The theme of the project is to analyse the trends before and after the cpdp went public. For example: to find changes in general behaviour of the officer, settlement amounts etc. and to predict for upcoming years. The questions are focussed towards finding out the things that have been tried to be hidden in the official record. For example, finding the relationship between the arrests and the allegation, because there are chances that if no arrest took place then the action taken by the officer might be illegal or looking for cases where things common between investigating officer and allegated officer impacted the case. So, for every question mentioned below, findings will be based on before and after cpdb release.

CP2- Visualization:

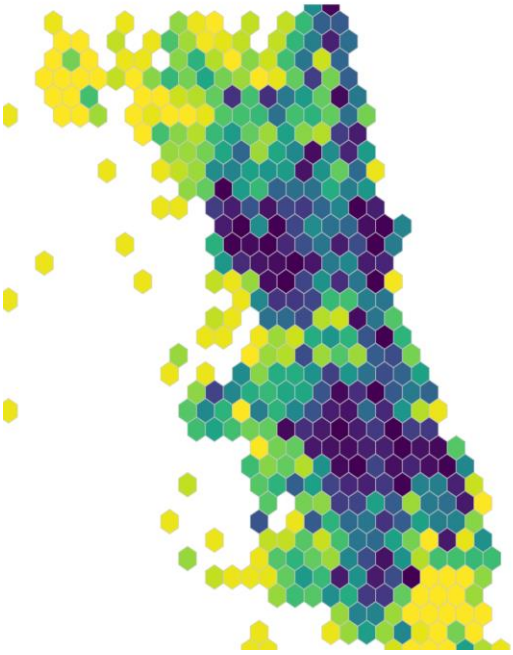
- **Heat map of areas in which the allegations take place while also considering the time of the day.**
  Using a slider bar for time and look at a heatmap of various complaint types over time for interactive visualization

  Observe the by hovering on the map you will see the number of allegations in that area.
  A slider has also been provided with the following labelling
  0: Allegation of years before 1950
  1: Allegation of years between 1950 - 2000
  2: Allegation of years between 2000 - 2003
  3:Allegation of years between 2003 - 2006
  4:Allegation of years between 2006 - 2009
  5:Allegation of years between 2009 - 2012
  6:Allegation of years between 2012 - 2015
  7:Allegation of years between 2015 - 2018
  8:Allegation of years between 2018 - present

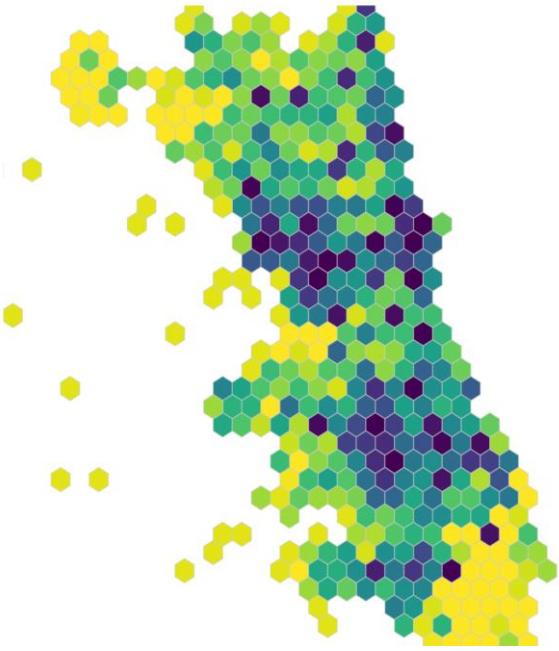
  https://observablehq.com/@rautnikita77/heatmap

The heat map shows distribution of the cases for given interval of years with a slider to vary it. There are some areas were the hexagons are very dark for most of the intervals of years signifying spots where these misconducts are very common. It would have been interesting to analyze officers that were assigned to these areas during these particular intervals of time.

Before 2015

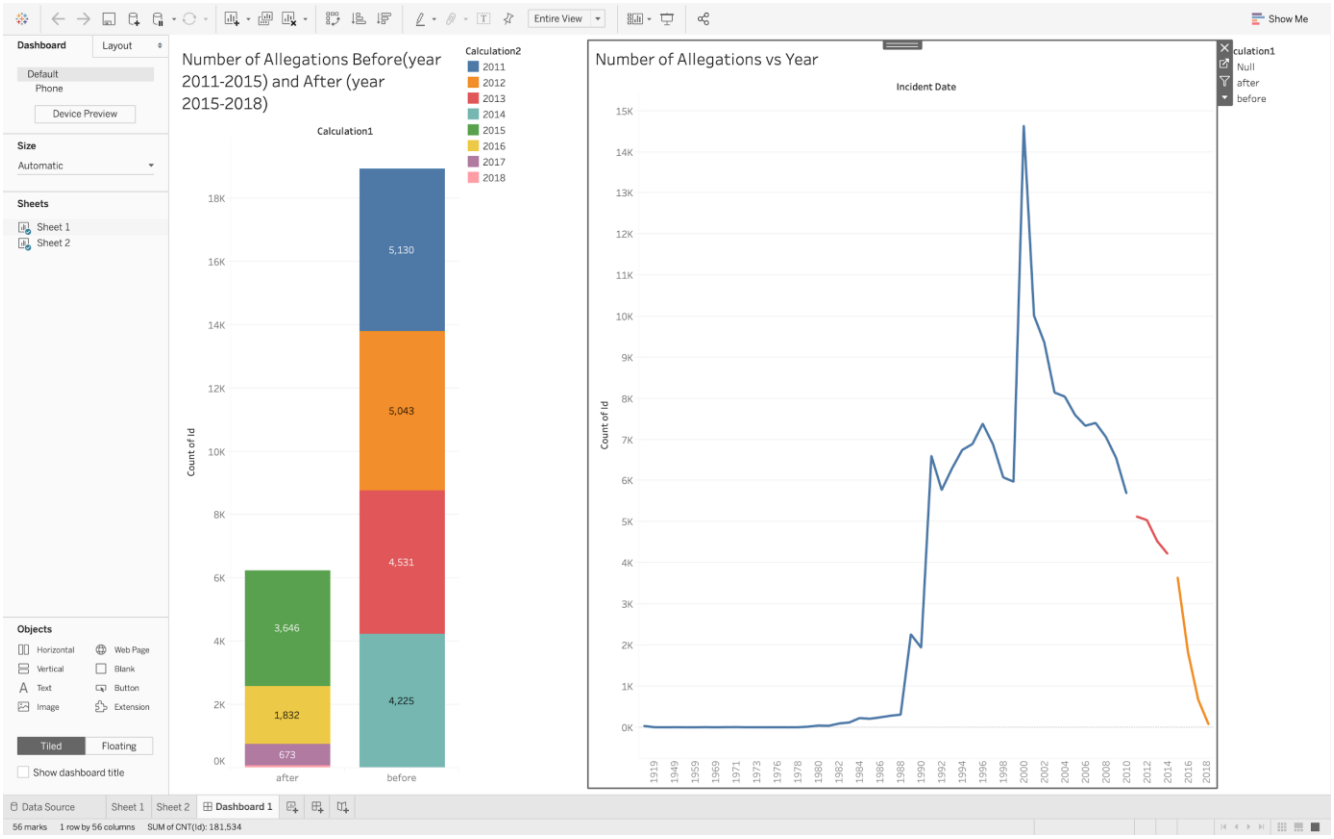


After 2015



We can see that highly are targeted areas remain same even after cpdb went public.

- **Visualizing whether there has been any decrease in allegations after the CPDP data has gone public with the help of line charts (Allegation vs Time) and histograms (Before vs After)**

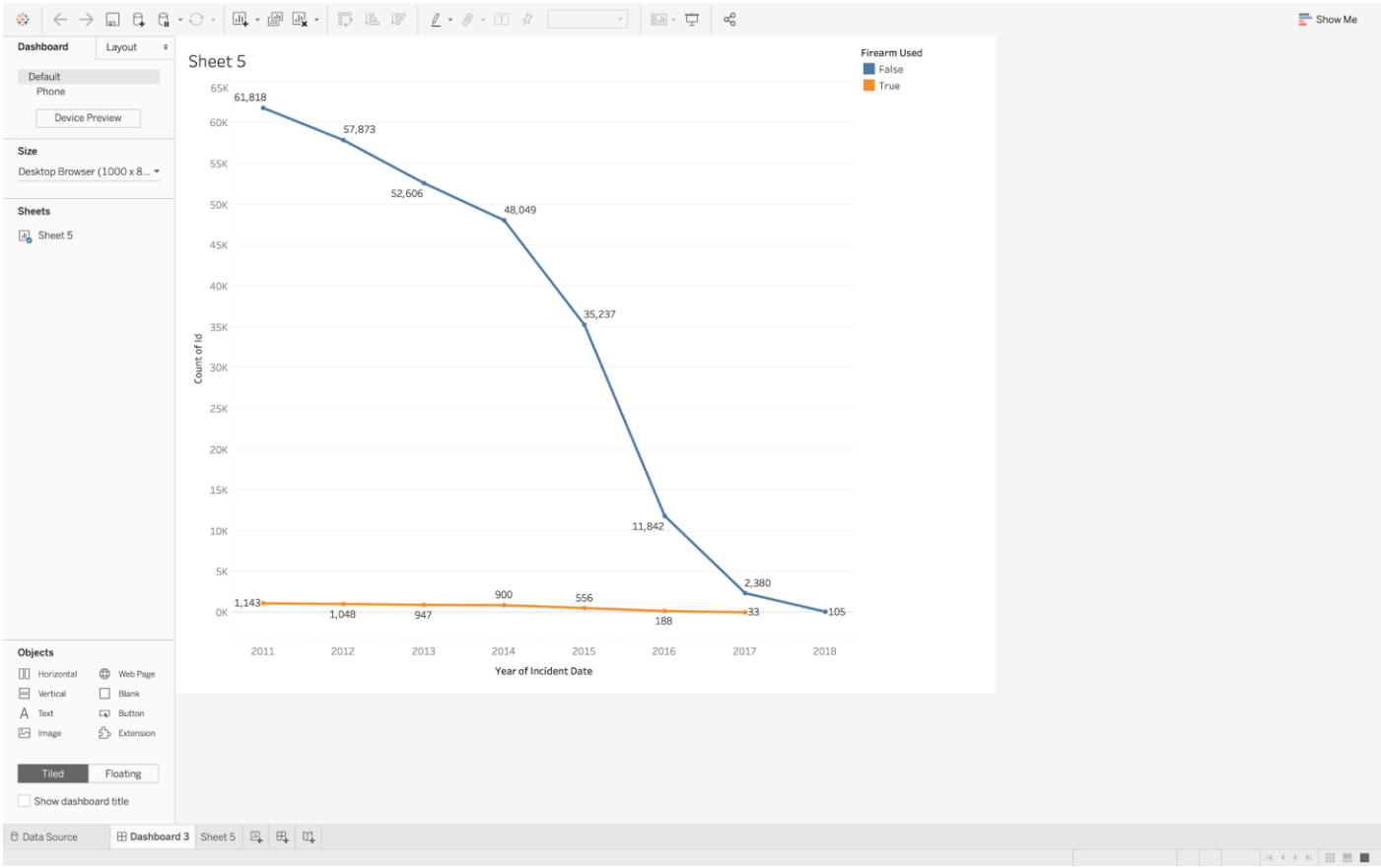


The graph on the right shows the distribution of the data year wise. It would be unfair to compare the of all the previous years to the 4 years since the data has gone public since the distribution is not even. Normalizing would not be a good idea since the years before 2000 are hardly contributing to the data. Therefore, in the graph on the left, we have compared 4 years before vs 4 years after the data went public. Before was significantly higher.

- **Visualizing the trends in the race of people affecting number of allegations.** With the data going public, the fact that people might be targeted based on race was highlighted. Visualizing using histograms for before and after or using line plot for the overall trend.

Similar to the last case, we have plots for 4 years before and after. For the graph on right side, there was significant decrease in the for all race, but especially for white and north American/Alaskan native after 2015. Graph on left side shows similar results.

- **Analysis of the number of incidents that used weapons vs the ones that did not.** With the data going public, did smaller incidents (assuming using gun makes the the incident much severe) reduce in number?



As expected, the number of cases not using firearm reduced significantly. One way to interpret this would be that since the dataset got public, most of the allegation cases now are the ones that absolutely had to use force and therefore possibly gun.