

The Study of Gun Violence



Zhijie Chen, Nilkanth Patel, Pooja Suresh Babu

Data Sets

- Gun Violence Data

240,000 gun crimes happened within the past 5 years in the US

Added zip code, as key, for each entry through `uszipcode` library

- US Household Income Statistics

320,000 household income within given areas in the US

- Merged Dataset - (without Duplicates)

215,000 gun crimes with household income of all states in the US

Features

| | |
|-----------------------------|--------|
| Unnamed: 0 | 0 |
| incident_id | 0 |
| date | 0 |
| state | 0 |
| city_or_county | 0 |
| address | 13631 |
| n_killed | 0 |
| n_injured | 0 |
| incident_url | 0 |
| source_url | 392 |
| incident_url_fields_missing | 0 |
| congressional_district | 3678 |
| gun_stolen | 92462 |
| gun_type | 92440 |
| incident_characteristics | 297 |
| latitude | 0 |
| location_description | 177250 |
| longitude | 0 |
| n_guns_involved | 92440 |
| notes | 73037 |
| participant_age | 83617 |
| participant_age_group | 37899 |
| participant_gender | 32657 |
| participant_name | 112147 |
| participant_relationship | 200650 |
| participant_status | 24682 |
| participant_type | 22124 |
| sources | 534 |

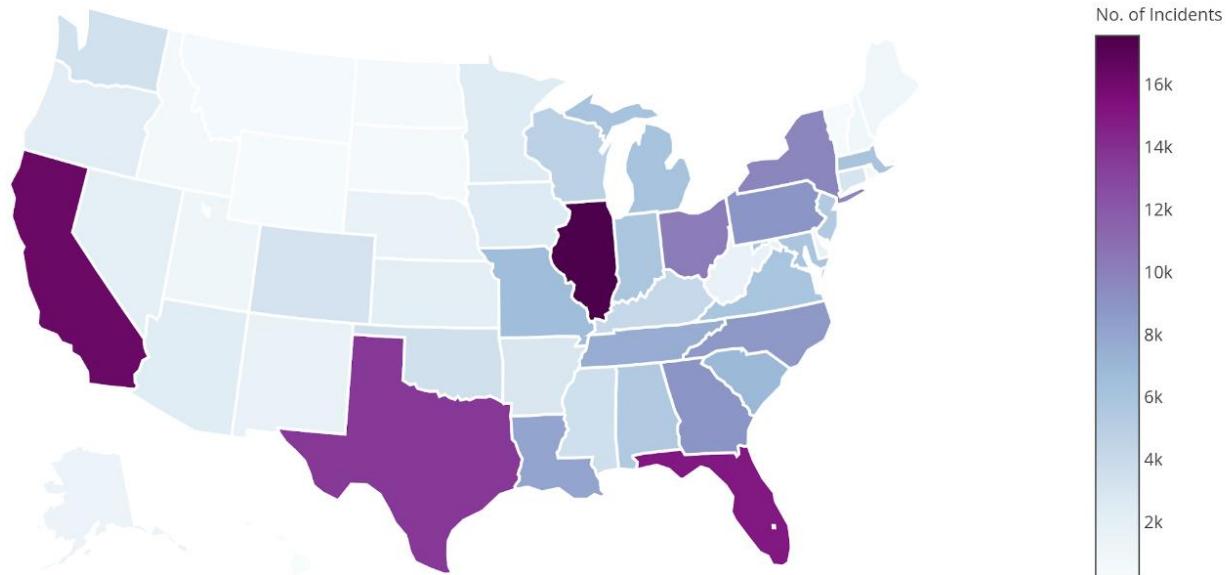
| | |
|-----------------------|-------|
| participant_type | 22124 |
| sources | 534 |
| state_house_district | 28019 |
| state senate district | 21581 |
| Zipcode | 0 |
| id | 0 |
| State_Code | 0 |
| State_Name | 0 |
| State_ab | 0 |
| County | 0 |
| City | 0 |
| Place | 0 |
| Type | 0 |
| Primary | 0 |
| Area_Code | 0 |
| ALand | 0 |
| AWater | 0 |
| Lat | 0 |
| Lon | 0 |
| Mean | 0 |
| Median | 0 |
| Stdev | 0 |
| sum_w | 0 |
| dtype: int64 | |

The Use of Algorithms

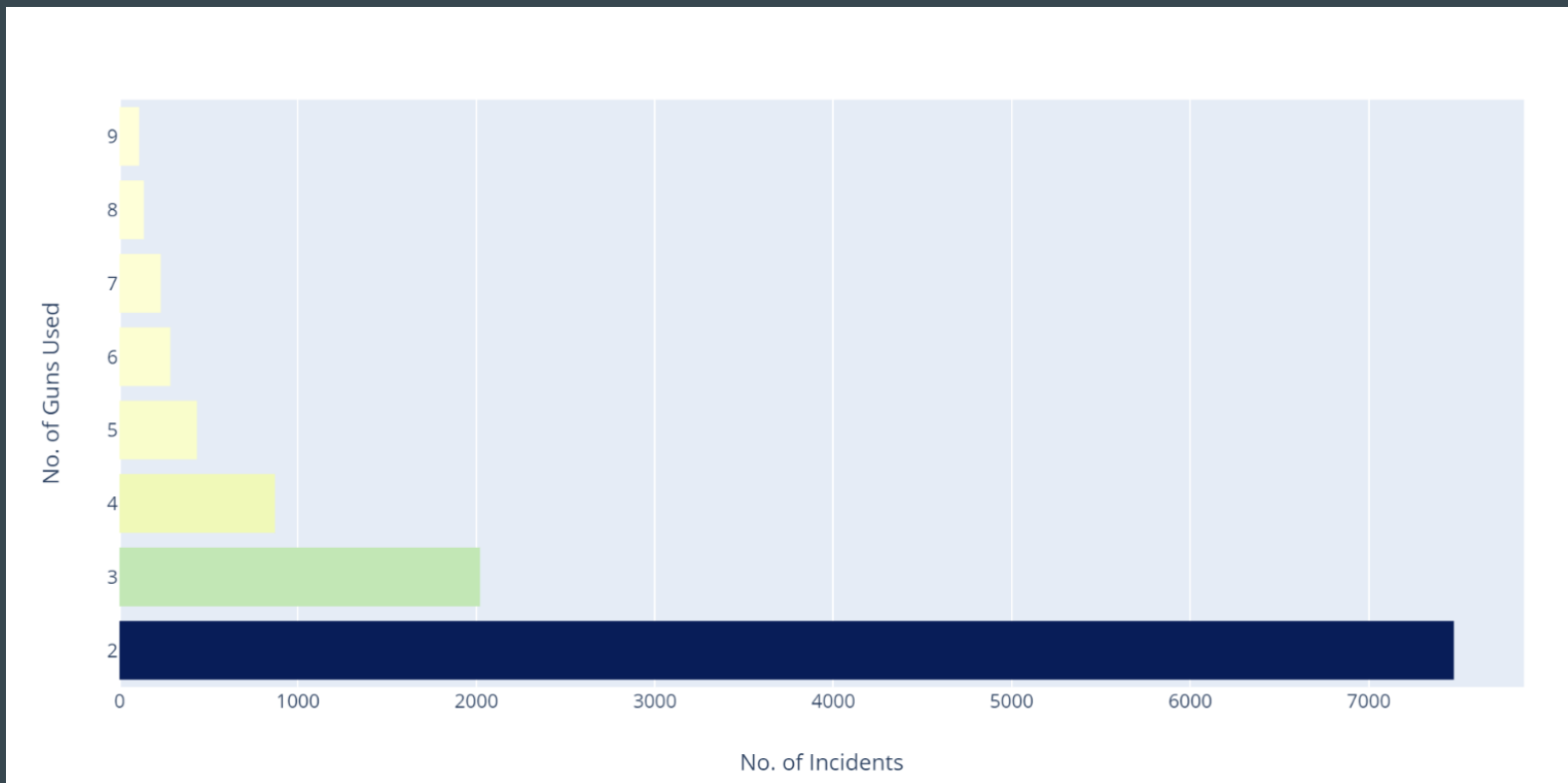
- Spark
 - Binding gun violence with household income with zip code as key
 - Merging data with respective keys (e.g. age, state)
- SON
 - Finding frequent pairs (e.g. gender/income and gun violence)
- CURE
 - Grouping gun crimes based on latitude and longitude

Number of gun incidents Vs State

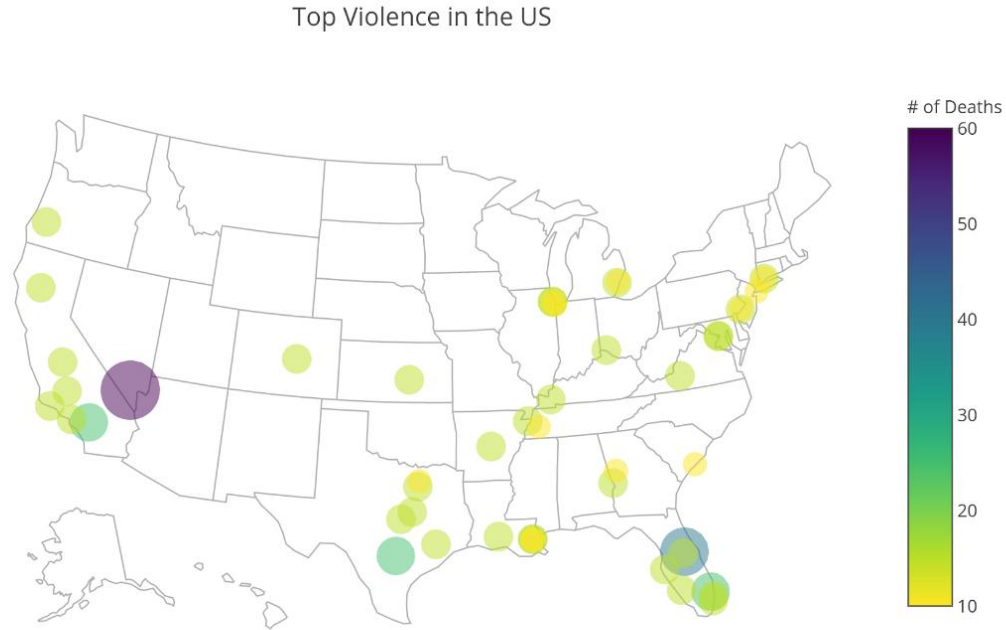
Total Gun Violence Incidents in the US



Number of Guns Used



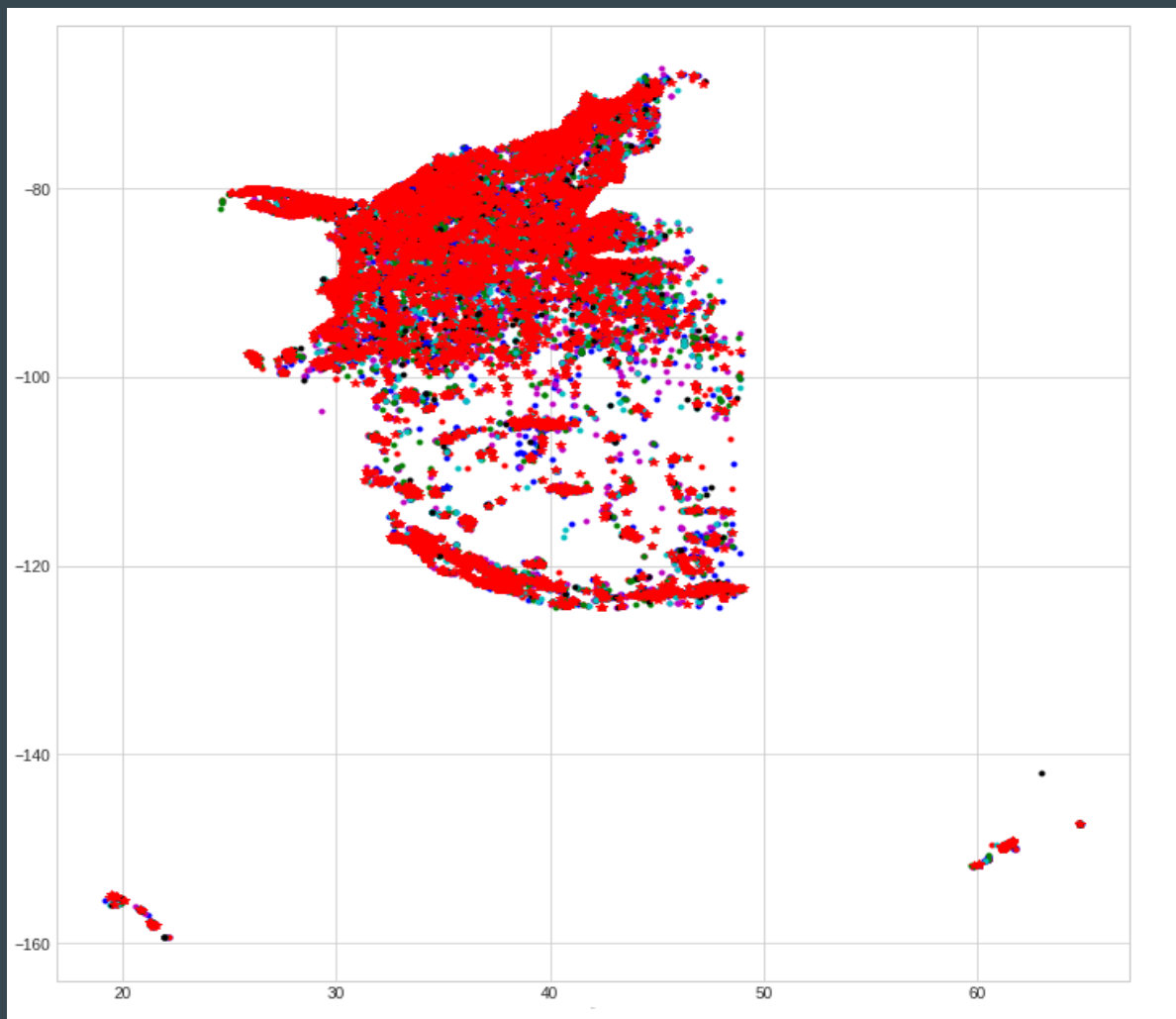
Most serious Gun Crimes by deaths



Suspects Age



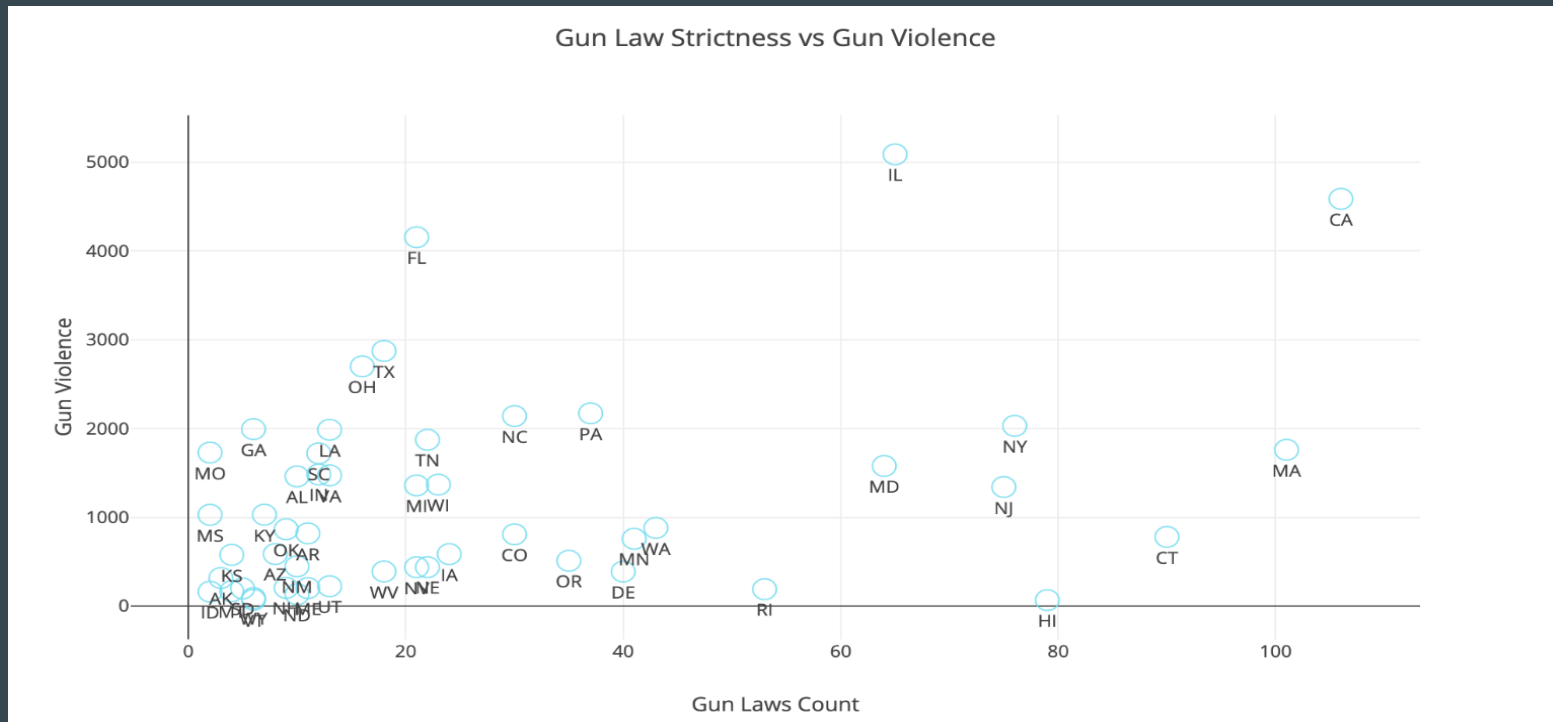
Clustering - CURE



Clustering - CURE



Findings - Gun Law and Gun Violence



gun law count source : <https://statefirearmlaws.org/national-data/>

Findings - Gender and Gun Violence

Gender: {-1000: Male, 1000 Female}

```
{(17,): 35, (6,): 71, (-1000, 6): 44, (50,): 78, (-1000, 50): 61, (5, 1000): 73, (-1000, 5): 104, (4, 1000): 203, (-1000, 4): 345, (3, 1000): 664, (-1000, 3): 1277, (2, 1000): 2809, (-1000, 2): 7286, (1, 1000): 8070, (0, 1000): 25423, (1,): 74329, (-1000, 1): 66259, (0,): 223586, (-1000, 0): 198163, (-1000,): 273627}
```

Rules:

```
{(50,): ((-1000,), 0.782051282051282), (1,): ((-1000,), 0.891428648306852), (0,): ((-1000,), 0.8862943118084317)}
```

- In most Mass Shootings, the suspect is likely a Male.
- Women suspects are most likely Not to Kill people but majorly Injure them.

Findings- Income and Gun Violence

0 - 0 killed, possibly injured

```
{(38100.0,): 875, (17900.0,): 880, (29100.0,): 893, (35000.0,): 980, (41400.0,): 1006, (51600.0,): 1041, (0, 51600.0): 883, (20300.0,): 1055, (0, 20300.0): 865, (25700.0,): 1543, (0, 25700.0): 1291, (2,): 3823, (1, 300000.0): 1778, (0, 300000.0): 8027, (1,): 40898, (0,): 160540}
```

Rules:

```
{(51600.0,): ((0,), 0.8482228626320846), (20300.0,): ((0,), 0.8199052132701422), (25700.0,): ((0,), 0.8366817887232664)}
```

- Higher income regions are most likely to Not have Mass shootings.
- Ex: California has Higher Income and More gun incidents

Findings - Special Cases

Rules:

```
{(27617,): ((60,), 0.6710526315789473), (23704,): ((1,), 0.5825242718446602), (23523,): ((1,), 0.5229357798165137), (33607,): ((60,), 0.6363636363636364), (33126,): ((60,), 0.5132743362831859), (33004,): ((60,), 0.51875), (77061,): ((60,), 0.5714285714285714), (75235,): ((60,), 0.6073619631901841), (82609,): ((7,), 0.5930521091811415)}
```

- Among Texas, California, Florida, Virginia, Connecticut, Oklahoma and Nevada. - Gun Incidents because of **TSA violation** are mostly in Florida and Texas regions (Zip Codes).

| | | | |
|-----------------|--------------|----|---------|
| 27617 - Raleigh | Wake | NC | |
| 33607 - Tampa | Hillsborough | FL | |
| 33126 - Miami | Miami-Dade | FL | 33004 - |
| Dania | Broward | FL | |
| 77061 - Houston | Harris | TX | 75235 - |
| Dallas | Dallas | TX | |