Exploring Racism in the U.S. Criminal Justice System with DCAT

August 3

2015

Over the past years there have been an escalating number of unjustifiable deaths to American citizens caused by the police and federal law enforcement agencies within this country.   Allegations of the use of excessive force by U.S. police departments are generating today's headlines. In Staten Island, NY, July 2014, the death of Eric Garner died because of the apparent use of a chokehold by an officer sparked outrage. A month later Michael Brown, in Ferguson MO, got shot by police officer Darren Wilson and the grand jury's decision not to indict him triggered further unrest. In November, Tamir Rice, who was only 12 years old and playing with a toy pistol, got shot by police in Cleveland, Ohio. On April 4, 2015, Walter L. Scott was shot by a police officer after a routine traffic stop in North Charleston, S.C. The same month, Freddie Gray died while in police custody in Baltimore, setting off widespread unrest. This paper presents a closely examined data analysis exploring the statistical data on topics dealing with Racial Profiling in the United States. I have used Python along with DCAT, FuXi and rdflib to prove that the U.S. criminal justice system is a race-based institution where African-Americans are directly targeted and punished in a much more aggressive way than any other race in America.

Semantic Web Search Engine

Contents

Term Project: Exploring Racism in U.S. Criminal Justice System using DCAT

Objective

This paper is based on using the DCAT- Data Catalog Vocabulary in order to seek specific information regarding racial profiling and police brutality. I have used semantic web programming combined with data analytics to conduct this project in proving certain effects of Racial Profiling in America based on various perspectives and sources.

Significance

Over the past years there have been an escalating number of unjustifiable deaths to American citizens caused by the police and federal law enforcement agencies within this country. Allegations of the use of excessive force by U.S. police departments are generating today's headlines. In Staten Island, NY, July 2014, the death of Eric Garner died because of the apparent use of a chokehold by an officer sparked outrage. A month later Michael Brown, in Ferguson MO, got shot by police officer Darren Wilson and the grand jury's decision not to indict him triggered further unrest. In November, Tamir Rice, who was only 12 years old and playing with a toy pistol, got shot by police in Cleveland, Ohio. On April 4, 2015, Walter L. Scott was shot by a police officer after a routine traffic stop in North Charleston, S.C. The same month, Freddie Gray died while in police custody in Baltimore, setting off widespread unrest.

A total of 682 people have already been killed this year throughout the US according to "The Counted" [1]. More disturbingly, half of the people that were killed by these police officers and federal agents were minorities, in which make up no greater than 38% of the US's total population. Being a young African American man residing in Florida, which has one of the highest numbers of people killed by the police in the US as to date, I am affected greatly by these statistics.

The purpose of this project is to be able explore and examine statistical data analysis on topics dealing with Racial Profiling in America. I will use multiple perspectives and sources in order to perform data analytics to answer questions such as: Are minorities in America more likely to be killed by cops than whites? Are minorities in America more likely to be injured by cops than whites? Are minorities in America more likely to die during medical operations than whites? The statistical data obtained from the provided datasets created by my search engine will prove that the U.S. criminal justice system is a race-based institution where African-Americans are directly targeted and punished in a much more aggressive way than whites.

Tools and Languages;

The tools used during this analysis were Python along with DCAT, FuXi and rdflib combined to collect multiple sets of data into one dataset. Other tools used to produce the data within my experiment were Google's "OpenRefine" and the Socrata database.

DCAT

DCAT is an RDF vocabulary designed to facilitate interoperability between data catalogs published on the Web. DCAT will be used to describe my datasets in data catalogs. This will increase the discoverability and enable applications easily to consume metadata from multiple catalogs. The namespace for DCAT is http://www.w3.org/ns/dcat#. DCAT itself defines a minimal set of classes and properties of its own including dcat, dct, dctype, foaf, rdf, rdfs, skos, vcard, and xsd. DCAT is well-suited for representing government data catalogs such as Data.gov, which I will be using to provide some of the data for my study.

Datasets

The datasets included in this study contain statistical information regarding a racial demographic breakdown of the U.S. population, the number of police shootings by race, the amount of justified and unjustified shootings by police, the percent of people killed by police, the total number of police shootings ordered by race, and the number fatal injuries by legal intervention.

- **A racial demographic breakdown of the U.S. population**
  http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk [2]

- **Justifiable Homicides by Police**
  https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2013/crime-in-the-u.s.-2013/offenses-known-to-law-enforcement/expanded-homicide/expanded_homicide_data_table_14_justifiable_homicide_by_weapon_law_enforcement_2009-2013.xls [3]

- **Arrest-Related Deaths**
- **The number of police shootings by race**
  http://www.bjs.gov/index.cfm?ty=pbdetail&iid=5259 [4]

- **The number fatal injuries by legal intervention**
  https://wisqars.cdc.gov:8443/cdcMapFramework/mapModuleInterface.jsp [5]

Inputs and Outputs

The main sources I will be collecting my data from will be Data.gov, the United States Census Bureau, and FBI.gov. This data will be going into the created search engine involves combing multiple sources in order to acquire a broader view of various datasets. In return, I will be able to perform data analytics by cross examining and comparing the retrieved stats.

The first input dataset being used is **a racial demographic breakdown of the U.S. population,** provided by **Data.gov**. This dataset will provide the information on the Annual Estimates of the Resident Population by Sex, Race, and Hispanic Origin for the United States, States, and Counties.

The second input dataset will be the **Justifiable Homicides by Police** data provided by **FBI.gov**. This data contains the number of justifiable deaths recorded by police and which weapon was used.

The third input dataset will be the total number of **Arrest Related Deaths** by police. This dataset, provided by the "Bureau of Justice Statistics" (BJS.gov), will have information on the number of police shootings by race, gender, demographics, etc.

The fourth input dataset being used is **The Number of Fatal Injuries by Legal Intervention,** provided by **CDC.gov,** will show information on number of people that died due to police brutality and fatal injuries.

The output of this data provided a beneficial correlation and was able to answer the following questions regarding racial profiling and police.

> **Questions**
> 1. Are minorities in America more likely to be killed by cops than whites?
> 2. Are minorities in America more likely to be injured by cops than whites?
> 3. Are minorities in America that are unarmed more likely to be killed during conflicts with the United States Law enforcement?

Challenges

Some of the challenges I ran across during this analysis was the lack of police shooting data actually entered into the government database. This is due to the reason that law enforcement agencies lack "sufficient incentives" to report officer-involved shootings. Local agencies also don't properly report the injuries and deaths of police offenders at times either. The Federal Bureau of Investigation captures data on justifiable homicides by law enforcement officers, but reporting is **voluntary** and limited to instances in which a civilian is killed while committing a felony.

Experiment

I compared datasets using RDF and CSV files to perform and conduct data analytics on if minorities in the US were really being killed more often than whites during confrontations with the law enforcement.

Total US Population

The first dataset I used was the US population Estimates provided by Data.gov to find out the total population of each race in the US. I discovered that there were 1014900744 Whites, 182689000 Blacks, 8100100 Asians, 26130392 American Indians, and 5876704 Native Hawaiians.

RACE	TOTAL POPULATION
White	1014900744
Black	182689000
Asian	81001000
American Indian	26130392
Native Hawaiian	5876704

Figure 1 Total Population in the US

Total Number of Deaths for Each Race Caused by the US Law Enforcement

The first dataset I used was the number of killing by the US Law Enforcement since January 1st 2015 provided by "The Gaurdian" in order to find out the total number of deaths for each race in the US so far. I discovered that 334 Whites, 176 Blacks, 100 Hispanics, 13 Asians and 8 American Indians were killed. 46 races were unknown.

RACE	TOTAL DEATHS
White	334
Black	176
Hispanic/Latino	100
Unknown	46
Asian/Pacific Islander	13
Native American	8
Arab-American	3

Figure 2 Total Number of Deaths for Each Race since Jan 1st 2015

Finding the Total Death Rate for Each Race

The last dataset I used for my analysis was a combination of the first two datasets using a search engine I implemented in Python. I took the total deaths of each race divided by the total number population number for each race to retain the correct rate. My results were as follows:

RACE	TOTAL DEATH RATE BY POLICE
White	3.20
Black	9.60
Asian	1.23
American Indian	4.97
Native Hawaiian	1.70

Figure 3 Total Death Rate for Each Race

Conclusion

After performing my data analysis, I concluded that my hypothesis was accurate. Although the data shows that more whites have been killed by Americas law enforcement to date, Blacks in fact are being killed at a higher rate than any other minority.

RACE	TOTAL DEATHS
White	334
Black	176
Hispanic/Latino	100
Unknown	46
Asian/Pacific Islander	13
Native American	8
Arab-American	3

COMPARING GRAPHS FOR TOTAL RATES

RACE	TOTAL POPULATION
White	1014900744
Black	182689000
Asian	81001000
American Indian	26130392
Native Hawaiian	5876704

--------------------OVERALL TOTAL DEATH RATE by PEOPLE KILLED BY THE POLICE------------------------

RACE	TOTAL DEATH RATE BY POLICE
White	3.20
Black	9.60
Asian	1.23
American Indian	4.97
Native Hawaiian	1.70

Figure 4 Screenshot of the output from the implemented search engine

Bibliography

[1] O. L. L. U. i. t. Jon Swaine, "The Counted," The Gaurdian, 3 August 2015. [Online]. Available: http://www.theguardian.com/us-news/ng-interactive/2015/jun/01/the-counted-police-killings-us-database. [Accessed 3 August 2015].

[2] F. Finder. [Online]. Available: http://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk.

[3] FBI.Gov. [Online]. Available: https://www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2013/crime-in-the-u.s.-2013/offenses-known-to-law-enforcement/expanded-homicide/expanded_homicide_data_table_14_justifiable_homicide_by_weapon_law_enforcement_2009-2013.xls.

[4] BJS.gov. [Online]. Available: http://www.bjs.gov/index.cfm?ty=pbdetail&iid=5259.

[5] WISQARS.cdc.gov. [Online]. Available: https://wisqars.cdc.gov:8443/cdcMapFramework/mapModuleInterface.jsp.