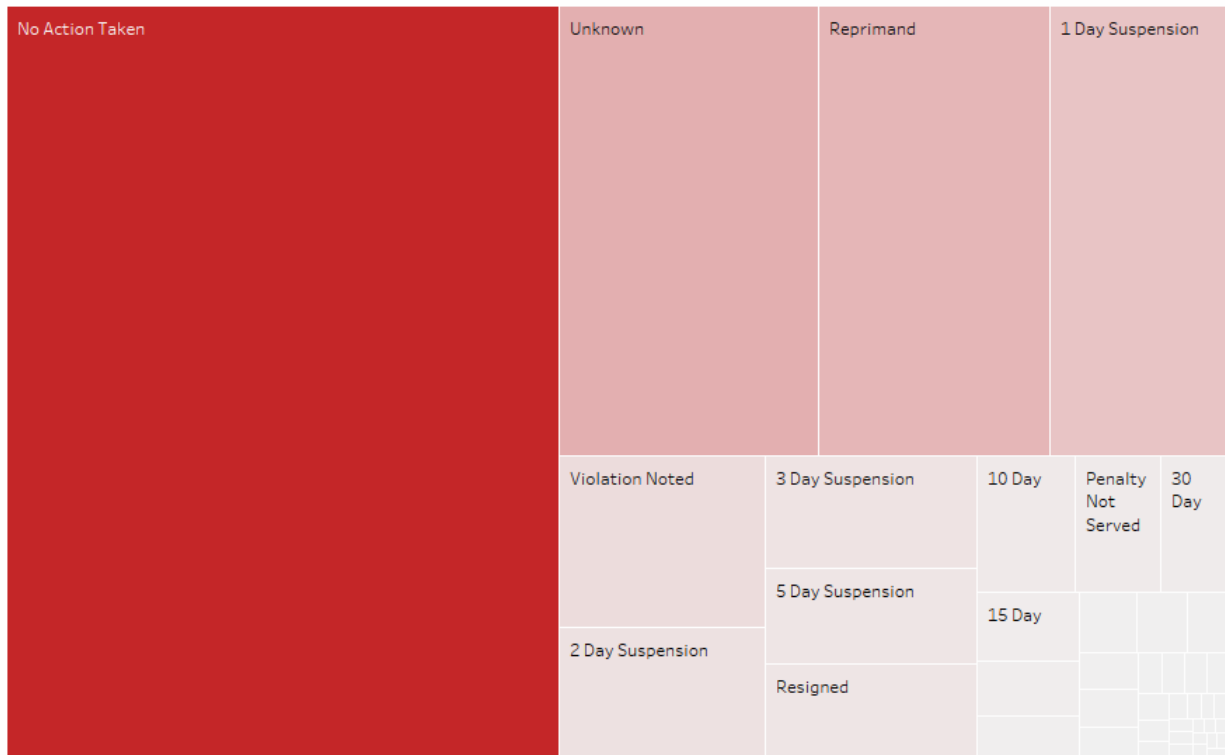


Figure 5: histogram of case resolutions for middle income neighborhoods  
 Similar to figure 9, for middle income neighborhoods, we have seen most of the outcomes for the cases are "no action taken", followed by "Reprimand" and "Unknown".

Among the officer allegations with complaints filed by high income people, what is the allegation outcome?

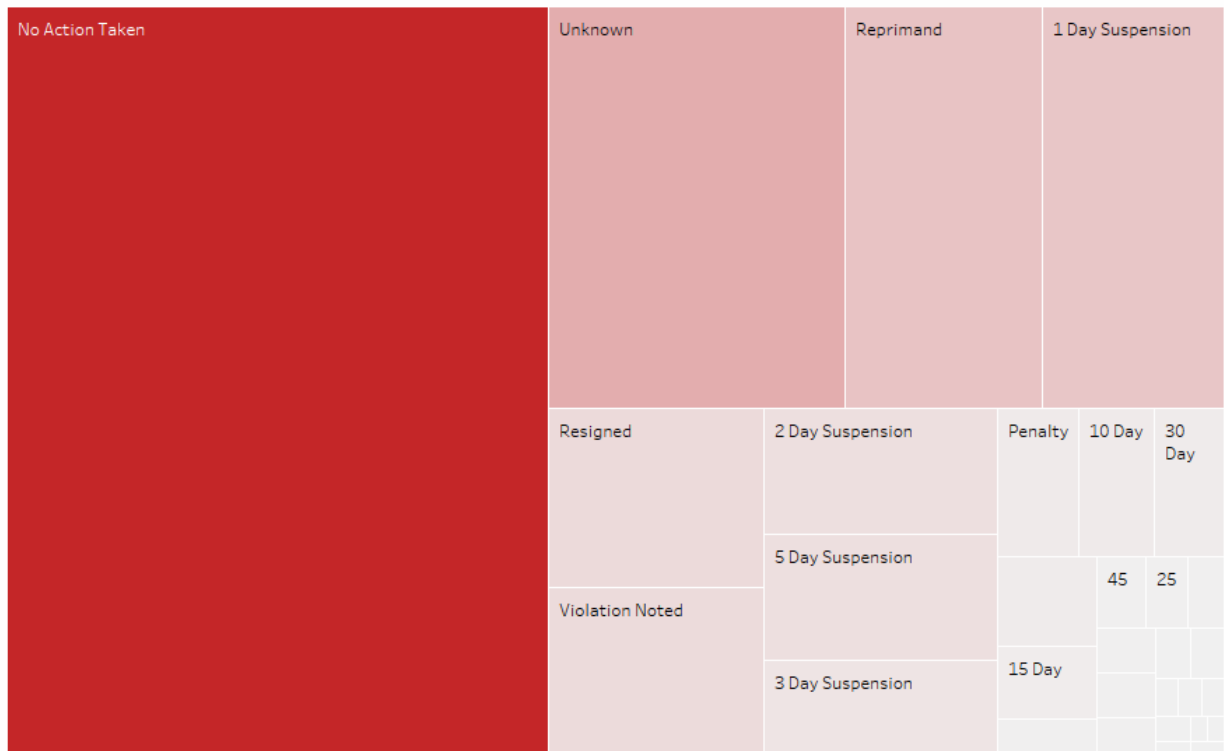


Figure 6: histogram of case resolutions for high income neighborhoods  
 Similar to both figure 9 and 10, for high income neighborhoods, we have seen most of the outcomes for the cases are "no action taken", followed by "Reprimand" and "Unknown".

### Experience with Tableau

- Query result is very slow.
- Tableau never saves the password.
- Poor documentation.
- Need to manually do the data cleaning - filter off the null value during visualization