Factors Related To Fatal Police Shootings

Team 2A
Cheng Chen | Jaya Nagesh | Mingwei Li | Selma Sentissi El Idrissi | Shamika Kalwe

Project Proposal

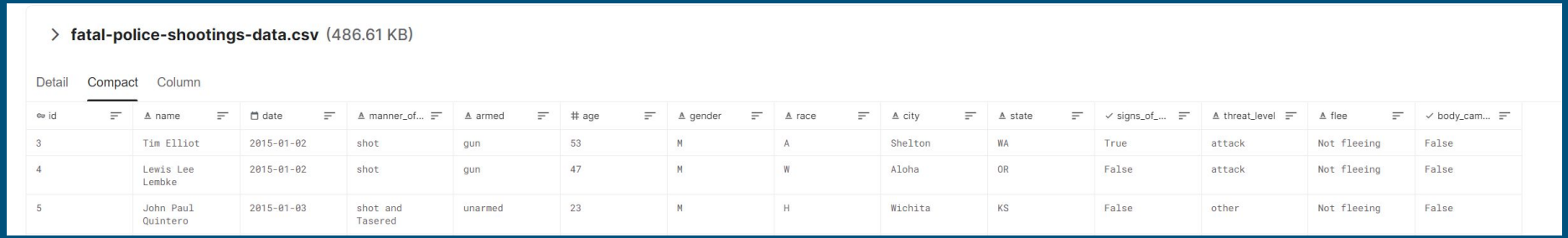
Project Goal: Examining the various factors that are related to police shootings in the US

Goal Description:

To analyze and visualize relationships between the independent variables (like gender, location, arms, age etc.) and the shooting incident. We can also explore incidents at different granularities like city or state or the arm used.

To try and predict (using multiple regression, or ML if possible) the combination of independent variables that are most likely to witness fatal shootings.

About the data



fatal-police-shootings-data.csv (486.61 KB)

Detail | Compact | Column

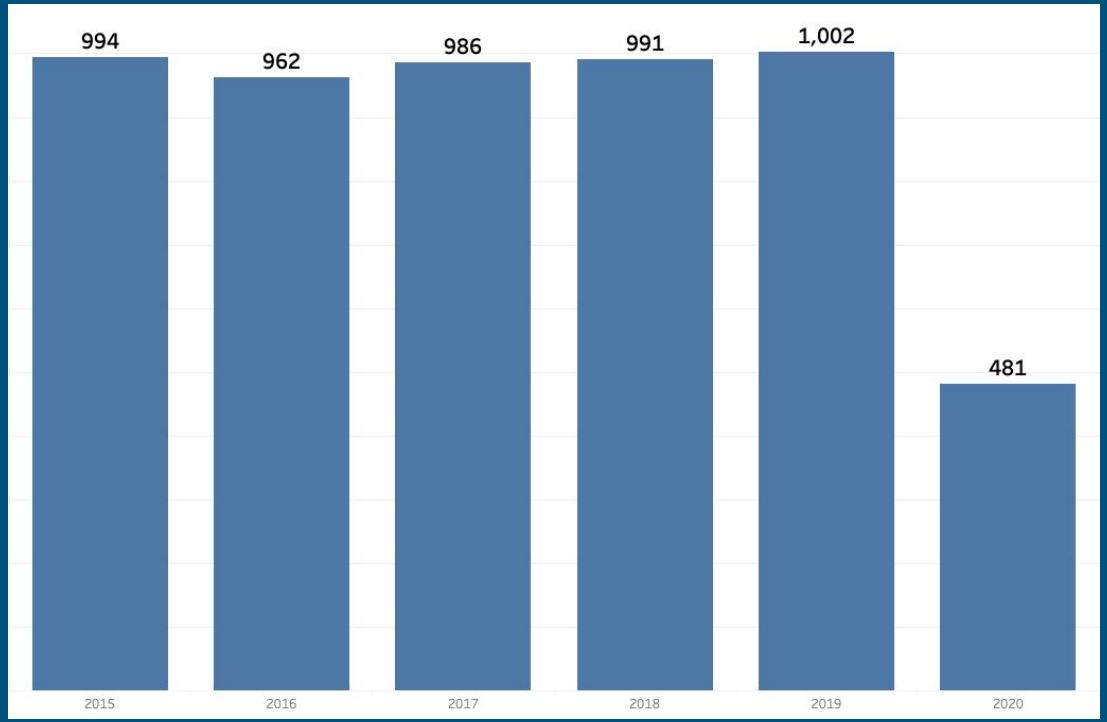
id	name	date	manner_of...	armed	age	gender	race	city	state	signs_of_...	threat_level	flee	body_cam...
3	Tim Elliot	2015-01-02	shot	gun	53	M	A	Shelton	WA	True	attack	Not fleeing	False
4	Lewis Lee Lembke	2015-01-02	shot	gun	47	M	W	Aloha	OR	False	attack	Not fleeing	False
5	John Paul Quintero	2015-01-03	shot and Tasered	unarmed	23	M	H	Wichita	KS	False	other	Not fleeing	False

- Source: Kaggle
- Date range: 1/2/2015 - 6/16 / 2020
- 5416 rows, 14 columns
  - Contained null values
- Both categorical and numeric data
  - data types - string, date, boolean, integer
- Required moderate level of data cleaning/filtering in BQ prior to visualizations

Exploratory questions

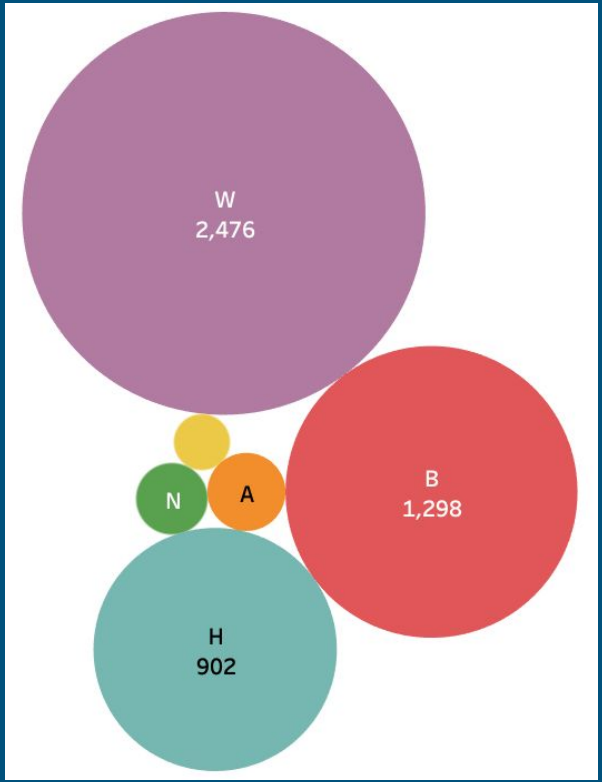
- Which state has the most deaths by a police officer?
- Which city has the highest number of deaths?
- What day has the highest deaths across the state level?
- How old are most of the shooting victims?
- Did they flee at the time of the shooting?
- What arms did the victims possess?
- What has been the timeline of these fatal shootings?
- What was the racial profile of the victims?

Year-wise number of fatal shootings



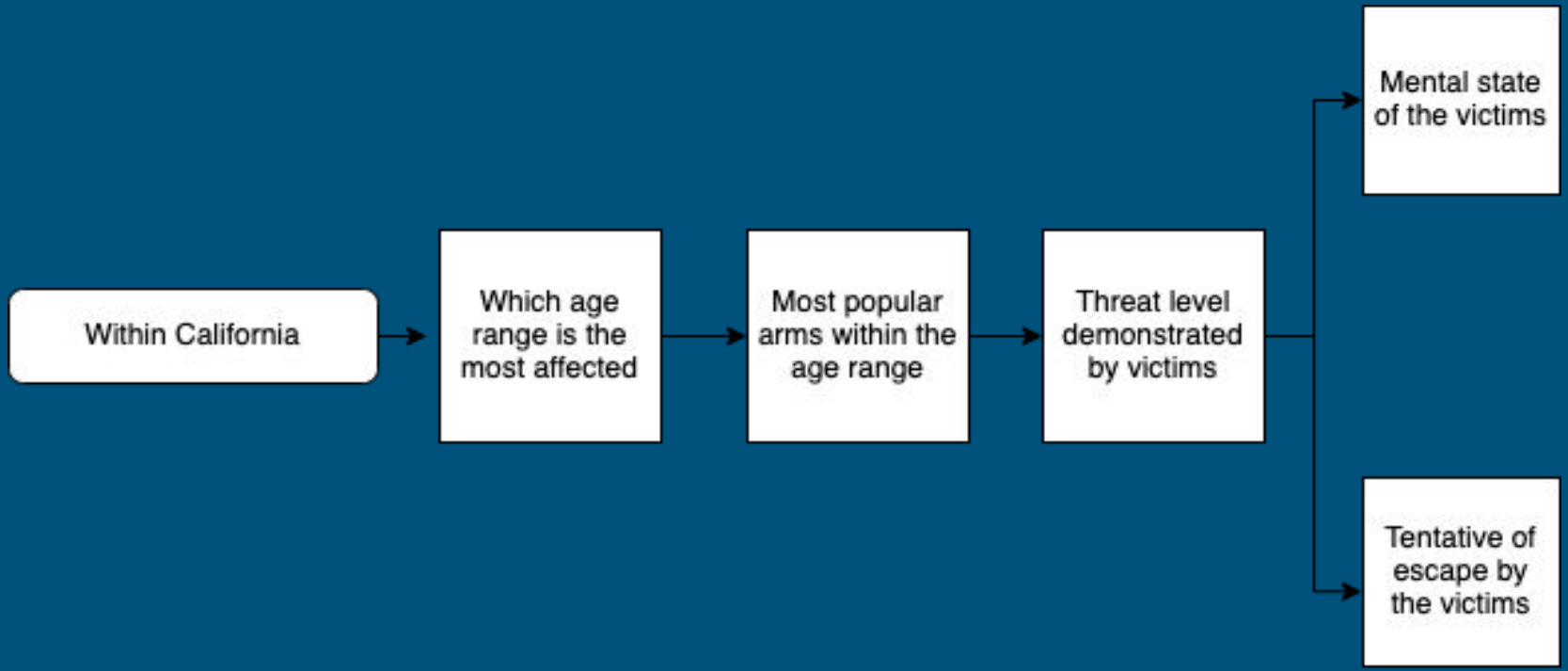
Note: The number for 2020 is only till Jun

Distribution of the victims by their race



W	White, non-Hispanic
B	Black, non-Hispanic
H	Hispanic
A	Asian
N	Native American
O	Other

Introduction to storyline



BQML: Predicting Signs of Mental Illness

Job information | Results | JSON | Execution details

Row	precision	recall	accuracy	f1_score	log_loss	roc_auc
1	0.5	0.03305785123966942	0.762278978388998	0.06201550387596899	0.5120102409792381	0.6959360639360639

BQML: Predicting Flee Method

Job information | Results | JSON | Execution details

Row	precision	recall	accuracy	f1_score	log_loss	roc_auc
1	0.32971887550200796	0.26875577100646353	0.7151277013752456	0.2449672052017377	1.8143612970489842	0.7135851648351648

Restrictions and limitations faced

Limitation	Details and solution
Raw Data	• While the dataset was overall of a refined quality but there were still some null values that had to be left out during our analysis • There were some columns with ambiguous entries for e.g. for the column *threat level* some of the entries were '*other', 'undetermined'* etc.
Coworking on Tableau	• Tableau Online was not very conducive to work in teams, we decided to finally use Tableau public and add all our charts during a zoom session
Data Analysis	• We had data for about 4 years, but there was no particular pattern in the yearly numbers, thus we chose to analyse the data as a whole rather than yearly

Thank you for your attention!

Any Questions?