**Discrimination in Police shootings in the United States**

Yingshi Lew, Himanshu Bhushan

[yingshi.lew@go.winona.edu](mailto:yingshi.lew@go.winona.edu) | [himanshu.bhushan@go.winona.edu](mailto:himanshu.bhushan@go.winona.edu)

**Contents**

[**1. Introduction** 2](#_Toc89285472)

[**1.1 Pur****po****se** 2](#_Toc89285473)

[**1.2 Project Scope** 2](#_Toc89285474)

[**2. Timeline** 2](#_Toc89285475)

[**3. Da****ta** 3](#_Toc89285476)

[**3.1 Data Acquisition** 3](#_Toc89285477)

[**3.2 Initial Database relation** 3](#_Toc89285478)

[**3.3 Constraints** 4](#_Toc89285479)

[**4. Problems (Exploratory Data Analysis)** 4](#_Toc89285480)

[**5. Database relation** 5](#_Toc89285481)

[**5.1 ERD** 5](#_Toc89285482)

[**5.2 Current Relation Schema** 5](#_Toc89285483)

[**6. Use cases** 6](#_Toc89285484)

[**7. Conclusion** 6](#_Toc89285485)

## **1. Introduction**

### **1.1 Purpose**

The purpose of this project is to design and curate a relational database to identify trends of discrimination in police shootings across the United States. The original database is constantly updated every time a police shooting is recorded. This database has been active from 2013-Present.

### **1.2 Project Scope**

This project focuses on utilizing the victim’s information to identify trends in encounters with the police, the alleged reasons for such encounters and shootings, as well as the police agencies/departments that are responsible for the victim’s death. These data were not collected by us and by no means should we be credited for the curation of the data. We are simply taking advantage of this dataset to aid our database design and creation using MySQL.

## **2. Timeline**

**Milestone I**

Design – 10/17/2021 – 10/22/2021 (Dennis & Himanshu)

**Milestone II**

Database Design/Creation – 10/23/2021 – 10/30/2021 (Dennis)

Data Insertion – 10/23/2021 – 10/30/2021 (Himanshu)

Data Analysis – 10/31/2021 – 11/06/2021 (Dennis & Himanshu)

Visualization – 11/08/2021 – 11/12/2021 (Dennis)

**Final Report**

Report Generation – 11/13/2021 (Dennis & Himanshu)

Presentation Slides – 11/15/2021 – 11/19/2021 (Dennis & Himanshu)

## **3. Data**

### **3.1 Data Acquisition**

This database stems from a real-world database that can be found [here](https://mappingpoliceviolence.org/). An intuitive dashboard can be found on the website to learn more about the climate of police shootings in the United States. The aforementioned website was just one of several inspirations for this project. The Washington Post also has another database that portrays a database with similar effectiveness in solving the problems at hand.

Other Links:

https://mappingpoliceviolence.org/

https://www.washingtonpost.com/graphics/investigations/police-shootings-database/

### **3.2 Initial Database relation**

victim(ID, date, name, age, gender, race, cause\_of\_death, ORI\_agency\_identifier)

location(ID, street, city, state\_abbrev, zip, county, geography)

alleged(ID, criminal\_charges, mental\_illness, armed\_unarmed, alleged\_weapon, alleged\_threat, fleeing)

department(ORI\_agency\_identifier, agency\_responsible)

population(state, state\_abbrev, total, black, Hispanic, native\_american, Asian, pacific\_islander, white, other)

officer(ORI\_agency\_identifier, body\_cam, offduty\_killing, encounter\_type, encounter\_reason, call\_for\_service)

### **3.3 Constraints**

*VarChar* – name, age, gender, race, cause\_of\_death, ORI\_agency\_identifier, street, city, state\_abbrev, zip, county, geography, criminal\_charges, mental\_illness, armed\_unarmed, alleged\_weapon, alleged\_threat, fleeing, agency\_responsible, state, body\_cam, offduty\_killing, encounter\_type, encounter\_reason, call\_for\_service

*Date* – date

*Int* – total, black, Hispanic, native\_american, Asian, pacific\_islander, white, other

## **4. Problems (Exploratory Data Analysis)**

* What were the number of shootings like?
* Where did the shootings take place?
* What are the demographics of the victims (age/race/gender)?
* What are the trends of these shooting occurrences with regards to the alleged charges imposed on the victims?

## **5. Database relation**

### **5.1 ERD**

Diagram

Description automatically generated

### **5.2 Current Relation Schema**

victim(ID, date, name, age, gender, race, cause\_of\_death)

location(ID, street, city, state\_abbrev, zip, county, geography)

alleged(ID, criminal\_charges, mental\_illness, armed\_unarmed, alleged\_weapon, alleged\_threat, fleeing)

police\_dept(ORI\_agency\_identifier, agency\_responsible, ID)

officer(ORI\_agency\_identifier, body\_cam, offduty\_killing, encounter\_type, encounter\_reason, call\_for\_service)

## **6. Use cases**

This database could provide several different angles to answer the necessary questions. The stated use cases are just a few examples. For now, this project aims to identify:

* the number of shootings during COVID-19
* killings by race
* killings by location
* if police officer(s) wore a body camera during their altercation with the victim(s)
* the types and reasons for these encounters
* if the victim had any charges imposed on them prior to/after the shooting.

## **7. Conclusion**

The victim’s information is used to identify trends in encounters with the police, the alleged reasons for such encounters and shootings, as well as the police agencies/departments that are responsible for the victim’s death. There are many directions to go into while performing exploratory data analysis. In the future, this project could be improved and stretched out to become like the websites mentioned above. Due to the complexity of database design itself, there will be much more to learn and to add to this initial phase.