KEOHANE SQUAD

Racial Disparities in Traffic Stops/Citations

Keohane sQUAD: Chris Liang, Andrew Qin, Bob Qian, and Katie Nash

Introduction and Purpose

Background: the US disproportionately incarcerates black people as opposed to white peers, and systemic racism may be seen in traffic stop data.

Research question: What is the relationship between a subject's demographic attributes (race and sex) and the likelihood of being stopped by police in traffic in Durham or receiving a citation upon being stopped?

Hypothesis: Race and the likelihood of being stopped or receiving a citation by police in traffic in Durham County are related, with black people disproportionately more stopped relative to their population proportion.

Data

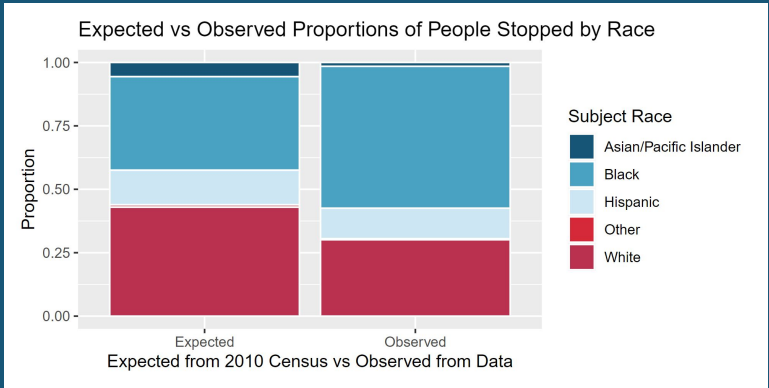
Dataset:

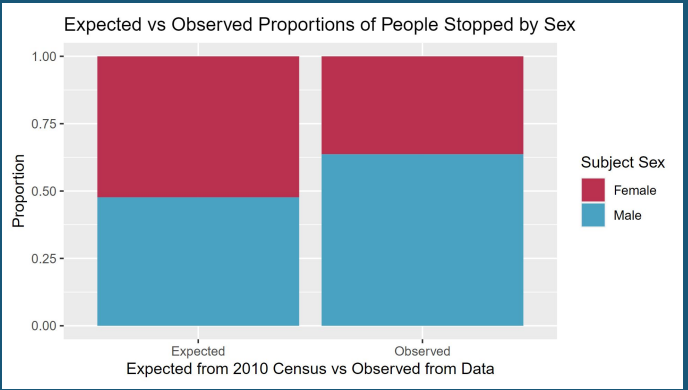
- Census of individual police stops in Durham from Stanford Open Policing Project with 323,147 observations—each an individual police stop recorded in Durham between December 2001 and December 2015
- We took stratified proportional samples of size 3231 or 3232 from the census data for 1st question, which is ~ 1% of the population size

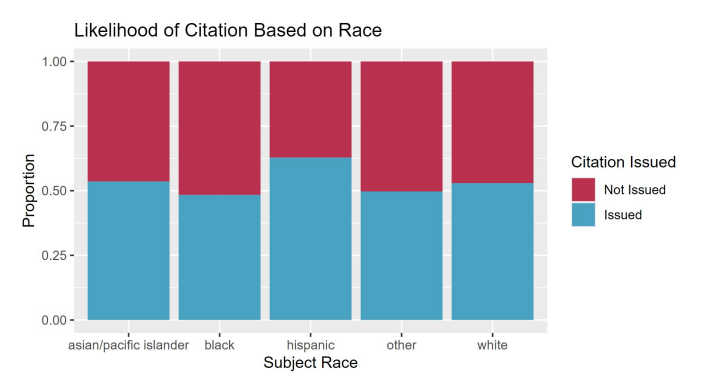
Variables:

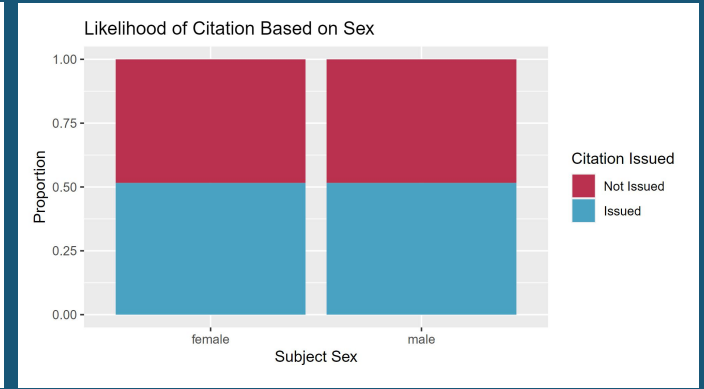
- There are 29 variables, with the most relevant ones being race, age, sex, and outcome of the stop (warning vs. citation, etc.)

EDA


Expected vs Observed Proportions of People Stopped by Race
Proportion
1.00
0.75
0.50
0.25
0.00
Expected
Observed
Expected from 2010 Census vs Observed from Data
Subject Race
Asian/Pacific Islander
Black
Hispanic
Other
White


Expected vs Observed Proportions of People Stopped by Sex
Proportion
1.00
0.75
0.50
0.25
0.00
Expected
Observed
Expected from 2010 Census vs Observed from Data
Subject Sex
Female
Male


Likelihood of Citation Based on Race
Proportion
1.00
0.75
0.50
0.25
0.00
asian/pacific islander
black
hispanic
other
white
Subject Race
Citation Issued
Not Issued
Issued


Likelihood of Citation Based on Sex
Proportion
1.00
0.75
0.50
0.25
0.00
female
male
Subject Sex
Citation Issued
Not Issued
Issued

Simulation

- After simulation, we receive 2 p-values of 0 for race and sex
- The data provides evidence to indicate that black people and males are disproportionately stopped

Chi-Squared Model

- Chi-Squared tests for independence outputted p-values of 0 and 0.98 for race and sex, respectively
- Data provides evidence to indicate race and likelihood of receiving citation are associated, but does not provide evidence of association for sex

Logistic Regression Model

- Variables included: Race, Sex, Age, and Race * Sex
- As age increases, log-odds of receiving citation decrease
- Black women's log-odds of receiving citation are higher than black men's
- HIspanic males have the highest log-odds of receiving a citation

term	estimate	std.error	statistic	p.value
(Intercept)	0.373	0.015	25.484	0.000
subject_raceasian/pacific islander	0.031	0.048	0.634	0.526
subject_raceblack	-0.127	0.013	-9.899	0.000
subject_racehispanic	0.357	0.025	14.458	0.000
subject_raceother	-0.253	0.105	-2.402	0.016
subject_age	-0.007	0.000	-26.459	0.000
subject_sexmale	0.032	0.013	2.443	0.015
subject_raceasian/pacific islander:subject_sexmale	-0.043	0.061	-0.715	0.474
subject_raceblack:subject_sexmale	-0.121	0.016	-7.364	0.000
subject_racehispanic:subject_sexmale	-0.004	0.029	-0.136	0.892
subject_raceother:subject_sexmale	0.151	0.125	1.206	0.228

Conclusion

- Black people are more likely to be stopped than white people in Durham, and are disproportionately more stopped than their population proportion
- Males are disproportionately more likely to be stopped
- Looking at sex within races, odds of citation for black females are much higher than for black males