A little look at the SPD

By Mitali Yadav, Thomas Hsiung, Yuwei Quan

Why this Data Set

- Even though was a limited dataset, it has a lot to offer when text analytics API was applied on Summary text
- In light of current events, there was a lot of tension and animosity between police and the public, we want to know if the police records echo this sentiment.
- Figure out if there is a pattern of police behaviour

Discovery from EDA

Data Cleaning

Interesting findings

Data Cleaning

- **Throw out** incidences that are not happening in Seattle (156 ➔ 151 rows)
- Found 12 **missing values** in policy injury columns.
  - Decided to drop the column instead of rows to leave as much data as possible
- Created **dummy variables** for categorical data.
  - `'Subject Weapon':{'Yes':1, 'No':0}` `'Fatal':{'Yes':1, 'No':0}`
- **Combined** the category in disposition from

Within Policy		Missing	0
Justified		Justified	1
Out of Policy	Into		Denoted by
Missing		Not justified	-1
Not Justified			

Findings

- Ratio of armed to unarmed subjects was **24 : 127**
- Ratio of fatal to non-fatal incidents reported was **79 : 72**
- Ratio of Male to Female officers was **133 : 18**
- The disposition column has 13 rows missing. The races of the officers with missing Dispositions were:

  White | Black | Hispanic/Latino

  11 | 1 | 1

Officer Race	Subject Race	Fatal	
AI/AN	Black or African American	0	5
	White	0	2
American Indian/Alaska Native	White	1	1
Asian	White	1	1
Asian/Pacific Islander	Black or African American	0	2
	White	1	2
Black	Black or African American	0	1
		1	1
	White	0	1
		1	1
Black or African American	Asian	1	1
	Black or African American	1	1
	Not Specified	1	1
	White	1	2
Hispanic or Latino	Black or African American	1	1
	White	1	2
Hispanic/Latino	Asian	0	1
	White	0	2
		1	1
Multi-Racial	Asian	1	1
	White	0	1
Nat Hawaiian/Oth Pac Islander	White	1	1
Two or More Races	Not Specified	1	1
White	Asian	0	8
		1	3
	Black or African American	0	26
		1	15
	Hispanic	0	1
	Nat Hawaiian/Oth Pac Islander	1	1

MultiIndex table of Officer, Subject and whether the shooting was fatal

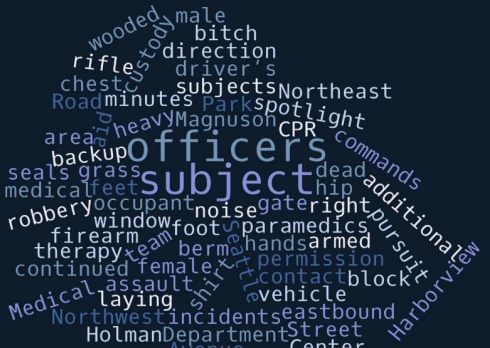
Outlier

When we are taking a look at the tone, language and length of the summaries, especially those without a weapon :

...an on duty Patrol Officer responded to assist in locating a possibly suicidal person...The subject was located in steep terrain and bushes after having jumped from the bridge. Officers approached the subject, and while attempting to take the subject into custody, the subject made attempts to grab both officers' firearms. An officer fired two rounds, striking the subject. The subject was not armed by attempted to take or control the firearms of both officers.

... two on duty Seattle Police Officers were in the area of 42nd Ave Southwest and Fauntleroy Way Southwest...The officers were approached by an employee of the Washington Federal Savings Bank (6428 California Avenue Southwest), who stated they had just been the subject of a robbery. The employee pointed to a subject walking eastbound from their location, wearing a black hooded sweatshirt and black pants, and identified him as the person who had committed the robbery. One of the officers approached the subject from behind, identified himself as 'Police' and ordered the subject, several times, to 'get on the ground.' The subject had his hood up and his hands in the pockets of his sweatshirt and did not respond to commands from the officer. Shortly after the officer made initial contact with the subject, a vehicle, later identified as the getaway car, pulled up and the subject turned to face the officer. The officer fired three rounds, striking the subject in the left leg.

Asian



Hispanic or Latino

Black

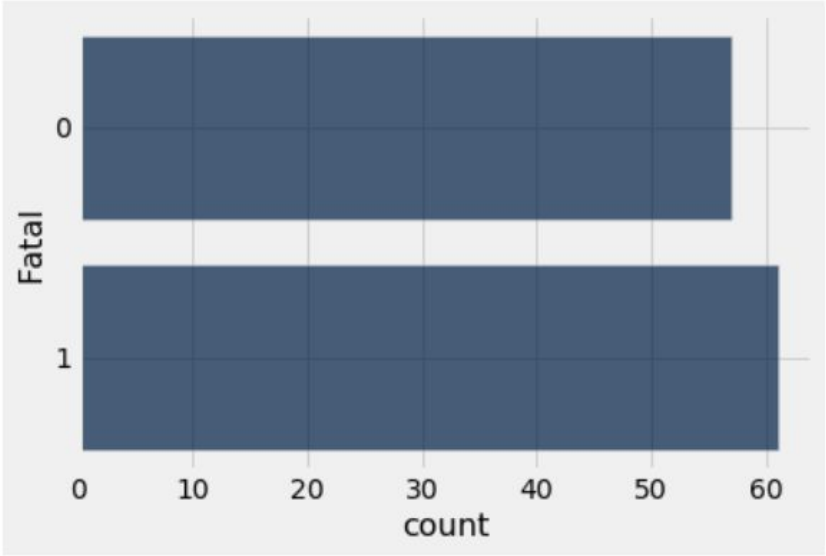
White

Research

White officers comprise the largest group within the Seattle police department, while non-white officers are much fewer in number, so we initially explored the data set by comparing white officers to non-white officers.

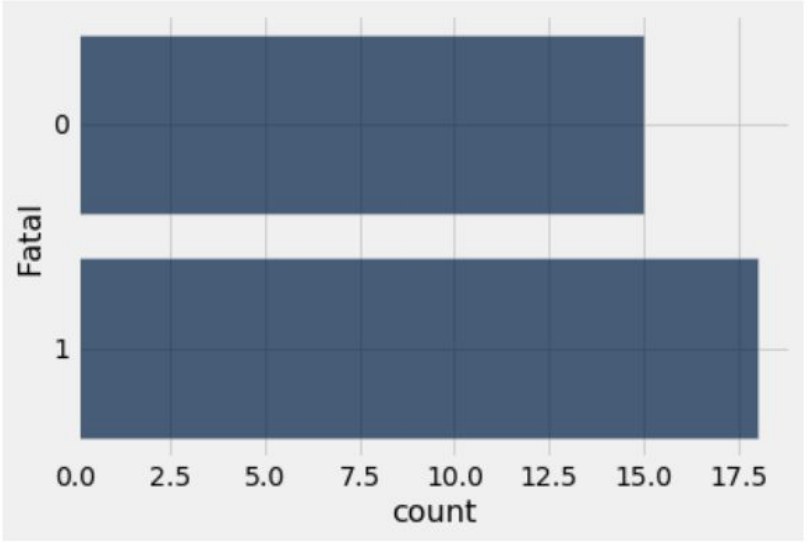
The results were fairly similar between both groups, and not particularly illuminating.

Fatality rates of white officers



Fatality rate: ~52%

Fatality rates of non-white officers



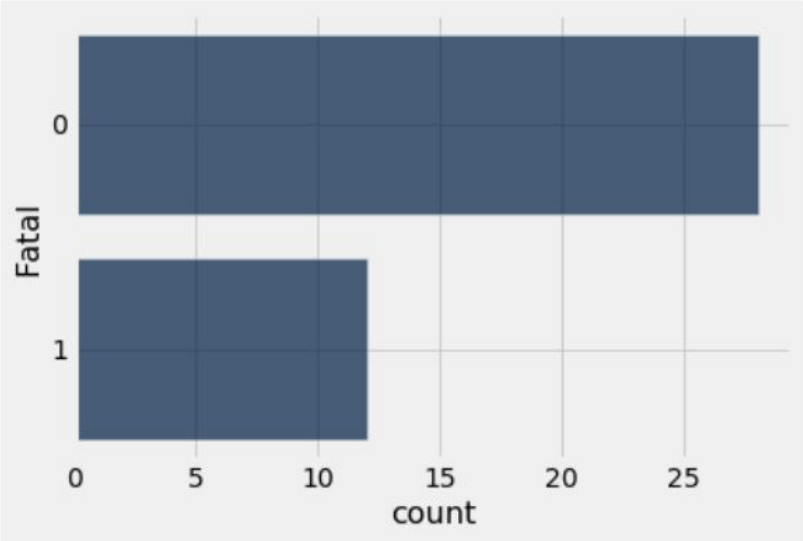
Fatality rate: ~55%

Research

Then we noticed many instances of separate police reports being filed for the same incident (noted by the same Longitude + Latitude, as well as sharing Key Phrases).

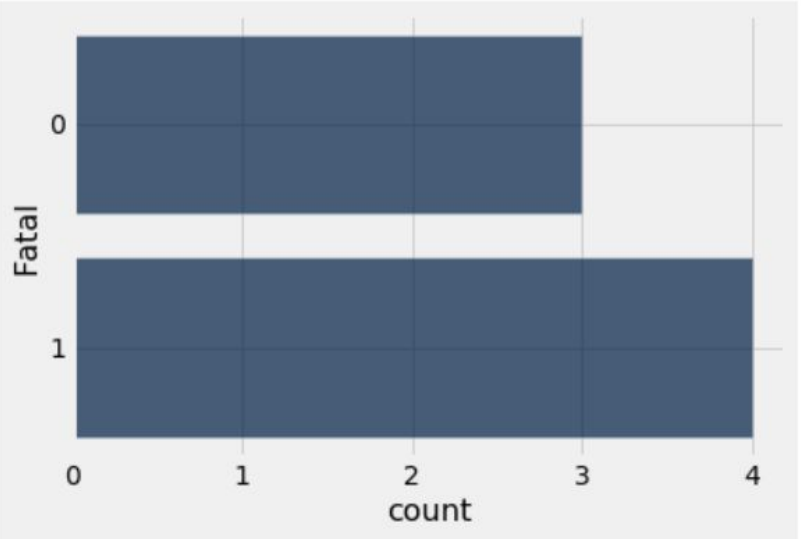
By separating the data set into individual police responders vs. group police responses, we found more interesting implications.

Fatality % of individual white officers



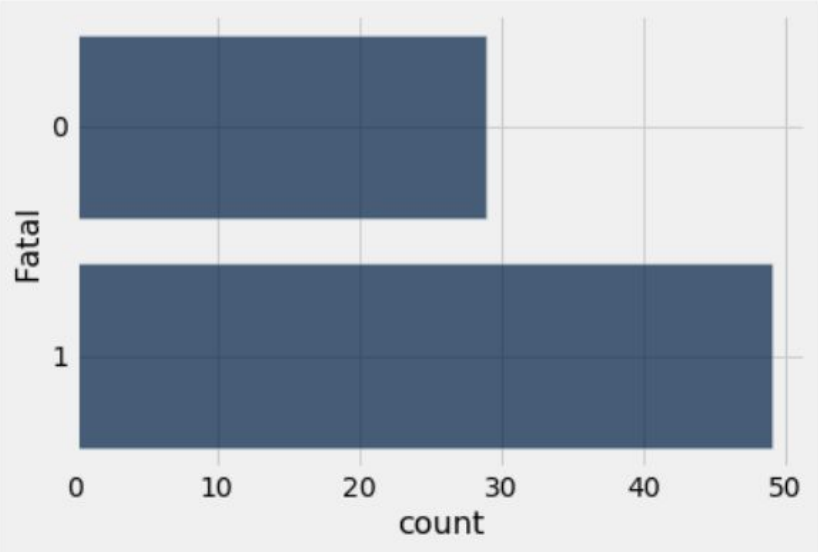
Fatality rate: ~30%

Fatality % of individual non-white officers



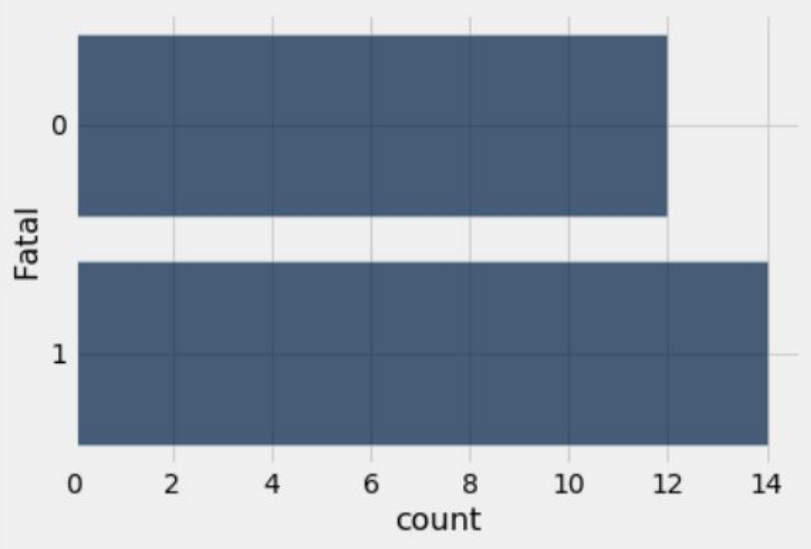
Fatality rate: ~57%

Fatality % of white officers in groups



Fatality rate: ~63%

Fatality % of non-white officers in groups



Fatality rate: ~54%

Hypothesis:

There is a higher rate of fatalities in shooting incidents involving white officers if they are in a group.

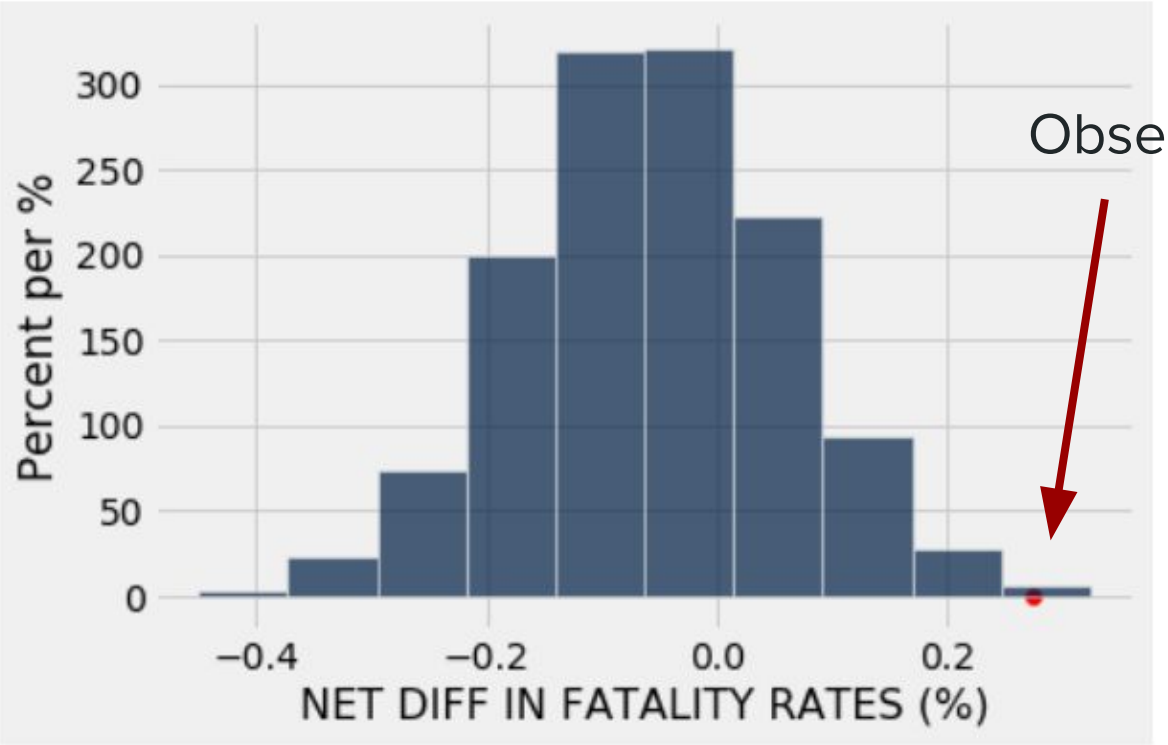
Hypothesis Testing

To test the hypothesis, we will assume the null is true and use A/B Testing with a 5% p-value cutoff.

Splitting the sample:

- Group A contains those NOT in a group when responding
- Group B contains those IN groups responding to incidents
- If the null is true, then the difference between the fatality rates of B average and A average should be close to zero.

Results from Bootstrapping 3000 resamples:



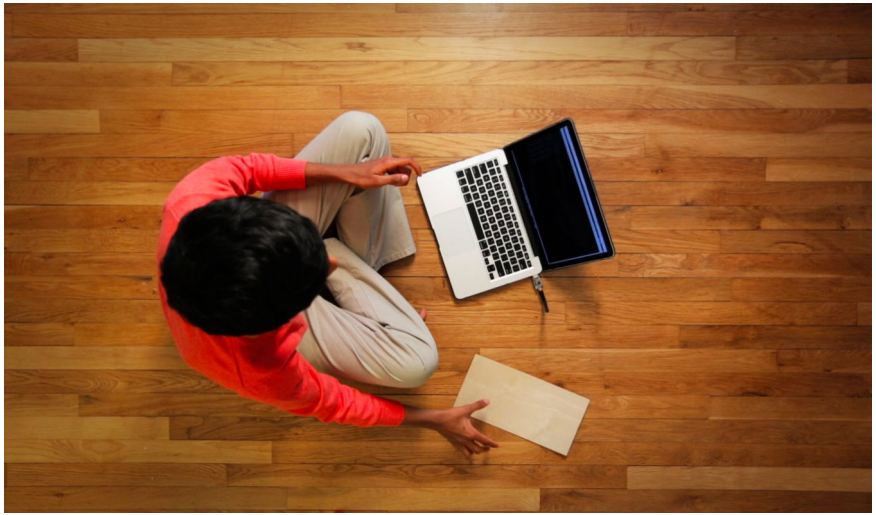
Observed difference: **0.27**

P-value: **0.0026**

The observed difference is around .27, and nearly all resampled values fall below that number.

With a p-value of 0.003 and a p-value cutoff of 5%, the data is inconsistent with the null hypothesis.

This means that the finding is statistically significant, and that it is likely that police shooting incidents involving white officers are much more likely to end in a fatality if the officers are in a group.



What will I do next?

Clean up the repeated high frequency words in the key phrase such as Policy, Department etc.. too see a more interesting word cloud.