

## Terry Stops in Seattle, WA

### Problem Statement:

A Terry Stop in the US allows the police to briefly detain a person based on reasonable suspicion of involvement in criminal activity. Under *Terry v. Ohio*, 392 U.S. 1 (1968), such contacts authorized under law and policy for purposes of investigating, based on an officer's reasonable suspicion, whether the individual is engaging, has engaged, or is about to engage in criminal activity. During the course of a Terry stop, an officer may develop probable cause to effect an arrest, but probable cause is not required to make the initial stop, nor does a stop that is based on probable cause to arrest fall within the category of a Terry stop.

There is a difference between one police officer stopping one individual, which is a tactical definition, and systematic promotion of this tactic on either the departmental or municipal level, which can damage police–community trust and lead to charges of racial profiling.

The goal is to investigate the Terry Stops in Seattle by race of the subject and officer, and gain insight if the percentage of stops match the demographic of the city. Also, to help identify if certain officers have a higher prevalence of stops of a certain race.

The data was obtained from the open data provided by the City of Seattle at:

<https://data.seattle.gov/Public-Safety/Terry-Stops/28ny-9ts8>

## Data Wrangling

After downloading the dataset, I took a look at the data and performed a few data cleaning / wrangling steps to help clean up the data.

I took a subset of the data, selecting 18 of the 23 columns that would be necessary for the analysis I was going to be conducting. I renamed the columns to remove spaces and capital letters to create consistency and ease during analysis.

Fifteen columns are object types and 3 are numeric. I converted the necessary columns to category and datetime types - these types better fit the data and will make it easier for analysis later on.

I made the following cleaning changes during exploration of the data:

### Missing Values:

- Replaced all '-' values with NaNs
- Observed the NaNs (1,015 rows) in the 'subject\_age' feature - these were left in because depending on the analysis, it may not affect the outcome

- There were 2 missing values in the 'officer\_age' feature, these were replaced with the mean age of the feature
- There were 'Unknown' and 'Not Specified' values for 'officer\_race' feature. The 'Unknown' values were changed to 'Not Specified' to keep consistency in the data. The 'Not specified' value was kept because I am assuming that those that were labeled 'not specified' were filled out by someone other than the officer performing the terry stop. Including these values will not affect the analysis.
- The rows with subject\_race NaN values were removed because this is the primary feature I will be looking into, also, the observations with a NaN in this column also corresponded to multiple features with NaN values (up to 10) mostly missing values for the subject. Leaving this data in will not contribute to the analysis.

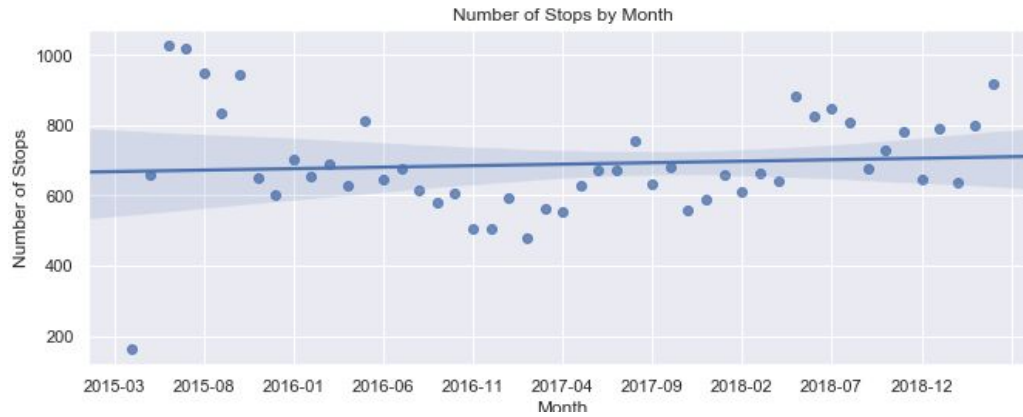
### **Cleaning Steps & Outliers:**

- The feature 'Officer YOB' was changed to 'age'. For readability a new feature was created with by subtracting this value with the date (year) of the observation to get an estimation of the age of the officer.
- Upon inspection, I found that there were a few officers that were over 100 years old. These observations had nonsensical data in other columns (ie. a negative officer\_id value, 'N' for officer\_gender etc.) These values were removed from the dataset.
- I inspected the remaining categories for any unique values
- Changed the 'frisk' & 'arrest' flag features from string categories to 1s and 0s
- The columns were rearranged and dataframe reindexed and saved in it's new cleaned state.

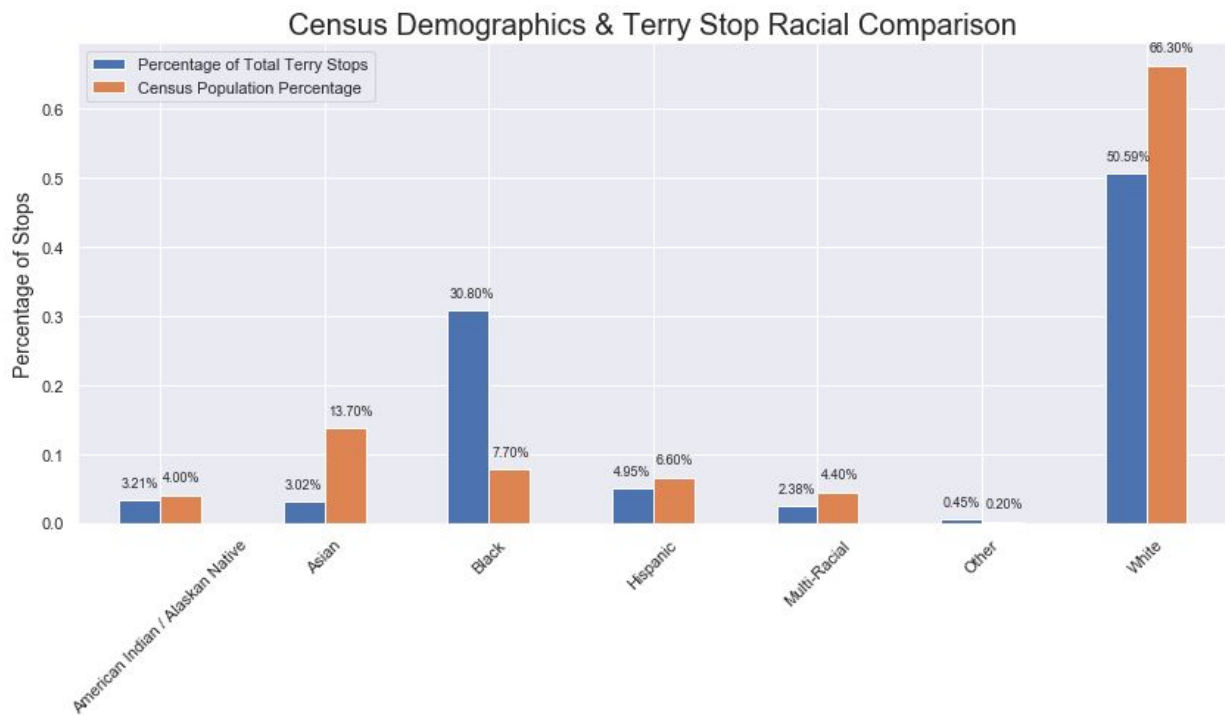
1. Initial findings from exploratory analysis (get this from your data story and inferential statistics reports)
  1. Summary of findings
  2. Visuals and statistics to support findings

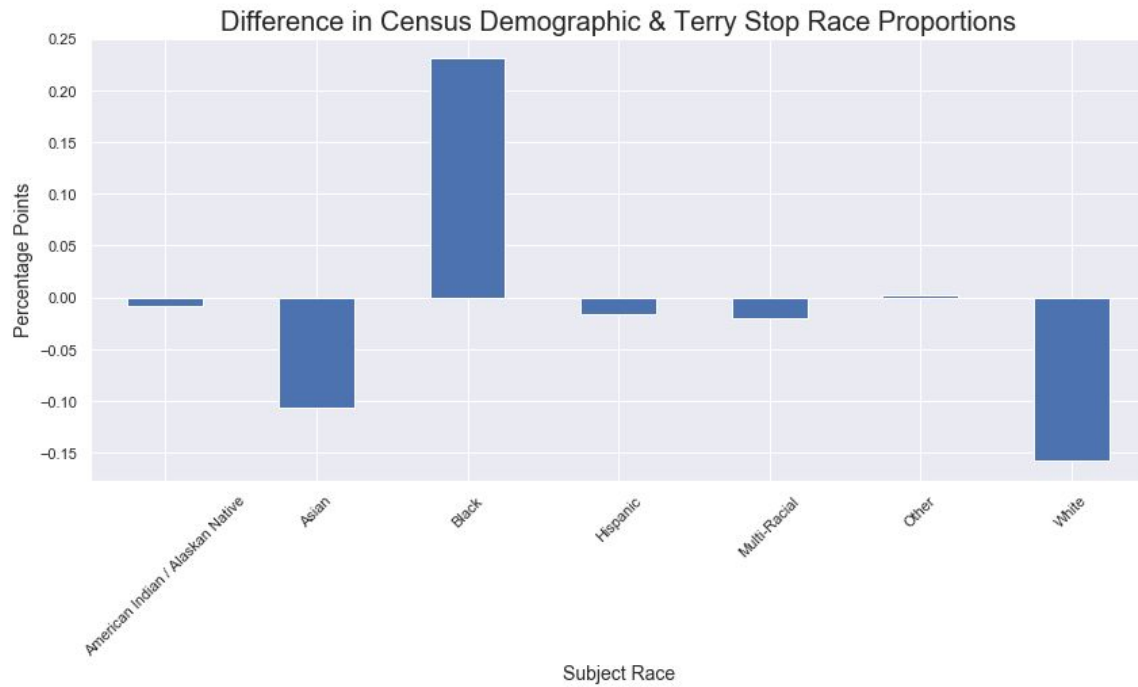
### **Initial Findings**

There was a large initial increase when the stops tracking first started, followed by steady decline over the next six months. Overall the number of stops have steadied, with a very slight increase over time.



White subjects make up the majority of stops in the data, however when looking at the proportion each race makes to the total, the proportion of Black race subjects is significantly larger than the demographic make up of black residents (demographic information taken from the 2010 census).

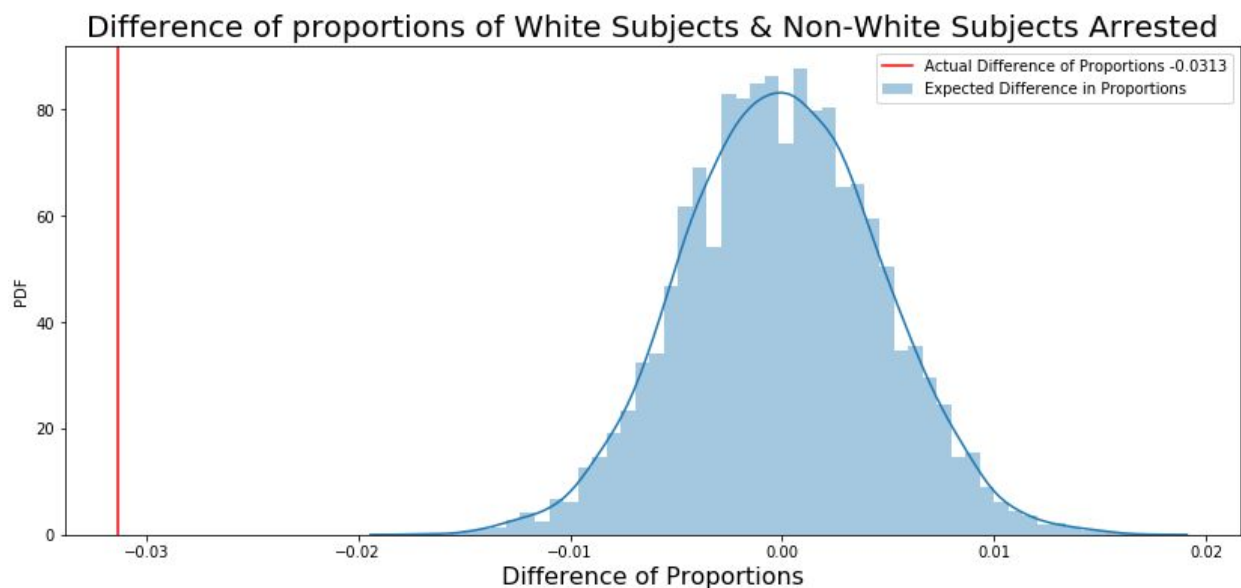
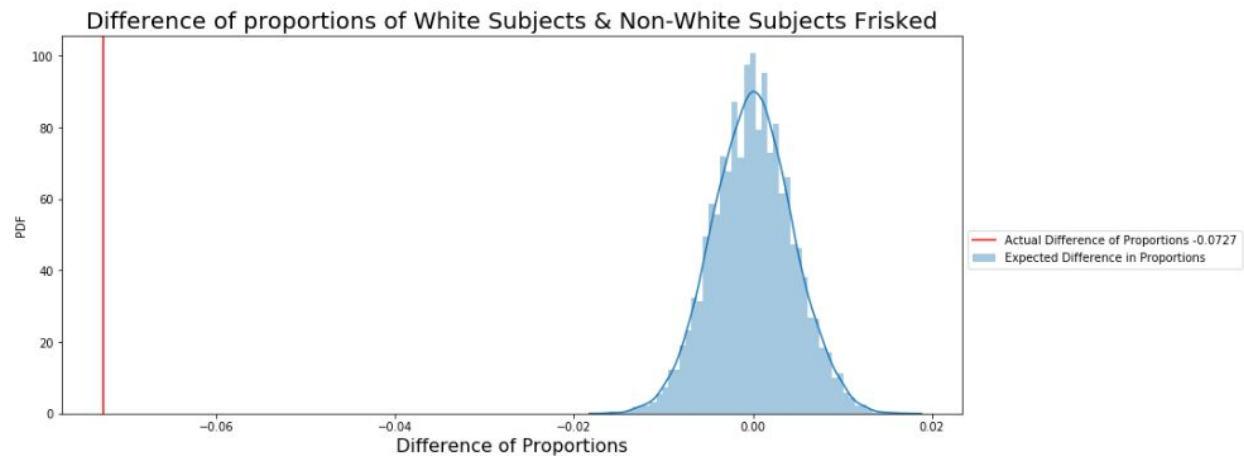




From an age perspective, the data suggests that Native Alaskans 36 years and older, and Black subjects 17 and younger are stopped more frequently than Non-Native Alaskans & Non-Black subjects respectively.



Also, White subjects are less likely to be arrested (-3.1%) or frisked (-7.3%) compared to Non-White subjects.



Further analysis is needed to identify some of these differences. For instance, what are the predominant neighborhoods where arrests occur, and what are the demographics of those neighborhoods? How are these stops initiated? Were these called in, or were they initiated on-the-spot by an officer? For those that were initiated by an officer, what portion of those did not lead to an arrest (no probable cause for arrest)? This could identify racial bias in Terry Stops.