Students: JaCoya Thompson (ID 2992892), Lianke Qin (ID 3229035), Victor S. Bursztyn (ID 3046613). Professor Jennie Rogers.

Checkpoint #2 — Visualization

Context for this checkpoint: Our group attempts to investigate how socio-demographic features such as race and ethnicity could play a role in the *outcomes* of police misconduct investigations, rather than focusing on the *occurrences* of such misconducts. Our project investigates the question: **"When a case of police misconduct victimizes groups that are historically marginalized, how unfavorable is its investigation?"** We will look at different hypothetical effects such as: (1) "Is an allegation made by a victim of color less likely of being sustained?"; (2) "Do the durations of misconduct investigations vary across victim races?"; and (3) "Do the disciplinary actions differ?"; to test for the presence of significant racial disparities in the outcomes of investigations involving people of color. This checkpoint allows us to move beyond the baselines extracted in Checkpoint #1 and explore the distributions of those effects using victim race as a potentially discriminant variable.

For full context, please refer to file "**full_proposal.pdf**."

Question 1: What's the percentage of allegations sustained across victim races?

Tableau Chart #1:
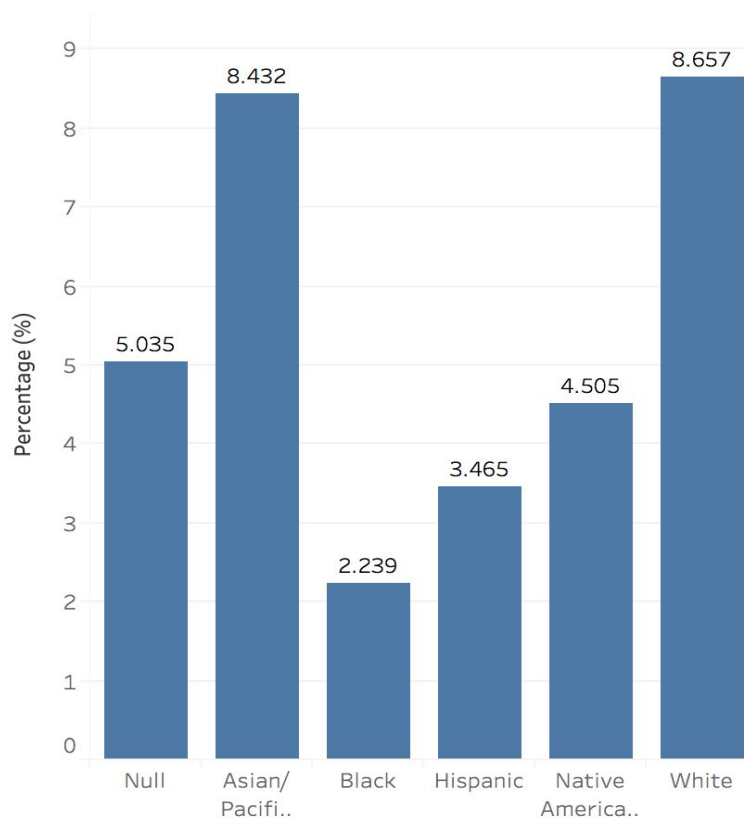
Results: The resulting bar graph displays the percentage of allegation sustained across victims races. This means the complaint resulted in an investigation and the allegation is supported by sufficient factual evidence of a violation of policy. **We observe Asian/Pacific Islander (8%) and White (9%) to have the highest percentages of allegations sustained with Black (2%) and Hispanic (3%) least likely to have a complaint sustained.** Regarding our overarching theme, **this is our first indication of the potential influence of race and ethnicity in the outcomes of investigations**. As discussed in our proposal, more evidence will be gathered and analyzed in future checkpoints, including the presence of lurking variables.

Question 2: What's the distribution of investigation times?
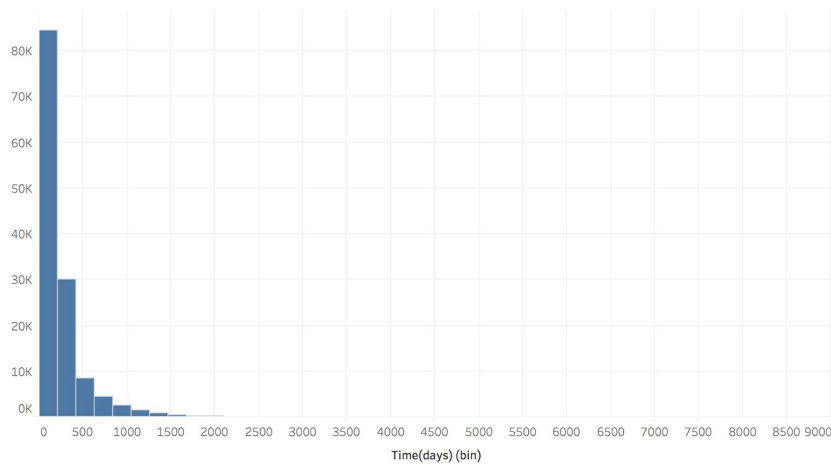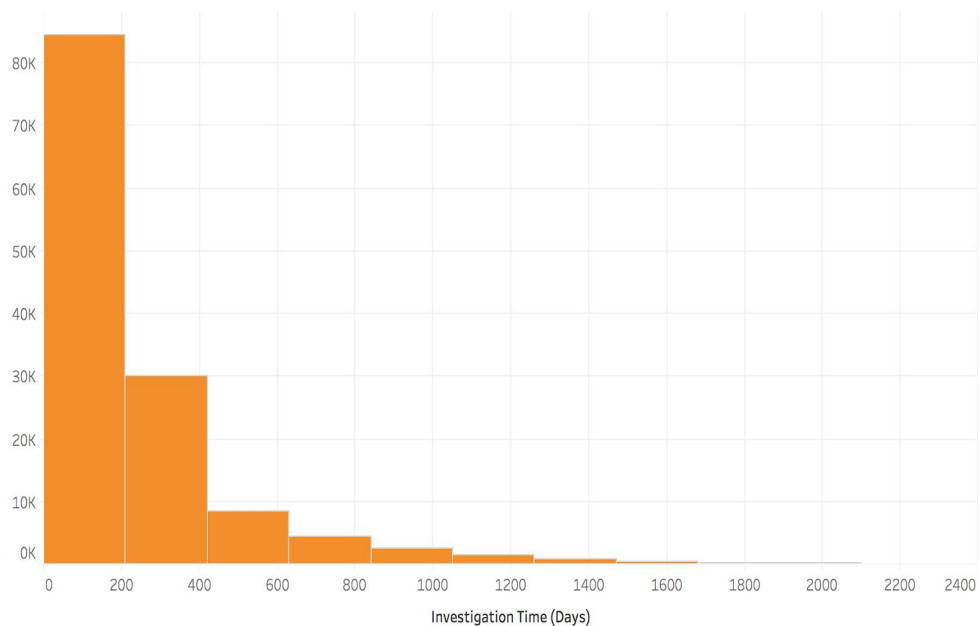
Tableau Chart #2:

Tableau Chart #2.1:

Results: Once the complaint has been completed it will be reviewed by a team to ensure accuracy and consistency through an investigation. Just because an allegation is made does not mean the complaint will be sustained. The resulting histogram shows the distribution of investigation times measured in days, where we can see most investigations happening between 0-200 days. An outlier of 8438 days exists, so we adjusted the bin values in Chart #2.1. Regarding our overarching theme, this helps to provide fine-grained information for our second baseline, established in Checkpoint #1 as an average investigation time of 267 days (or 8.7 months) for the entire population. **We will continue to explore this distribution in Question #4 further below.**

Question 3: What's the frequency of each disciplinary action?

Tableau Chart #3:

Disciplinary Action	frequency		Disciplinary Action	frequency
No Action Taken	214,458		365 Day Suspension	18
Unknown	8,446		28 Day Suspension	9
Reprimand	5,280		90 Day Suspension	9
1 Day Suspension	4,222		99 Day Suspension	9
2 Day Suspension	1,701		16 Day Suspension	7
3 Day Suspension	1,410		Suspended For 180 Days	6
5 Day Suspension	1,256		11 Day Suspension	5
Sustained-No Penalty	1,133		120 Day Suspension	5
Penalty Not Served	926		14 Day Suspension	5
Resigned -Not Served	900		150 Day Suspension	4
Violation Noted	827		40 Day Suspension	4
Resigned	644		Separated Other Case	4
10 Day Suspension	562		13 Day Suspension	3
Separation	502		17 Day Suspension	3
30 Day Suspension	413		21 Day Suspension	3
15 Day Suspension	349		29 Day Suspension	3
Suspended Over 30 Days	271		35 Day Suspension	3
20 Day Suspension	211		18 Day Suspension	2
4 Day Suspension	182		23 Day Suspension	2
Reinstated By Police Board	139		270 Day Suspension	2
Administrative Terminati..	132		9 Day Suspension	2
7 Day Suspension	105		180 Day Suspension	1
25 Day Suspension	69		22 Day Suspension	1
6 Day Suspension	49		24 Day Suspension	1
8 Day Suspension	34		27 Day Suspension	1
45 Day Suspension	31		31 Day Suspension	1
60 Day Suspension	23		326 Day Suspension	1
12 Day Suspension	21		540 Day Suspension	1
Reinstated By Court Action	21		75 Day Suspension	1
Suspended Indefinitely	21		900 Day Suspension	1

frequency: 1 — 214,458

Tableau Chart #3.1:

Results: After the complaint is sustained and the investigation happens, the officer is given some type of disciplinary action. We explored two visualizations to compare the frequencies of each disciplinary action, a frequency table and word cloud. The most common types of disciplinary action is "No Action Taken", followed by "Unknown", then "Reprimand" and "1 Day Suspension". Regarding our overarching theme, this helps to provide background information for our third baseline from Checkpoint #1. **We will continue to explore this distribution in Question #4 further below.**

**Question 4: What's the influence of victim race on the last two distributions?**

∵

Question 4a: What's the influence of victim race on the distribution of investigation time?

Interactive Visualization: Please find our interactive visualizations fully available at ObservableHQ [https://observablehq.com/@vbursztyn/interactive-visualizations-of], with all steps properly documented in that Notebook.

The four screenshots below (**from left to right, top to bottom: Black Victims, White Victims, Hispanic Victims, and Asian/Pacific Islander Victims**) illustrate our analysis, and the GUI component to select victim race is at the bottom of screenshots.
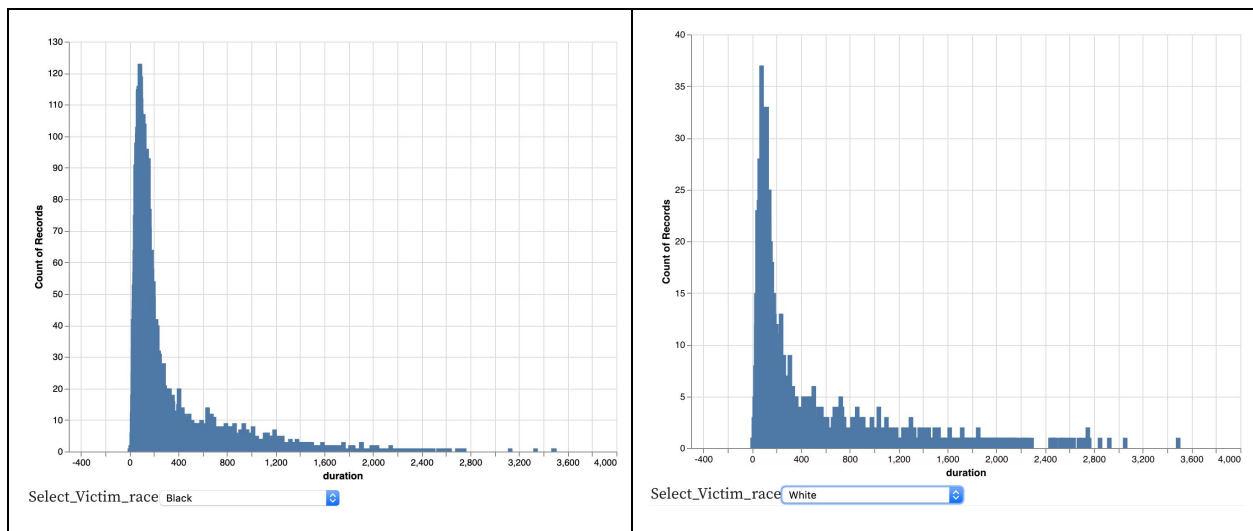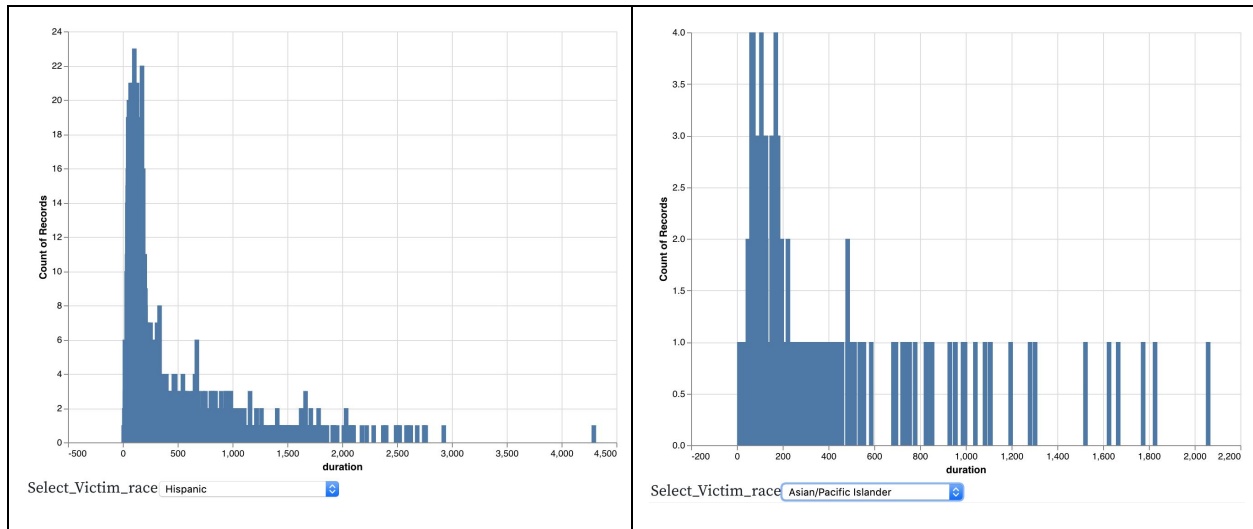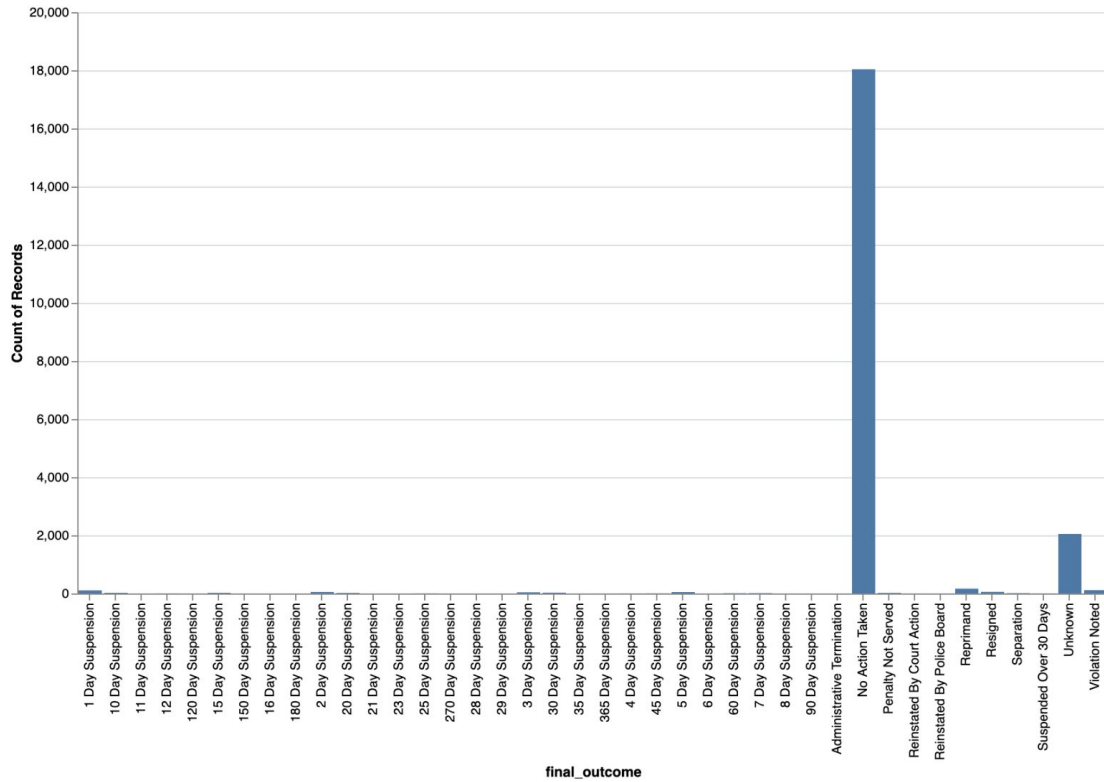
Results: To continue our exploration on Question #2, we resorted to an interactive visualization powered by Vega-Lite (Satyanarayan *et al.*, 2016). We interacted with the distribution of investigation times by filtering the underlying dataset according to victim race. **From this standpoint, however, we note that most of the more targeted distributions — for Black Victims, White Victims, and Hispanic Victims — have similar shapes.** Somewhat worth-noting is that Hispanic Victims seem to have a slightly higher concentration of cases taking more than 500 days to investigate, i.e., a slightly taller tail. We consider the fourth distribution — for Asian/Pacific Islander Victims — to be less reliable due to a smaller sample.

In general, until we use more quantitative measurements, we see no striking, globally evident sign of racial disparities in the duration of police misconduct investigations.

Question 4b: What's the influence of victim race on the distribution of disciplinary action?

Interactive Visualization: Please find our interactive visualizations fully available at ObservableHQ [https://observablehq.com/@vbursztyn/interactive-visualizations-of], with all steps properly documented in that Notebook.

The fours screenshots below (from top to bottom: **Black Victims and White Victims with all disciplinary actions included; then Black Victims and White Victim without majority class "No Action Taken" nor "Unknown"**) illustrate our analysis, and the GUI component to select victim race is at the bottom of screenshots.

Count of Records (y-axis, top chart: 0 to 20,000)

final_outcome (x-axis)

1 Day Suspension, 10 Day Suspension, 11 Day Suspension, 12 Day Suspension, 120 Day Suspension, 15 Day Suspension, 150 Day Suspension, 16 Day Suspension, 180 Day Suspension, 2 Day Suspension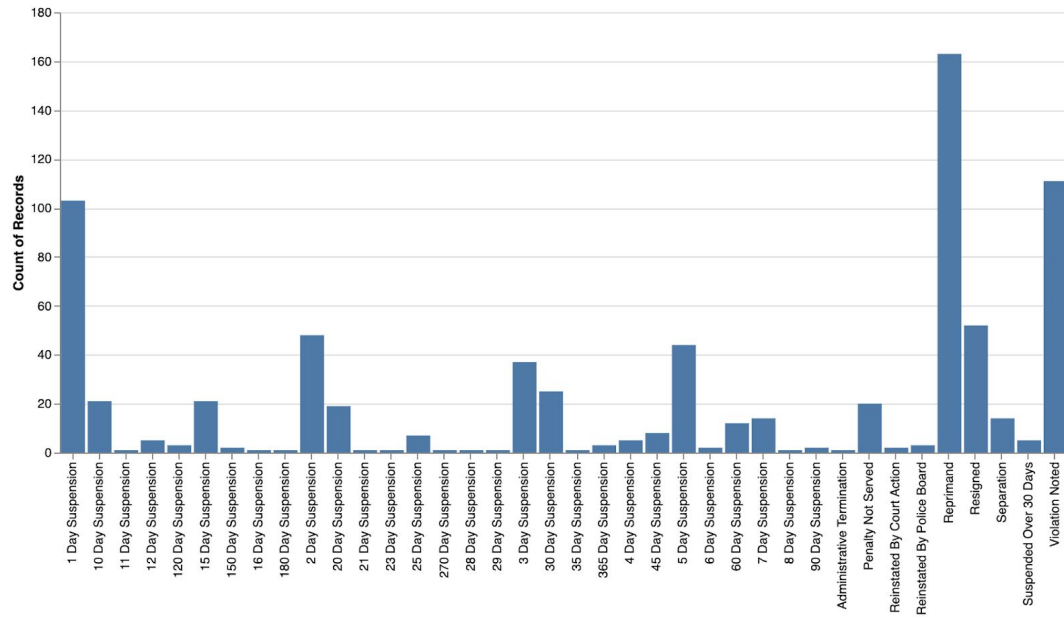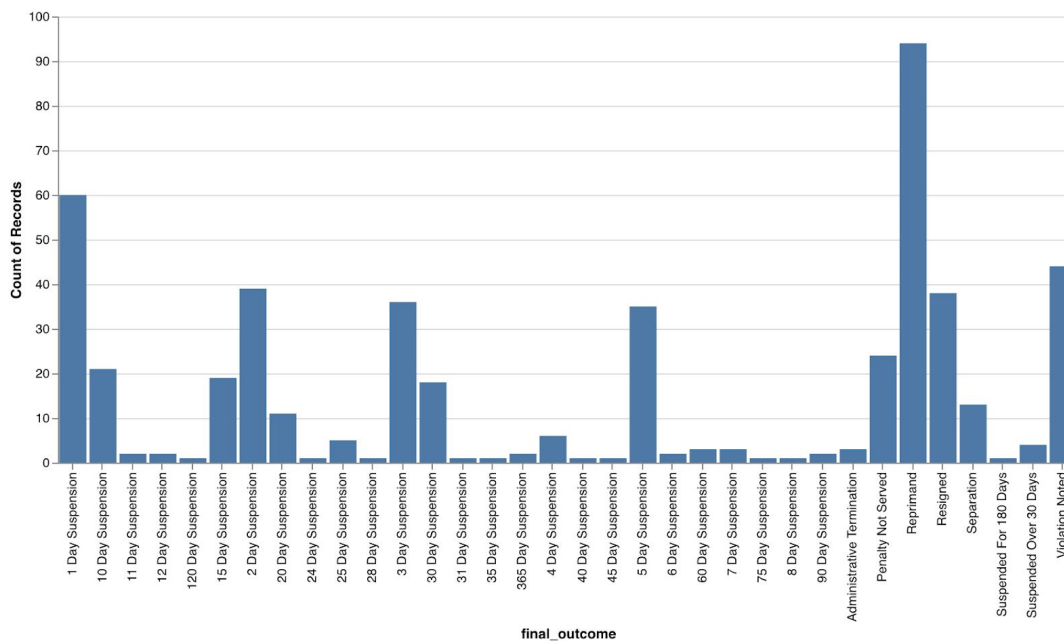, 20 Day Suspension, 21 Day Suspension, 23 Day Suspension, 25 Day Suspension, 270 Day Suspension, 28 Day Suspension, 29 Day Suspension, 3 Day Suspension, 30 Day Suspension, 35 Day Suspension, 365 Day Suspension, 4 Day Suspension, 45 Day Suspension, 5 Day Suspension, 6 Day Suspension, 60 Day Suspension, 7 Day Suspension, 8 Day Suspension, 90 Day Suspension, Administrative Termination, No Action Taken, Penalty Not Served, Reinstated By Court Action, Reinstated By Police Board, Reprimand, Resigned, Separation, Suspended Over 30 Days, Unknown, Violation Noted

Select_Victim_race [Black ▼]

Count of Records (y-axis, bottom chart: 0 to 4,000)

final_outcome (x-axis)

1 Day Suspension, 10 Day Suspension, 11 Day Suspension, 12 Day Suspension, 120 Day Suspension, 15 Day Suspension, 2 Day Suspension, 20 Day Suspension, 24 Day Suspension, 25 Day Suspension, 28 Day Suspension, 3 Day Suspension, 30 Day Suspension, 31 Day Suspension, 35 Day Suspension, 365 Day Suspension, 4 Day Suspension, 40 Day Suspension, 45 Day Suspension, 5 Day Suspension, 6 Day Suspension, 60 Day Suspension, 7 Day Suspension, 75 Day Suspension, 8 Day Suspension, 90 Day Suspension, Administrative Termination, No Action Taken, Penalty Not Served, Reprimand, Resigned, Separation, Suspended For 180 Days, Suspended Over 30 Days, Unknown, Violation Noted

Select_Victim_race [White ▼]

Select_Victim_race [Black ▾]

Select_Victim_race [White ▾]

Results: Similarly to Question #4a, we also continued our exploration on Question #3 by resorting to another interactive visualization powered by Vega-Lite (Satyanarayan *et al.*, 2016). We interacted with the distribution of disciplinary actions by filtering the underlying dataset according to victim race. Since "No Action Taken" is a majority class in this set of nominal values, our first attempts to visualize racial disparities were slightly overshadowed by this high-occurring mode value (see the first two screenshots above or refer to our Notebook at ObservableHQ). However, we could still see that the proportion

of cases leading to "Reprimanded" was somewhat higher among Black Victims. For this reason, we repeated the visualization after filtering both "No Action Taken" and "Unknown" from our dataset (see the last two screenshots above).

From this new standpoint, we could see distinctions more clearly. Although "Reprimanded" became the new upper bound in both segments of the population — Black Victims and White Victims —, **we can see that several mid-severity disciplinary actions are relatively more frequent among investigations of White Victims. For instance: White Victims and Black Victims have roughly the same absolute amounts of "15 Day Suspension" and "10 Day Suspension", despite the population of Black Victims being much larger.** In general, compared to Black Victims, White Victims seem more likely to have investigations leading to "x Day Suspension" with $x > 1$ (see "10 Day Suspension", "5 Day Suspension", "3 Day Suspension" and "2 Day Suspension").

Finally, regarding our overarching theme, **this is another indication of the potential influence of race and ethnicity in the outcomes of investigations**. As discussed in our proposal, more evidence will be gathered and analyzed in future checkpoints, including the presence of lurking variables.

References:

Satyanarayan, A., Moritz, D., Wongsuphasawat, K., & Heer, J. (2016). Vega-lite: A grammar of interactive graphics. IEEE transactions on visualization and computer graphics, 23(1), 341-350.