BRSM Project Report

Team Name: The Police did not investigate this murder

Radhikesh Agrawal
2021113013

Table of Contents

1. Topic Selection
2. A Brief Introduction about the study
    2.1. A Brief Description of the dataset
    2.2. Statistical Tests done in the study and Conclusions
    - Veil of Darkness test
    - Threshold test
    - Difference-in-differences strategy
3. Beyond the Paper - Our Team's Analysis
4. Hypothesis Tested
5. Data Preprocessing
6. Exploratory Analysis
7. Testing our Hypothesis
8. Appendix
9. References

Tip

If the report seems too long, you can go through these specific sections first

- 2.1
- 3
- 4
- 5
- 7

State Abbreviations

We used state state abbreviations instead of their full names as it helps to keep the body of the report concise and readable, especially when multiple states are mentioned frequently.

The abbreviations and corresponding state names are provided in the Appendix towards the end of the report.

Paper: A large-scale analysis of racial disparities in police stops across the United States

1. Topic Selection

> Martin Luther King Jr.
>
> I have a dream that my four little children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character.

Racial discrimination remains a widespread problem that affects many areas of society, slowing down progress and posing a serious threat to the principles of equality and justice.

In this Project, we intend to explore racial differences in police stops within the **United States**, using data from the ***Stanford Open Policing Project***. The nation's difficult history and ongoing issues with racial tensions make it a fitting context for such an examination.

Many of us might have faced some sort of discrimination in our lives, discrimination based on gender, religion, race and creed not only is a hindrance in development of a system but is also a serious threat to mankind in general. The United States, with its staggering contrast between races, and its history of racial discrimination offers a relevant setting for this research.

An inherent problem in doing a study on India is that the Indian population by large is very diverse and the notion of contrasting races is diminished. Drawing inspiration from prior research, we recognize that systematic racial discrimination in public services and bodies is a huge problem to justice and is a social evil at large. By carefully studying racial differences in police stops, we aim to shed light on this critical issue and contribute to the ongoing efforts toward a more fair and just society.

2. A Brief Introduction about the study

Each year in the United States, over 20 million traffic stops occur, making these interactions between police and the public extremely common. However, due to the decentralized nature of law enforcement agencies across the country and a lack of comprehensive, standardized data collection, it has been difficult to rigorously assess the role that race plays in traffic stops.
The most widely cited national statistics come from the **Police-Public Contact Survey (PPCS)**, which surveys around 50,000 people who report recently being stopped by police. While this provides some data, surveys can suffer from issues like selection bias and recall errors.

Some local and state police agencies have periodically released reports on traffic stops in their jurisdictions and made data available to outside researchers. However, this data is limited, providing only a partial picture. The data comes from select agencies, lacks comprehensive details, and is inconsistent across different jurisdictions.

2.1 A Brief Description of the dataset

To try to address these challenges, the researchers compiled and analyzed a massive dataset of nearly **100 million** traffic stops conducted by **21 state patrol agencies** and **35 municipal police departments** over almost a **decade**. This dataset was built by filing public records requests in all 50 states. The records were then redistributed in a standardized format to facilitate future analysis by researchers.

- Each data point captures a single stop with various factors like time, date, driver demographics (age, race, sex), citation information, vehicle type, "raw defendant race" etc.
- For consistency, the authors further restrict to stops occurring in 2011–2018, as many jurisdictions did not provide data on earlier stops.
- The authors limit their analysis to drivers classified as white, black or Hispanic, as there were relatively few recorded stops of drivers in other race groups.

Some of the factors present in the dataset

Column name	Column meaning	Example value
date	The date of the stop, in YYYY-MM-DD format.	"2017-02-02"
time	The 24-hour time of the stop, in HH:MM format.	20:15
subject_age	The age of the stopped subject.	54.23
subject_race	The race of the stopped subject. Values are standardized to white, black, hispanic, asian/pacific islander, and other/unknown	"hispanic"
subject_sex	The recorded sex of the stopped subject.	"female"
violation	Specific violation of stop where provided. What is recorded here varies widely across police departments.	"SPEEDING 15-20 OVER"
arrest_made	Indicates whether an arrest made.	FALSE
citation_issued	Indicates whether a citation was issued.	TRUE
warning_issued	Indicates whether a warning was issued.	TRUE
outcome	The strictest action taken among arrest, citation, warning, and summons.	"citation"
contraband_found	Indicates whether contraband was found.	FALSE
search_conducted	Indicates whether any type of search was conducted, i.e. driver, passenger, vehicle.	TRUE

2.2 Statistical Tests done in the study and Conclusions

The paper took a different route from the usual testing methods like t-tests or ANOVA. Instead, it delved into the realm of specialized techniques uniquely suited to its domain. These approaches, while perhaps not widely known, showcase a profound depth of insight and a knack for innovative analysis.

2.2.1 *Veil of Darkness test*

The authors first assessed potential racial bias in police stop decisions by applying the ***veil of darkness test*** developed by **Grogger** and **Ridgeway**. This test is based on the observation that since sunset times vary throughout the year, one can examine the racial breakdown of stopped drivers as a function of daylight while controlling for the time of day. Specifically, they utilized the discontinuity created by the start and end of daylight saving time (DST). They compared the racial distribution of drivers stopped in the evenings right before DST began (when it was dark) to the distribution after DST began (when it was light during the same times). If a smaller proportion of stopped drivers were black when it was dark out and therefore harder to determine race, that suggests black drivers were being stopped during daylight hours in part due to their race.

Conclusion from the test

In both state patrol and municipal police stops, they find that black drivers comprise a smaller proportion of drivers stopped after sunset, suggestive of discrimination in stop decisions.

2.2.2 *Threshold test*

Secondly, the researchers investigated potential racial bias in the decision by officers to search drivers for contraband after making a stop. To do this, they applied the ***threshold test*** developed by **Simoiu et al**. and refined by **Pierson et al**. This test incorporates both the rate at which searches occur as well as the success rate of those searches in order to infer the standard of evidence used to determine who to search. Building on traditional outcome analysis methods, a lower search success rate for one group compared to another suggests bias, as it implies a lower evidentiary bar was used for that group when deciding to conduct searches.

Conclusion from the test

The threshold test indicates that black and Hispanic drivers were searched on the basis of less evidence than white drivers, both on the subset of searches carried out by state patrol agencies and on those carried out by municipal police departments

2.2.3 *Difference-in-differences strategy*

Lastly, the researchers examined how drug policy impacts racial disparities in outcomes from traffic stops. They compared policing patterns in **Colorado** and **Washington**, the two states that legalized recreational marijuana at the end of 2012, to patterns in 12 states where recreational marijuana remained illegal during the study period. Utilizing a ***difference-in-differences strategy***.

Conclusion from the test

They found that legalization in **Colorado** and **Washington** was associated with reduced rates of vehicle searches and misdemeanor drug offense rates for white, black and Hispanic drivers compared to the states without legalization

3. Beyond the Paper - Our Team's Analysis

In addition to the analyses presented in the original paper, our study explored two further areas.

Firstly, while the paper exclusively investigated racial disparities in traffic stops, we expanded our scope to examine potential **gender-based disparities** as well. This allowed us to analyze whether significant differences existed in how police interactions played out for individuals of different genders during traffic stops.

Secondly, the paper conducted an overarching analysis to identify disparities across the **United States** collectively. However, our study went a step further by performing a state-level analysis. Rather than just looking at national trends, we analyzed the data regionally to uncover any Geographic variations and differences in policing patterns between distinct regions of the United States. By taking this granular regional approach, we could pinpoint areas where disparities may be more or less pronounced compared to others.

There were two parts to our analysis.

For the first part, we were quite skeptical about making certain assumptions regarding the data, as they did not seem appropriate or justified. Consequently, we conducted our analysis while avoiding those assumptions.

However, after receiving feedback from the professors, they instructed us to proceed by making those assumptions for the sake of the project's objectives. This would enable us to carry out additional analyses that were previously precluded due to our reluctance to make those assumptions. In the report, we will incorporate everything together and provide justifications for our initial skepticism as well.

4. Hypothesis Tested

4.1 Hypothesis 1 - Disparity in Searches

Null Hypothesis (H_0) **:** The distribution of drivers stopped and searched by police is independent of the driver's race across the states.

Alterntive Hypothesis (H_1) The distribution of drivers stopped and searched by police differs by the driver's race across the states, with certain racial groups being stopped more frequently.

4.2 Hypothesis 2 - Disparity in Outcomes

Null Hypothesis (H_0) **:** The outcome of stops does not depend on or relate to the race of drivers across the states.

Alternative Hypothesis (H_1) **:** The outcome of stops depends on or relates to the race of drivers, with certain racial groups experiencing different outcomes more frequently in certain states.

4.3 Hypothesis 3 - Gender Disparity

Null Hypothesis (H_0) **:** The distribution of drivers stopped by police is independent of the driver's sex across the states.

Alternative Hypothesis (H_1) **:** The distribution of drivers stopped by police differs by the driver's sex across the states, with certain sex groups being stopped more frequently.

Work done by each Team Member

Figure

Workdone	By
Data Preprocessing	Arnav and Myself
Analysis of Hypothesis 1	Myself
Analysis of Hypothesis 2	Arnav
Analysis of Hypothesis 3	Brahad
Data Analysis	Everyone

I'll talk in detail about the work I've done, i.e. Data Preprocessing and Analysis of Hypothesis 1. For the rest of the work, I'll mention in brief.

5. Data Preprocessing

Data preprocessing consumed a significant portion of our effort. We did most of that data preprocessing in python due to higher familiarity and ease as compared to R.

1. **Data Volume and Efficient Code**
   The sheer volume of data that we had required efficient code to handle the preprocessing tasks was in GBs. Therefore, we needed to be sure that the code we write did not contain redundancies and was efficient.
2. **Variable Selection and Aggregation**
   A critical aspect of data preprocessing involved was determining which variables to retain and what aggregations to perform.
3. **Memory Constraints**
   The CSV files containing the raw data were often too large to fit into available memory. This limitation posed a significant challenge, as it prevented the simultaneous loading and processing of the entire dataset. To overcome this obstacle, we adopted a streaming approach, reading and processing the data one file at a time, extracting and retaining only the necessary information in a single pass.
4. **Variable Inconsistency across States**
   The data exhibited inconsistencies in the presence of variables across different states. To create a cohesive and representative dataset, we identified regions within each state where a common set of variables was available. ***This approach allowed us to construct a dataset that captured the diversity of the United States while maintaining consistency in the analyzed features.***
5. **Key Variables for Analysis**
   For our primary analysis, we focused on specific variables: `subject_race`, `subject_sex`, `search_conducted`, `arrest_conducted`, and `contraband_found`. These variables provided critical insights into potential patterns and trends related to law enforcement practices and related to our proposed hypothesis
6. **Representativeness and State Coverage**
   To ensure the dataset's representativeness and capture a comprehensive picture of the United States, we aimed to include data from as many states as possible. By incorporating a diverse range of state-level data, we could enhance the generalizability of our findings and conclusions.
7. **Temporal and Age-related Patterns**
   In addition to the core variables for analysis, we also explored `date` and

`subject_age` variables. These variables enabled us to investigate potential temporal trends and patterns in law enforcement practices concerning age demographics.

6. Exploratory Analysis

Age Analysis

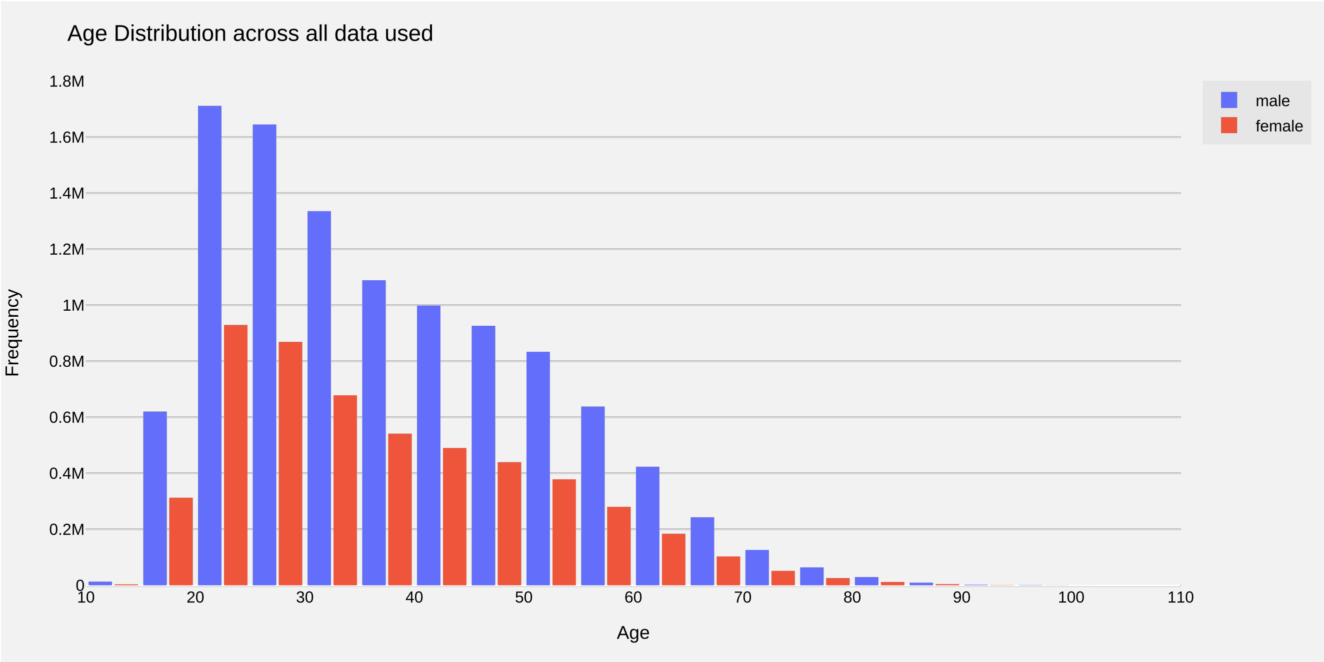

Fig. Age distribution of Stops made across all states

1. **Age Disparity in Stop Rates**
   The graph reveals a higher rate of stops for younger individuals, particularly those under the age of 30. This pattern could be attributed to several potential factors:
   - Younger individuals may be more likely to engage in behaviors that draw police attention, such as traffic violations, petty crimes, or being in high-crime areas. The impulsivity and risk-taking behavior often associated with youth could contribute to this.
   - Police officers may have implicit biases or preconceptions that lead them to scrutinize younger individuals more heavily, regardless of actual behavior. This could be a result of societal stereotypes or past experiences.
   - Younger individuals may be more likely to be in situations or environments where police presence and enforcement is higher, such as in urban areas or near entertainment venues.
2. **Gender Differences in Stop Rates**
   According to the data represented in the graph, females were subjected to fewer stops than males. This disparity may be attributable to some potential factors:
   - A lower number of females operating vehicles on the roads compared to their male counterparts.
   - Differences in stop rates between genders could be influenced by societal norms and expectations, where females may be perceived as less threatening or engaging in less "risky" behavior, leading to lower stop rates.
   - Factors such as socioeconomic status, race, and neighborhood crime rates could also play a role in shaping the patterns observed in the data, though these factors are not directly evident from the information provided.

3. **Similarity in Age Trends across Genders**
   Despite the observed differences in stop rates between males and females, the age trend appears to be remarkably similar for both genders. This similarity suggests that, at a preliminary glance, there is no apparent gender disparity in police stops based solely on age factors.

Temporal Analysis

1. **Temporal Stability in Police Stops** For most states, the number of police stops remained relatively consistent over time, with no major increases or decreases observed.
2. **Racial Variations in Stop Patterns**
   - For white individuals, some states experienced significant fluctuations in the number of stops, with notable increases and decreases.
   - For other racial groups, the number of stops tended to be consistently low, with a few exceptions, such as Pennsylvania, which has a larger Black population.

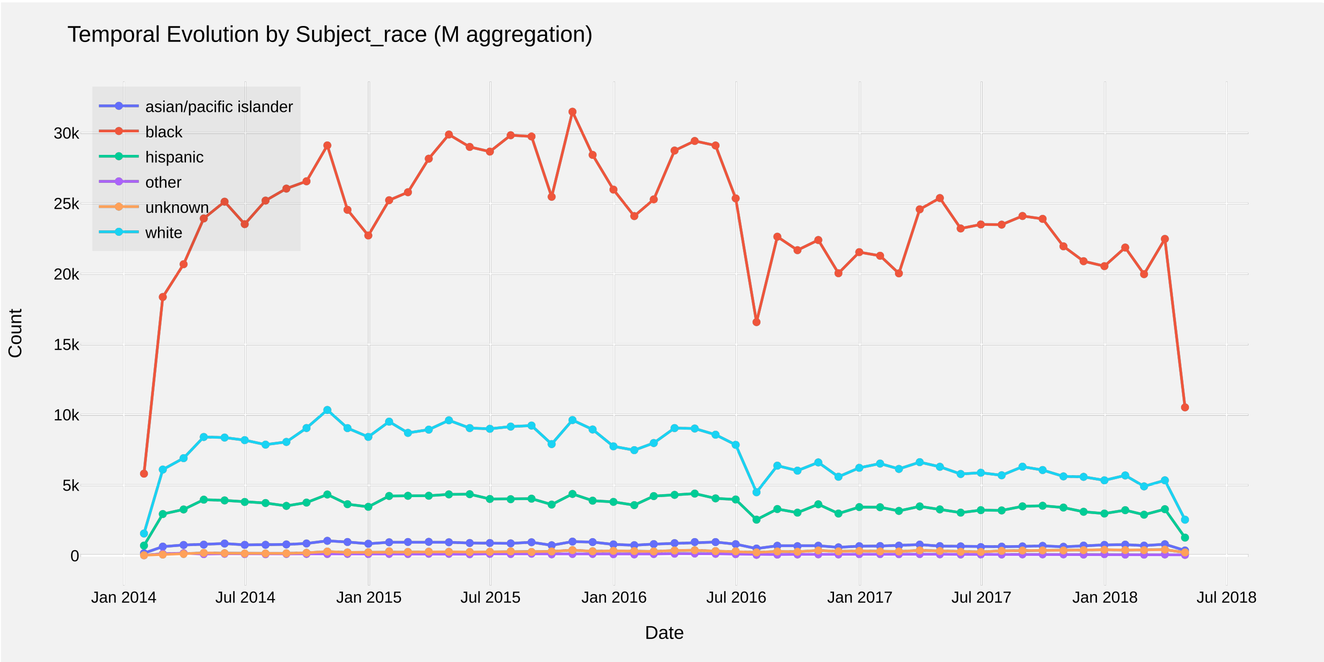

Fig. Time Evolution of Stops by Race for Pennsylvania

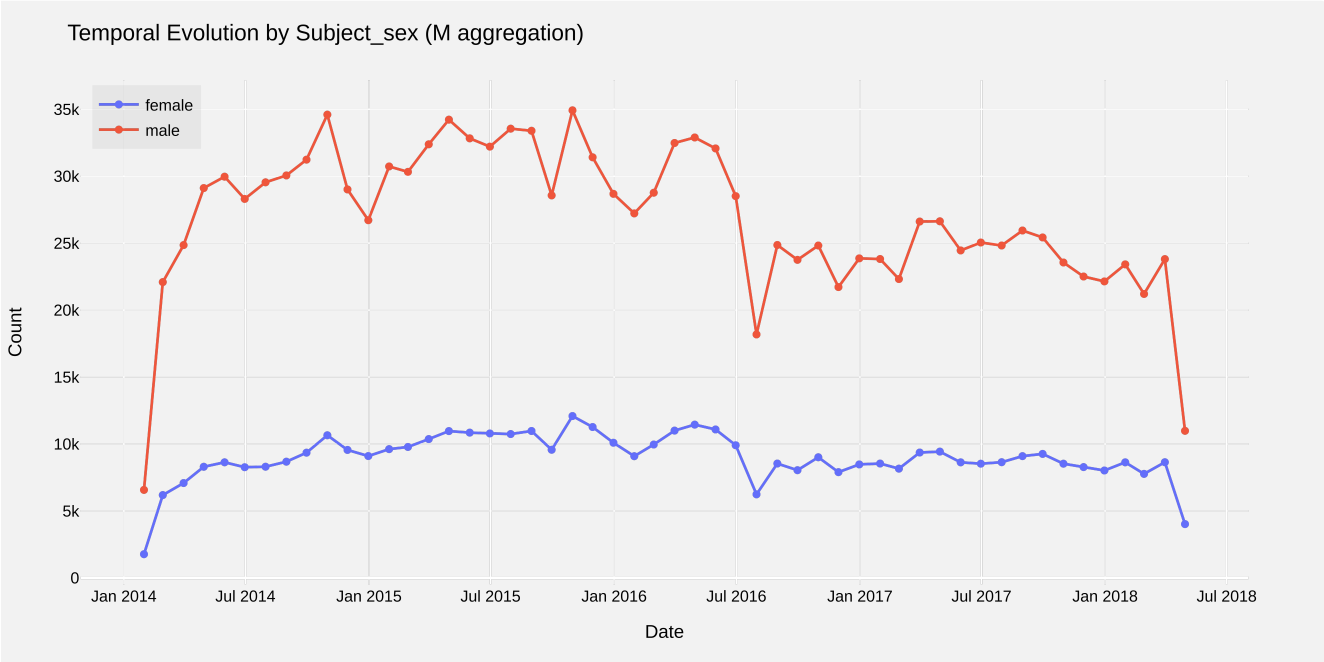

Fig. Time Evolution of Stops by Sex for Pennsylvania

3. **Gender Variations in Stop Patterns** The trend for stops based on gender follows the racial variation in stops
4. **State-Specific Trends** In certain states, like Tennessee, the number of stops initially increased and then decreased towards 2018, suggesting changes in policing practices or priorities during that time period in those particular locations.
5. **Decrease in Stops in 2014** In many states, including Arizona, California, Connecticut, North Carolina, Tennessee, Vermont, and Wisconsin, there was a decrease in police stops observed in 2014. This decline was likely attributed to high-profile incidents that occurred around that time, such as the death of Eric Garner and the shooting of a 12-year-old boy in Cleveland, Ohio, by a police officer.

It's important to note that these are potential explanations, and further analysis and contextual information would be needed to draw more definitive conclusions about the underlying factors driving the observed trends.

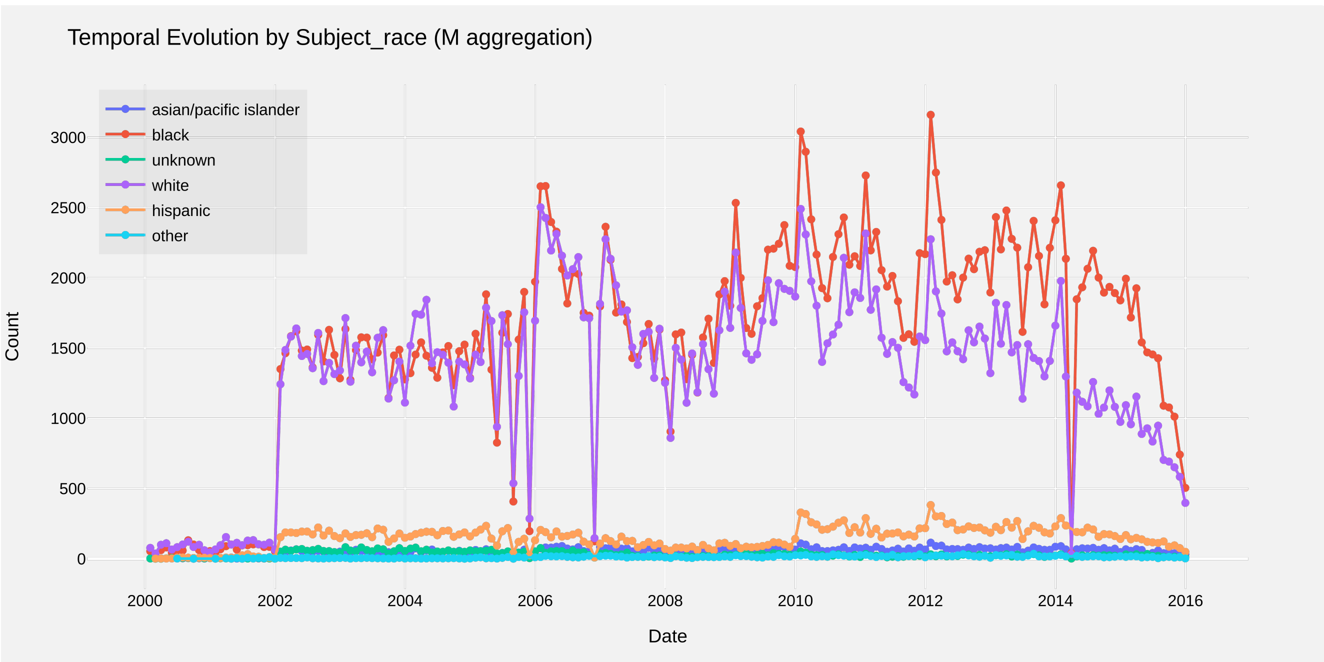

Fig. Descrease in stop in 2014 for North Carolina

In 2014, several significant events in the United States contributed to a heightened awareness and tension over policing, particularly concerning racial disparities and the use of force by law enforcement. This period saw a series of high-profile incidents and subsequent protests that called for reform in policing practices.

- The death of Eric Garner in Staten Island, New York, after an altercation with police, where the officer was not indicted for using an unauthorized chokehold, sparked widespread protests and demands for police reform.
- The fatal shooting of John Crawford III in a Walmart in Ohio, where he was carrying a BB gun, and the shooting of a 12-year-old boy in Cleveland, Ohio, by a police officer, further fueled the national conversation on police accountability and the use of force.
- The shooting of Michael Brown in Ferguson, Missouri, by a white police officer, Darren Wilson, after a struggle and foot chase, ignited months of intense protests and calls for reform. This incident, along with the subsequent protests and investigations, highlighted the racial tensions and the need for change in policing practices.

7. Testing our Hypothesis

7.1 Hypothesis 1

After stopping a driver, officers may carry out a search of the driver or vehicle if they suspect more serious criminal activity. We try to analyze how this distribution varies across various US states for different races.

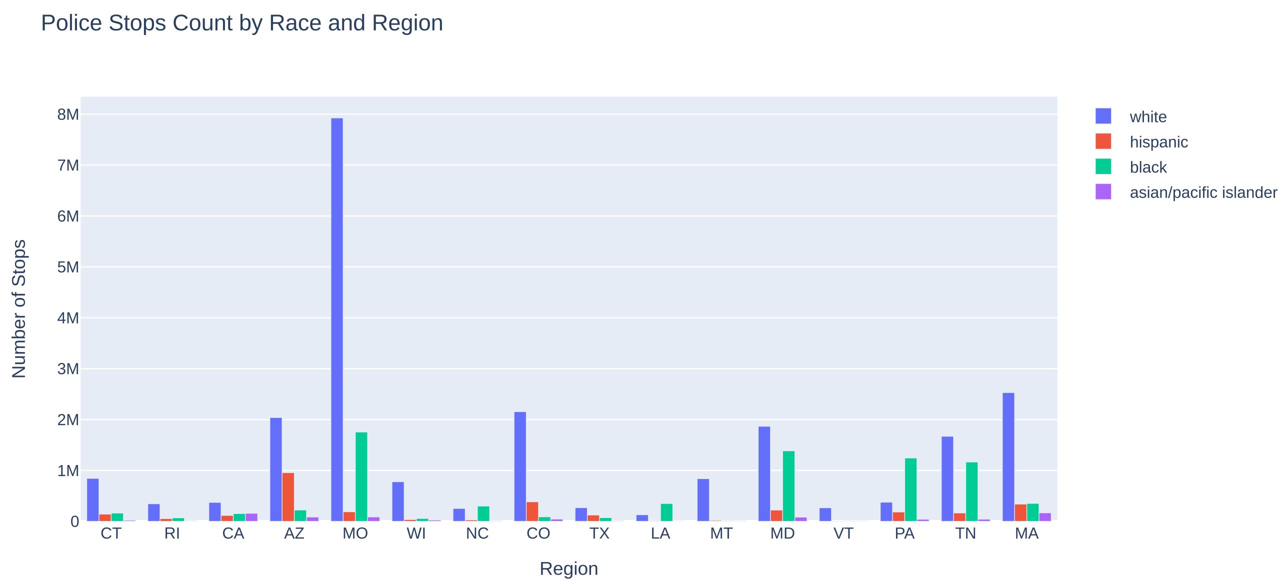

Fig. Police Stops Demographics

Check Miscellaneous Section for US State Fullforms

The above plot illustrates the variation in police stops across different regions. Individuals identified as White had the highest rate of stops, followed by those identified as Black. A potential factor contributing to these statistics could be the demographic makeup of many states, where White and Black populations tend to be larger than other racial/ethnic groups.

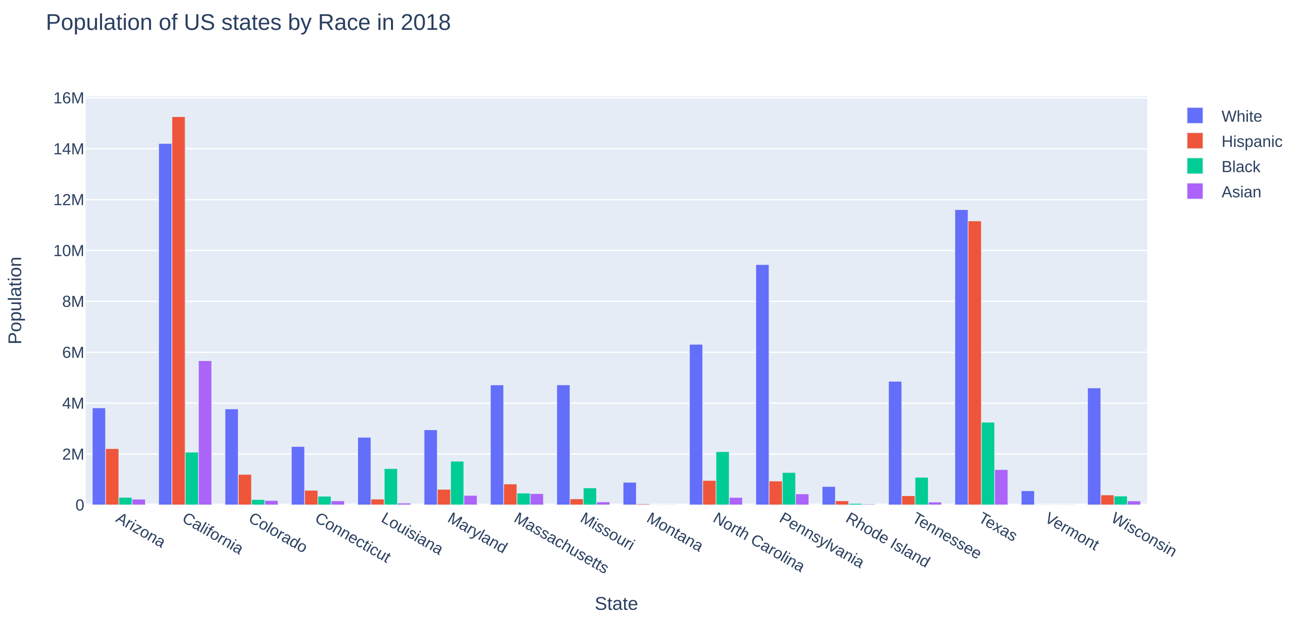

Fig. Population of US states in 2018

The above plot shows demographics of some of the states in 2018. This has been the trend more or less in the past 5 to 6 years too.

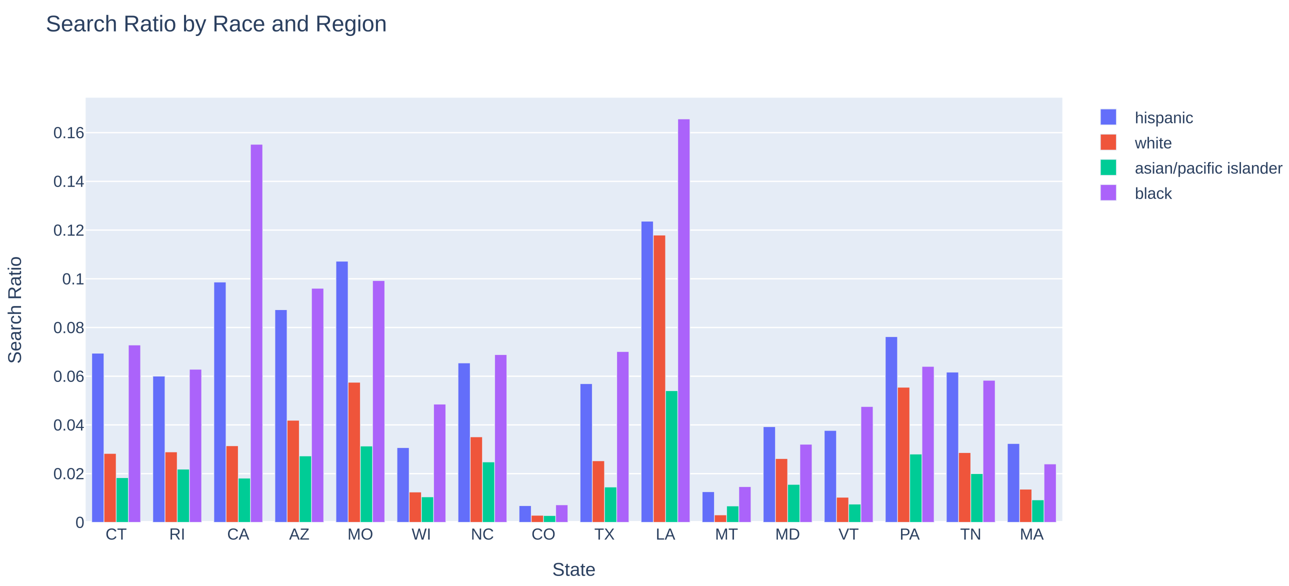

Fig. Search Rates

The above plot shows how search rates varies across different states and races. The search rates are calculated as

$$Search\ Rate(race, region) = \frac{\#\ of\ Searches(race, region)}{\#\ of\ Stops(race, region)}$$

and the hit rates are calculated as

$$Hit\ Rate(race, region) = \frac{\#\ of\ Contraband\ Found(race, region)}{\#\ of\ Searches(race, region)}$$

The above plot on search rates indicates that individuals identified as Black experienced higher rates of searches compared to those identified as Hispanic or White, despite having a smaller overall population proportion. This observation raises potential concerns about disparities in search practices across racial/ethnic groups.

However, further statistical analysis would be necessary to determine if these differences are statistically significant and to understand the complex factors that may contribute to such patterns. It's important to approach this data objectively and avoid drawing premature conclusions about the underlying causes, which may be multifaceted and require a nuanced examination. I'll talk about some of these nuances in a while.

Note that directly comparing the raw numbers of stops or searches would be misleading, as the overall population sizes differ between these groups. To allow for a proper comparison, analyzing the rates rather than absolute numbers is crucial, as this helps normalize the data and account for the varying population baselines.

Another thing to note is that the disparities we see in search rates are not necessarily the product of discrimination. Black and Hispanic drivers might, hypothetically, carry contraband at higher rates than white drivers, and so elevated search rates may result from routine police work even if no racial bias were present. To measure the role of race in search decisions, we apply a statistical strategy: ***outcome analysis***.

To assess whether the gaps resulted from biased decision-making, we apply the outcome test, originally proposed by **Becker** to circumvent omitted variable bias in traditional tests of discrimination. The outcome test is based not on the search rate but on the "**Hit Rate**": *the proportion of searches that successfully turn up contraband*. If searches of minorities are successful less often than searches of whites, it suggests that officers are applying a double standard, searching minorities on the basis of less evidence.

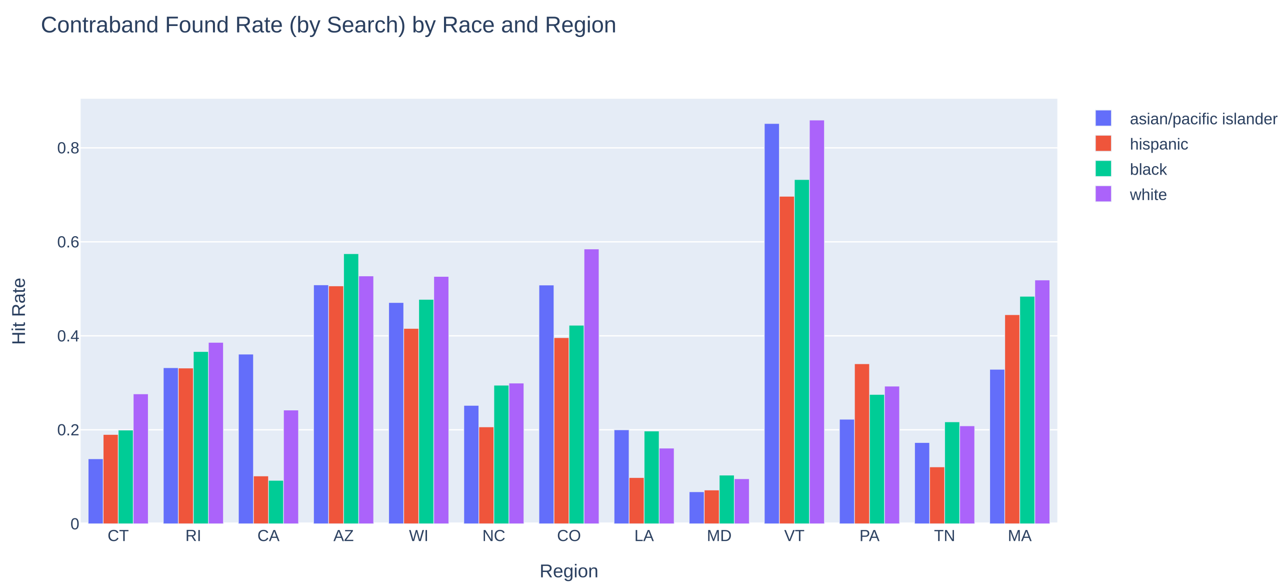

Fig. Hit Rates

The graph indicates that within a given state, the hit rates, or rates of finding contraband during searches, are relatively consistent across racial/ethnic groups. However, there appears to be variation in the hit rates across different states.

At first glance, this data does not seem to suggest disparities in search outcomes based on race/ethnicity within each state. Instead, the observed differences appear to be at the state level, where some states have higher overall hit rates than others, regardless of the racial/ethnic composition.

7.1.1 Analysis of Data

A) Search Rate

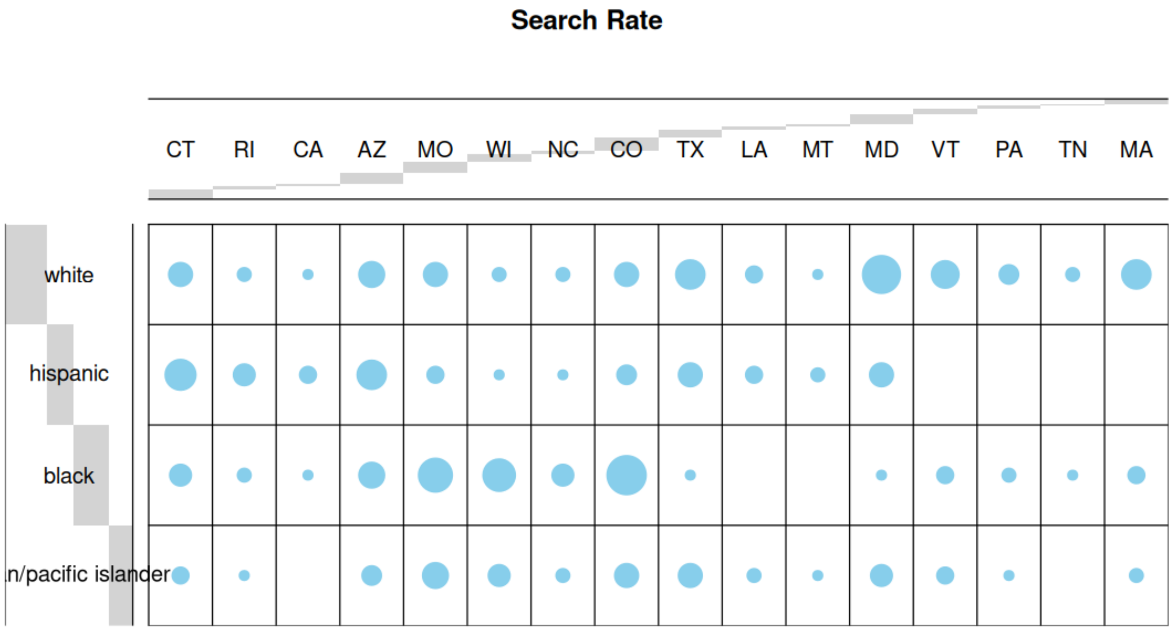

Fig. Bubble Plot for Search Rate

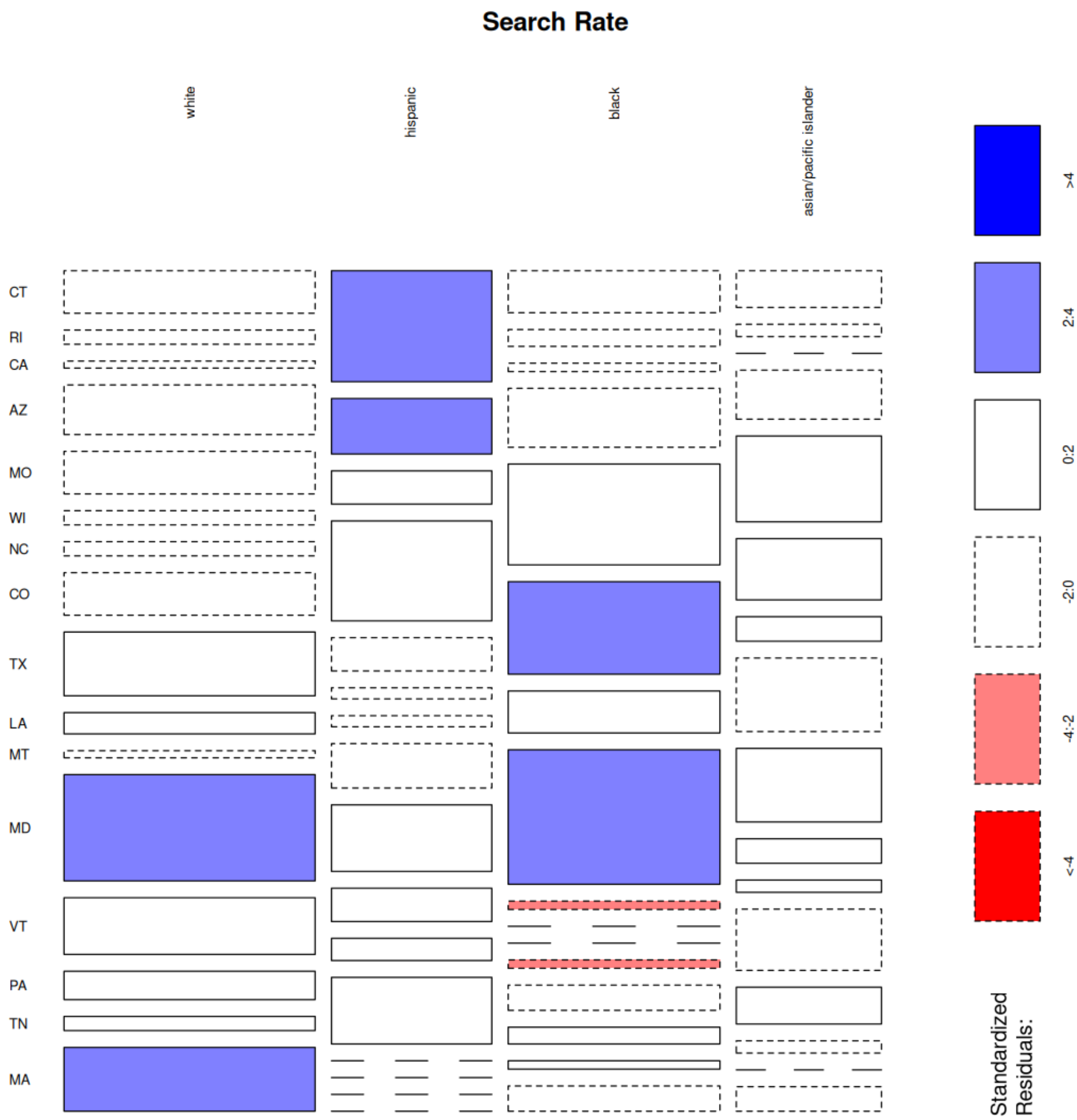

Fig. Mosaic Plot for Search Rate

Standard residuals are calculated based on the expected frequencies under the assumption of independence between the two variables. A standard residual greater than 2 (or less than -2) typically indicates a significant deviation from independence.

In essence, they tell you how much each cell's observed frequency deviates from what would be expected if the two variables were independent. Positive residuals suggest over-representation, while negative residuals suggest under-representation.

And we can say by looking at the plot above that there are certain residuals greater than 4 (> 2) which suggests that there is a significant deviation from independence for them.

B) Hit Rates

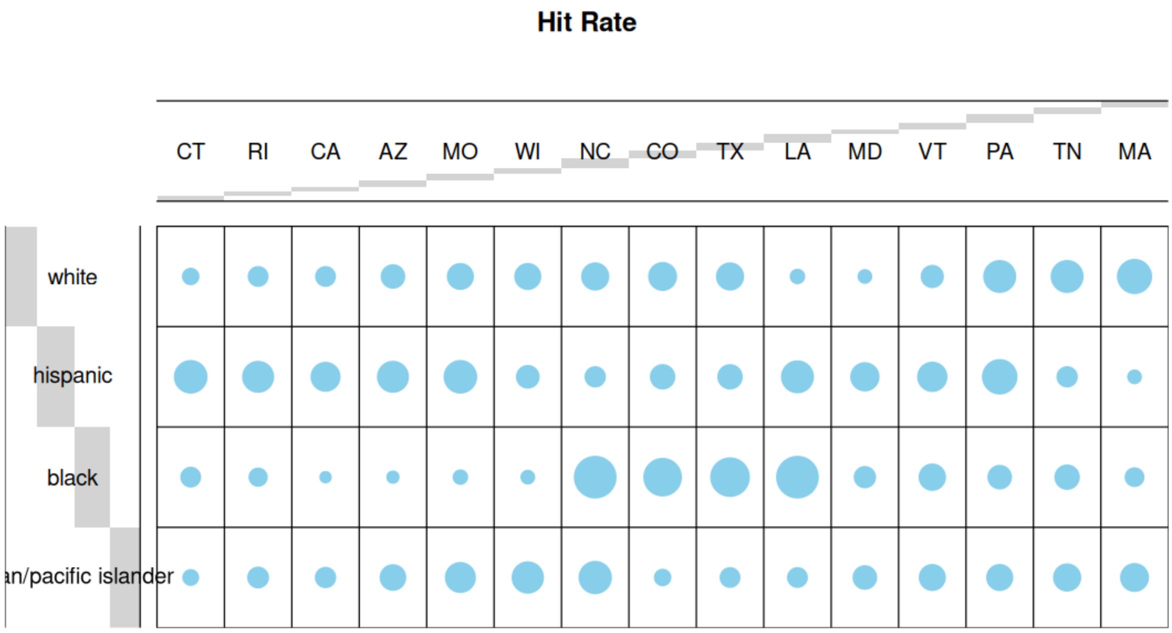

Fig. Bubble Plot for Hit Rate

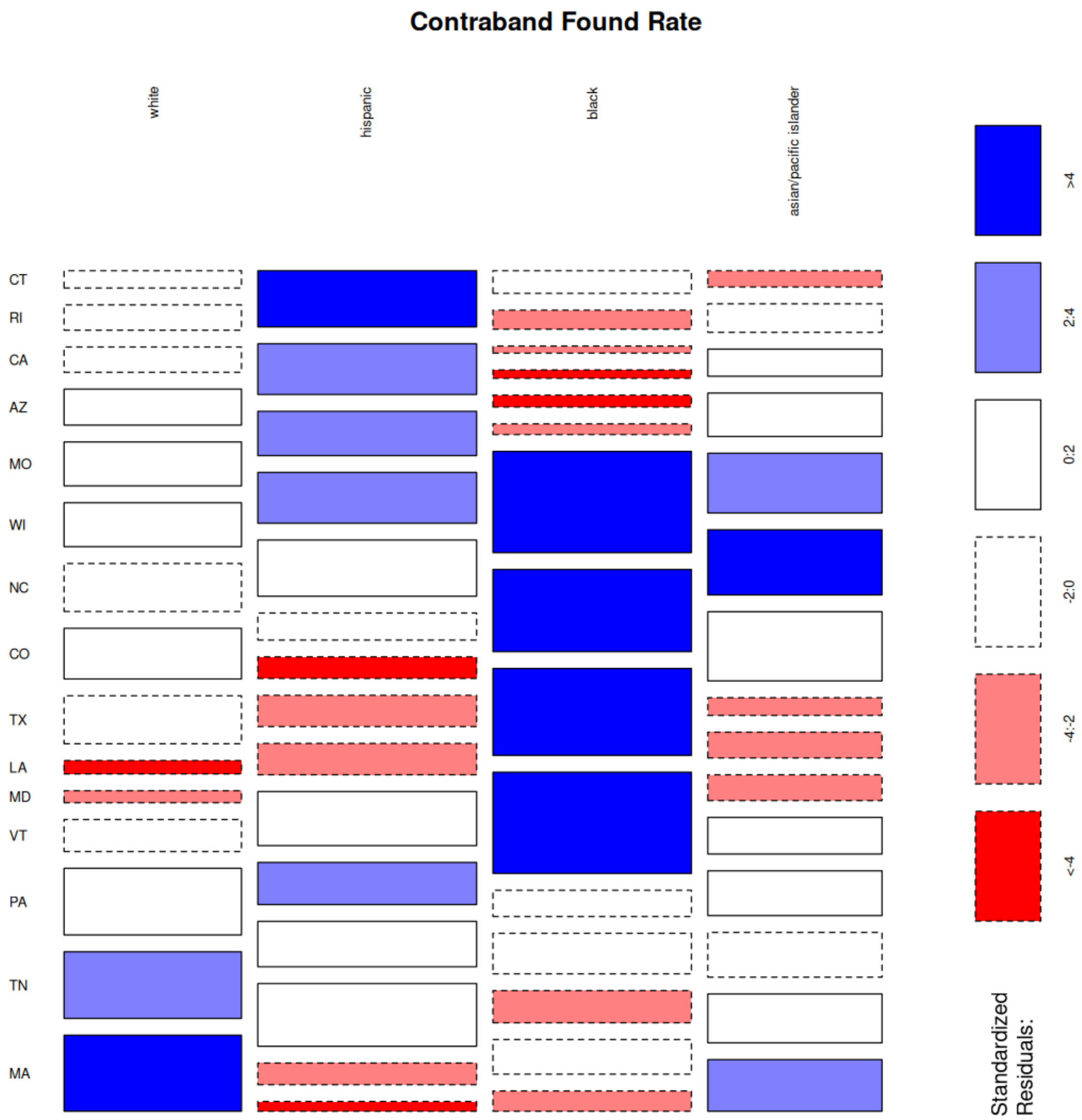

Fig. Mosaic Plot for Search Rate

By looking at the plot above, we can say that there are certain residuals greater than 4 (> 2) which suggests that there is a significant deviation from independence for them

7.1.2 Tests Conducted

A) Search Rates

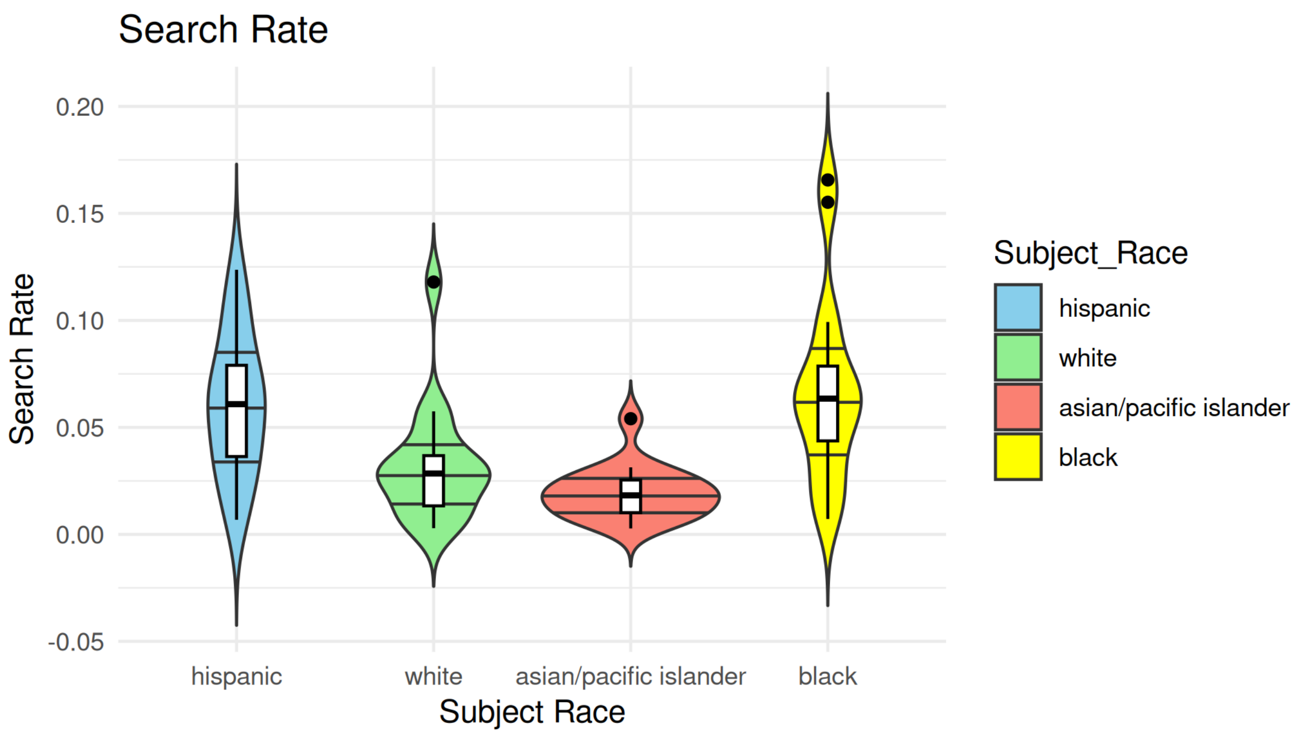

Fig. Violin Plot for Search Rate

1. χ^2 Test of Independence

I applied this on the the Search rate data. The columns of my data were the different races (Categorical Variable 1) and the rows of my data were different states (Categorical Variable 2).

Note that χ^2 tests are not meant for ratios, so I multiplied by a fixed number, 100, to get frequencies which is what χ^2 test requires. Note that it is not wrong to do this because our data is of the order of at least 10^5, so multiplying it by 100 is no big deal.

1.1 Contingency Table

```
                        CT RI CA AZ MO WI NC CO TX LA MT MD VT PA TN MA
white                    6  2  1  7  6  2  2  6  9  3  1 15  8  4  2  9
hispanic                10  5  3  9  3  1  1  4  6  3  2  6  0  0  0  0
black                    5  2  1  7 12 11  5 16  1  0  0  1  3  2  1  3
asian/pacific islander   3  1  0  4  7  5  2  6  6  2  1  5  3  1  0  2
```

1.2 Test Result

```
Pearson's Chi-squared test

data:  contingency_table
X-squared = 88.81, df = 45, p-value = 0.0001072
```

1.3 Conclusion and Interpretation

$p < 0.05$ so we can say that the variables are not independent of each other. This gives us a green flag to do further analysis on our data

2. ANOVA

2.1 Check for Normality: Shapiro-Wilk Normality Test (N<50)

Subject Race	W	p-value
hispanic	0.96892	0.8208
white	0.7887	0.001935
asian/pacific islander	0.84059	0.009936
black	0.91408	0.1354

NORMALITY VIOALTED
Since not all of the groups satisfy the normality condition, we'll have to apply non-paramateric ANOVA - *Kruskal Wallis ANOVA*

2.2 Kruskal-Wallis ANOVA

```
data:  Search_by_Stop by Subject_Race
Kruskal-Wallis chi-squared = 22.857, df = 3, p-value = 4.326e-05
```

NULL HYPOTHESIS REJECTED
$p < 0.05$ means that there is variance in means of the groups. We can proceed to apply post-hoc tests

2.3 Post-Hoc Test - Dunn's Test

Comparison	Z	P	Adjusted P
asian/pacific islander - black	-4.0076817	3.065884e-05	0.0001839531
asian/pacific islander - hispanic	-3.8737735	5.358151e-05	0.0003214891
black - hispanic	0.1339082	4.467376e-01	1.0000000000
asian/pacific islander - white	-1.4347309	7.568190e-02	0.4540913731
black - white	2.5729508	5.041777e-03	0.0302506646
hispanic - white	2.4390426	7.363117e-03	0.0441787019

Individuals identified as `Asian/Pacific Islander` experienced lower search rates compared to all other groups. The analysis further indicates that the differences in search rates between `Asian/Pacific Islander` and `Black` or `Hispanic` groups were statistically significant. However, the difference between `Asian/Pacific Islander` and `White` groups did not reach statistical significance. Additionally, the search rates for individuals identified as `Black` and `Hispanic` were higher than those identified as `White`.

3. Pairwise Permutation using BH Correction

Comparison	Stat	p.value	p.adjust
hispanic - white	2.37	0.01777	0.026660

Comparison	Stat	p.value	p.adjust
hispanic - asian/pacific islander	3.591	0.0003297	0.001978
hispanic - black	-0.5522	0.5808	0.580800
white - asian/pacific islander	1.655	0.09792	0.117500
white - black	-2.467	0.01363	0.026660
asian/pacific islander - black	-3.391	0.0006956	0.00208

Surprisingly, permutation test gives same result as the Dunn's post-hoc test.

- Permutation tests provide a robust alternative to traditional parametric tests, especially when assumptions such as normality or homogeneity of variance are not met and in our case, both of them were violated.
- By randomizing the data through permutations, you eliminate the risk of biased estimates or inflated Type I error rates. This ensures that the significance we observe is not due to spurious correlations or chance occurrences in the data.

This enforces our findings even further

Conclusion

The analysis indicates that there is a correlation between geographic region and search rates, which aligns with the patterns visible in the graphs. Also, across states, the data shows that individuals identified as `Asian/Pacific Islander` and `White` experienced lower rates of police stops compared to those identified as `Black` and `Hispanic`. This suggests potential racial/ethnic disparities in police stop practices.

However, as mentioned earlier, this is not enough to arrive at a definitive conclusion. Therefore we'll do our analysis further on Hit Rates and see what results we get.

B) Hit Rates

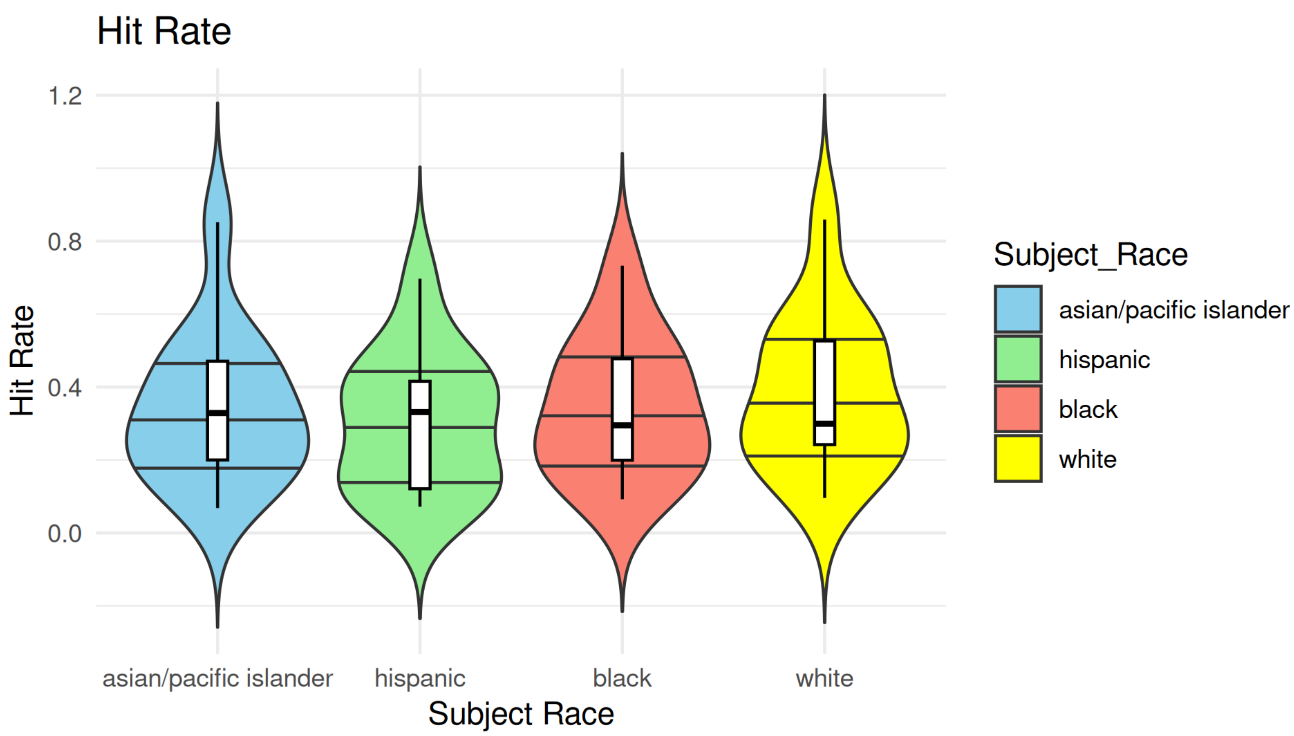

Fig. Violin Plot for Hit Rate

1. χ^2 Test of Independence

I applied this on the the hit rate data. The columns of my data were the different races (Categorical Variable 1) and the rows of my data were different states (Categorical Variable 2).

Note that χ^2 tests are not meant for ratios, so I multiplied by a fixed number, 100, to get frequencies which is what χ^2 test requires. Note that it is not wrong to do this because our data is of the order of at least 10^5, so multiplying it by 100 is no big deal.

1.1 Contingency Table

```
                        CT RI CA AZ MO WI NC CO TX LA MD VT PA TN MA
white                   13 19 19 27 33 33 36 38 36 10  9 24 50 50 57
hispanic                52 47 41 47 52 25 20 29 29 50 39 42 58 20  9
black                   19 16  6  7 10  9 85 69 73 85 22 34 27 29 17
asian/pacific islander 12 21 20 32 44 48 51 13 19 19 27 33 33 36 38
```

1.2 Test Result

```
Pearson's Chi-squared test

data:  contingency_table
X-squared = 474.26, df = 42, p-value < 2.2e-16
```

1.3 Conclusion and Interpretation

$p < 0.05$ so we can say that the variables are not independent of each other. This gives us a green flag to do further analysis on our data

2. ANOVA

2.1 Check for Normality: Shapiro-Wilk Normality Test (N<50)

Subject Race	W	p-value
asian/pacific islander	0.91455	0.2116
hispanic	0.9308	0.3491
black	0.95296	0.6438
white	0.93305	0.3734

NORMALITY RETAINED
Since all of the groups satisfy the normality condition $(p > 0.05)$, we can apply parametric ANOVA test

2.2 Check for Homogeneity: Levene test

```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  3  0.0235 0.9951
      48
```

HOMOGENEITY RETAINED
Since homogeneity of variance is retained $(p > 0.05)$ and normality is satisfied as well,

we can apply ONE-WAY ANOVA

2.3 ONE-WAY ANOVA

```
             Df Sum Sq Mean Sq F value Pr(>F)
Subject_Race  3 0.0431 0.01435   0.358  0.783
Residuals    48 1.9221 0.04004
```

NULL HYPOTHESIS RETAINED

We don't have a statistically significant effect which means that the variance in means of various groups is same (similar).

3. Pairwise Permutation using BH Correction

Comparison	Stat	p.value	p.adjust
asian/pacific islander - hispanic	0.4937	0.6215	0.7458
asian/pacific islander - black	-0.02313	0.9815	0.9815
asian/pacific islander - white	-0.535	0.5926	0.7458
hispanic - black	-0.541	0.5885	0.7458
hispanic - white	-1.032	0.3022	0.7458
black - white	-0.5351	0.5926	0.7458

None of the p-values are below 0.05, implying that we don't have a statistically significant effect.

7.1.3 Conclusion

The analysis of hit rates reveals a correlation between geographic region and these rates, as the data shows variation across different states. However, in contrast to the findings related to search rates, the hit rate analysis does not indicate racial disparities within each state. In other words, while hit rates differ between states, they appear to be consistent across racial groups within a given state.

Based on the arguments presented, the hit rate data may provide a more appropriate basis for drawing conclusions about potential disparities in search outcomes. The lack of observed racial disparities in hit rates suggests that the previously identified differences in search rates across groups may not necessarily translate into differences in the likelihood of finding contraband during those searches.

Result of Hypothesis 1

In conclusion, the analysis of search data yielded contrasting findings when examining search rates versus hit rates across racial/ethnic groups. While disparities in search rates were observed, with certain groups experiencing higher rates of searches, the subsequent analysis of hit rates did not indicate racial or ethnic disparities in the likelihood of finding contraband during those searches. Specifically, the hit rate analysis showed that while hit rates varied across different geographic regions, they were relatively consistent across racial and ethnic groups within each state.

7.2 Hypothesis 2

7.2.1 Conclusion

The analysis revealed regional variations in outcomes such as arrests made, citations issued, and warnings given. However, we did not find evidence of racial disparities in the issuance of citations or warnings. Notably, individuals identified as `Asian/Pacific Islander` experienced lower rates of arrest compared to those who identified as `Black` or `Hispanic`.

One potential contributing factor could be immigration patterns. The data analyzed spans roughly 2008-2018, and research suggests first-generation immigrants tend to exhibit lower crime rates than subsequent generations and non-immigrants. This phenomenon may be influenced by the selective nature of immigration policies, the personal motivations and characteristics of those who choose to immigrate, and the fact that Europeans and Asians were among the earliest immigrant groups to the United States in the early 1800s (History.com, Asian American Timeline). Additionally, systemic bias and over-policing of certain racial groups could play a role in higher arrest rates for Black and Hispanic individuals.

The observed racial disparities in arrest rates are unlikely to be due to chance alone, as evidenced by the non-significant results of the permutation test. The data shows consistent variations across multiple years, making it implausible that these disparities occurred randomly. Instead, a more plausible explanation lies in the evolution of systematic biases and policing practices over time, coupled with shifts in societal viewpoints and attitudes.

Consequently, it is more reasonable to attribute these disparities to systemic factors that have evolved over time, such as changes in law enforcement policies, practices, and training, as well as broader societal attitudes and perceptions towards different racial and ethnic groups. As these systemic factors have gradually shifted, their impact on arrest rates may have manifested in the observed patterns across multiple years.

> **Result of Hypothesis 2**
>
> There are regional variations in outcomes such as arrests made, citations issued, and warnings given. However, we did not find evidence of racial disparities in the issuance of citations or warnings. Moreover, individuals identified as `Asian/Pacific Islander` experienced lower rates of arrest compared to those who identified as `Black` or `Hispanic`.

7.3 Hypothesis 3

7.3.1 Conclusion

The results of the four factors that my teammate looked at are as follows:

- Search rate: No strong evidence against the null hypothesis, but a close case (further analysis may be necessary).
- Warning rate: Null hypothesis holds true. No evidence against it.
- Citation rate: Null hypothesis holds true. No evidence against it.
- Arrest rate: Evidence against null hypothesis, thus it’s rejected.

Result of Hypothesis 3

Overall, **there is some bias observed, with regard to gender of the subject, in policing practices in stops**. This is alarming and should be looked further into.

Appendix

State Abbreviations

State	Full Form	State	Full Form	State	Full Form
AL	Alabama	LA	Louisiana	OH	Ohio
AK	Alaska	ME	Maine	OK	Oklahoma
AZ	Arizona	MD	Maryland	OR	Oregon
AR	Arkansas	MA	Massachusetts	PA	Pennsylvania
CA	California	MI	Michigan	RI	Rhode Island
CO	Colorado	MN	Minnesota	SC	South Carolina
CT	Connecticut	MS	Mississippi	SD	South Dakota
DE	Delaware	MO	Missouri	TN	Tennessee
FL	Florida	MT	Montana	TX	Texas
GA	Georgia	NE	Nebraska	UT	Utah
HI	Hawaii	NV	Nevada	VT	Vermont
ID	Idaho	NH	New Hampshire	VA	Virginia
IL	Illinois	NJ	New Jersey	WA	Washington
IN	Indiana	NM	New Mexico	WV	West Virginia
IA	Iowa	NY	New York	WI	Wisconsin
KS	Kansas	NC	North Carolina	WY	Wyoming
KY	Kentucky	ND	North Dakota		

References

[1] E. Pierson, C. Simoiu, J. Overgoor, S. Corbett-Davies, D. Jenson, A. Shoemaker, V. Ramachandran, P. Barghouty, C. Phillips, R. Shroff, and S. Goel. ***"A large-scale analysis of racial disparities in police stops across the United States"***. Nature Human Behaviour, Vol. 4, 2020.

[2] https://openpolicing.stanford.edu/data/

[3] Class Slides and resources