

Racial Disparities in Traffic Stops/Citations

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

Update date

Introduction and Data

Our data is a census of individual police stops in Durham created by the Stanford Open Policing project. The Stanford Open Policing Project “[collects] and [standardizes] data on vehicle and pedestrian stops from law enforcement departments across the country”(https://openpolicing.stanford.edu/). We would like to see if that same kind of racial bias is evident in police stops in Durham. In doing so, we also wish to examine if other demographic characteristics (such as sex or age) influence traffic stops. Our general research question is the following: what is the relationship between a subject’s demographic attributes (sex, race, or age) and the likelihood of being stopped by police in traffic in Durham?

We hypothesize that race and the likelihood of being stopped by police in traffic in Durham are related, with black people representing disproportionately more of the people being stopped relative to their proportion within the population. We hypothesize that younger people (roughly 18-30) have a disproportionately higher chance of being stopped in traffic (not necessarily due to bias but other lurking variables, such as inexperienced driving). We also hypothesize that sex has no significant relationship with being stopped in traffic. To find the true population proportions of people by race, sex, and age in Durham, we will utilize the 2010 Durham census data (http://censusviewer.com/city/NC/Durham).

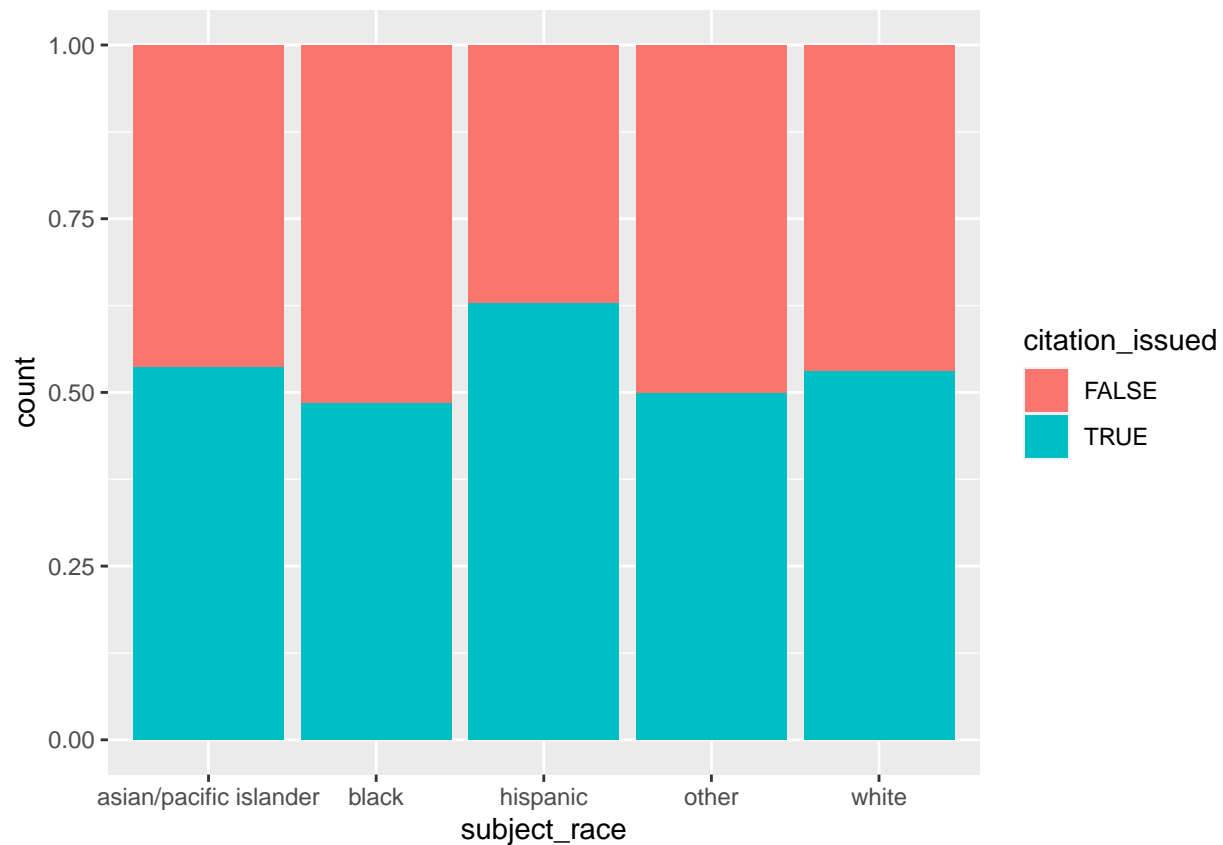
Additionally, we will examine whether race, sex, or age are related to the outcome of the traffic stop (whether a citation will be issued). We hypothesize that race and the likelihood of receiving a citation are related, with black people more likely to receive a citation upon being stopped. We additionally hypothesize that younger people have a higher chance of receiving a citation upon being stopped and that sex has no significant relationship with being stopped in traffic.

It has 29 variables and 323432 observations, and each observation in the data set is an individual police stop recorded in Durham during 2001 to 2015. A categorical variable in the data set is **subject_race**, which describes the race of the subject involved in the traffic stop. A discrete numerical variable in the data set is **subject_age**, which describes the age of the subject at the time of the traffic stop. A continuous numerical variable in the data set is **time**, which describes the hour, minute, and second that the stop was recorded. Other variables in the data set include **outcome**, which is what resulted from the stop (a warning or a citation, for example); **reason_for_stop**, which describes what the violation leading to the stop was; and **search_conducted**, whether a search of the subject was conducted during the stop.

Methodology

race_{citation} age_{citation} sex_{citation}

segmented bar graph w/ probability of citation based on race.



Results

```
## # A tibble: 1 x 3
##   statistic chisq_df p_value
##   <dbl>     <int>   <dbl>
## 1    2789.         4       0
```

Discussion

The chi squared test tells us that there is a significant relationship between subject race and traffic citations; however, the chi squared test does not tell us the exact nature of the relationship between subject race and traffic citations. Once we created a barplot of our data we discovered that hispanics are the most likely to receive a traffic citation after being pulled over. TALK ABOUT LIMITATIONS AND WHAT WE COULD HAVE DONE BETTER