IS 425 Group 2

Team Responsibilities

Garret Wong – Data Cleaner

- Gathered and preprocessed dataset

Luke Dyer – Data Analyst

- Utilized Weka to perform data analysis

Ephraim Schlessinger – Visualization Modeler

- Utilized Tableau to perform data visualization

Problem: Gun Violence in the United States

We aim to create a Decision Support System Application that will utilize a visualization tool to help lawmakers make informed and effective decisions to reduce gun violence in the United States.

Dataset (Part 1)

Our 1st dataset comes from Kaggle:
(https://www.kaggle.com/datasets/jameslko/gun-violence-data/)
and contains data gathered from the Gun Violence Archive.

Contains >230,000 US gun violence incidents from 2013-2018.

Our next 2 data sets come from the U.S. Census Bureau:

https://www.census.gov/data/datasets/time-series/demo/popest/2010s-state-total.html

https://www.census.gov/data/datasets/time-series/demo/popest/2010s-total-cities-and-towns.html

We used the first excel spreadsheet for each, which are the annual estimates.

Dataset (Part 2)

Dataset Fields:

incident_id, **date**, **state**, **city_or_county**, address, **n_killed**, **n_injured**, incident_url, source_url, incident_url_fields_missing, congressional_district, gun_stolen, gun_type, incident_characteristics,

latitude, location_description, **longitude**, n_guns_involved, notes, participant_age, participant_age_group, participant_gender, participant_name, participant_relationship, participant_status, participant_type, sources, **state_house_district**, **state_senate_district**

Dataset (Part 3)

Both population data sets had numerous rows and fields. We picked only those that were relevant to the few cities and states that we tabulated values for by hand. This was done by hand because we were unable to blend the data in Tableau and get it to show in a readable/understandable manner.

The fields used from the State Populations dataset were California, Texas, Florida, Illinois, and Maryland along with their population fields from 2014-2017.

The fields used from the City Populations dataset were Chicago, Houston, Baltimore, Los Angeles, Jacksonville, San Antonio, and Oakland along with their population fields from 2014-2017.

Data Pre-Processing Of Gun-Violence Dataset

Missing values under notes and data_description, however unnecessary.

Data Analytics:

state, n_killed, n_injured, latitude, longitude, state_house_district, and state_senate_district

Data Visualization:

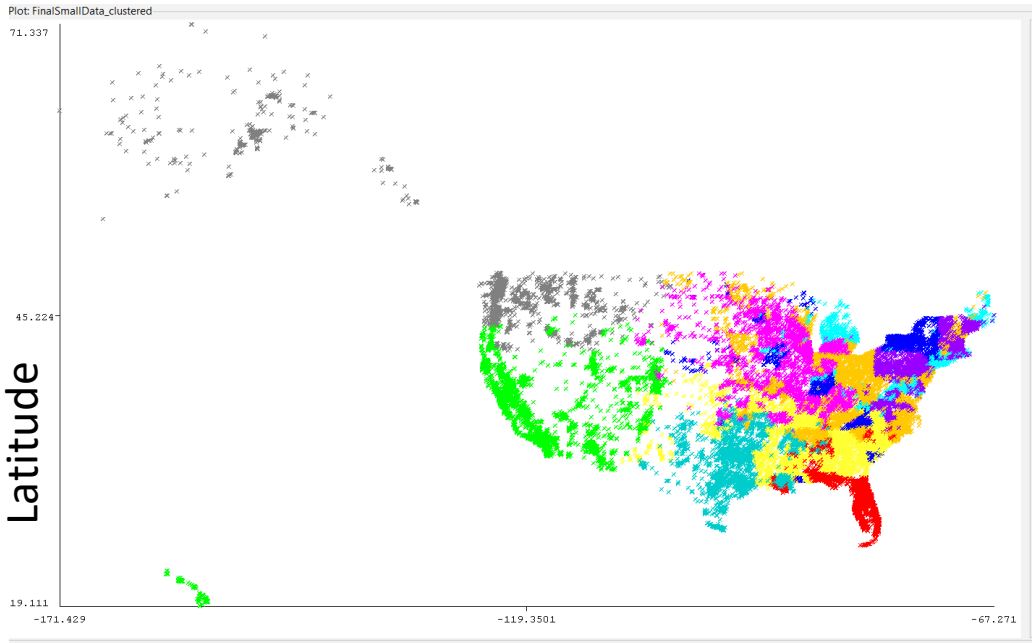
date, city_or_county, state, n_killed, n_injured, total incidents, participant_age, participant_age_group

Analytics

- Method Used: Kmeans Clustering
- Reason: Find areas with disproportionately high rates of gun violence
- Attributes Used: State, num killed, num injured, State house district, state senate district, longitude, and latitude.
- Number of Cluster: 7,10,12
- Problems: Couldn't include specifics about incidents

Notable clusters: 5,8,2

10 Clusters Including State Districts



Final cluster centroids:

Attribute	Full Data (239654.0)	Cluster# 0 (12627.0)	1 (19047.0)	2 (25615.0)	3 (15946.0)	4 (26729.0)	5 (40921.0)
state	Illinois	New York	Florida	California	New York	Pennsylvania	Illinois
n_killed	0.2517	0.2015	0.2622	0.3387	0.2263	0.1817	0.2148
n_injured	0.4926	0.4705	0.4747	0.4384	0.5212	0.4503	0.5639
state_house_district	55.4461	93.9159	53.2765	30.7359	58.843	60.0193	48.5612
state_senate_district	20.4772	49.4273	18.4275	15.3551	29.0694	19.8157	20.6642
latitude	37.547	40.173	29.4602	36.3381	40.6043	40.2187	40.342
longitude	-89.3421	-81.4452	-82.6755	-116.8667	-77.8729	-75.4821	-89.574

Time taken to build model (full training data

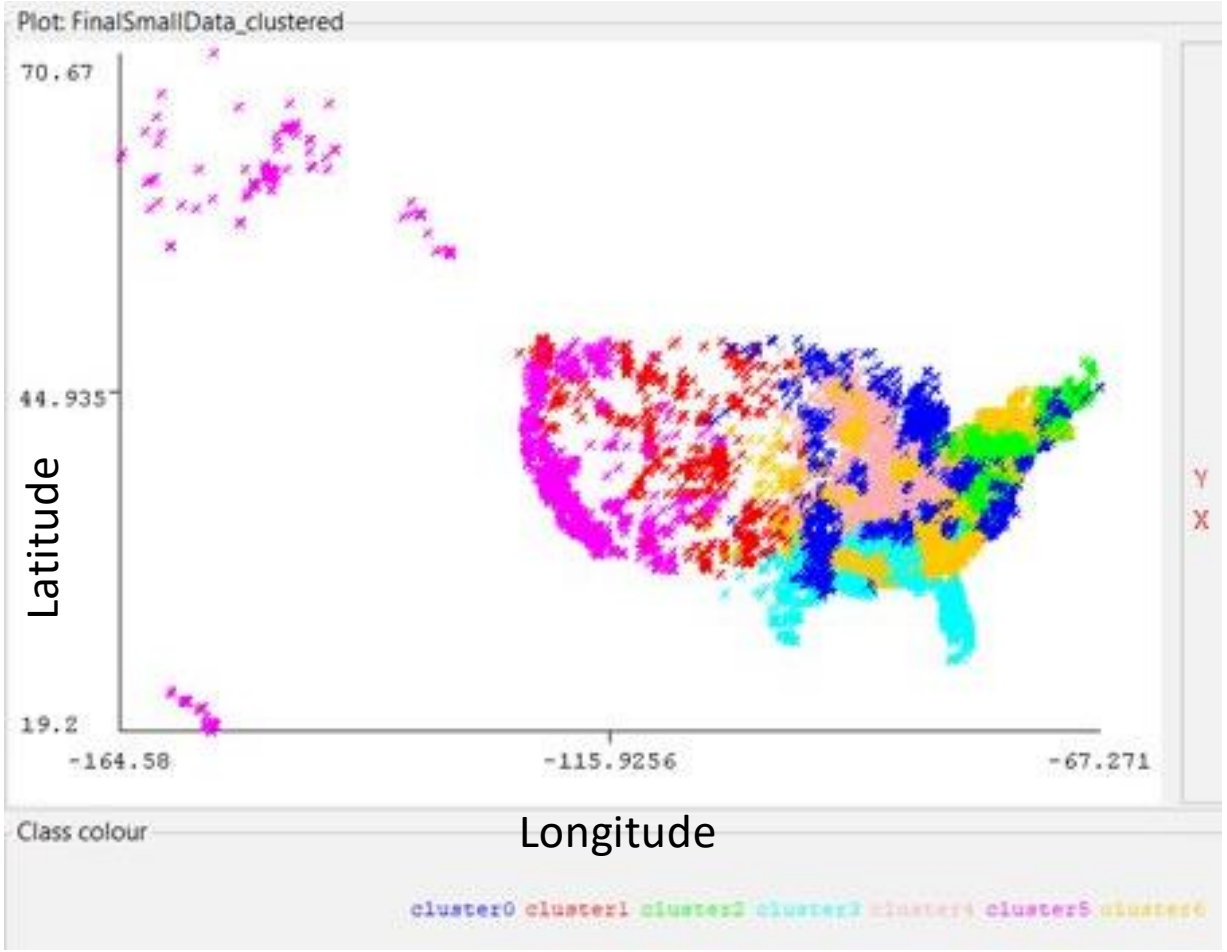
=== Model and evaluation on training set ===

Clustered Instances

0 12627 (5%)
1 19047 (8%)
2 25615 (11%)

	6 (40155.0)	7 (29831.0)	8 (21128.0)	9 (7655.0)
	Ohio	Georgia	Texas	Washington
	0.2417	0.27	0.3521	0.2174
	0.5363	0.4912	0.4679	0.3087
	34.0018	84.3971	78.1891	30.7291
	8.4892	30.1023	13.858	22.0189
	39.3515	33.8975	31.2689	48.8263
	-83.0523	-87.0423	-94.928	-124.4398

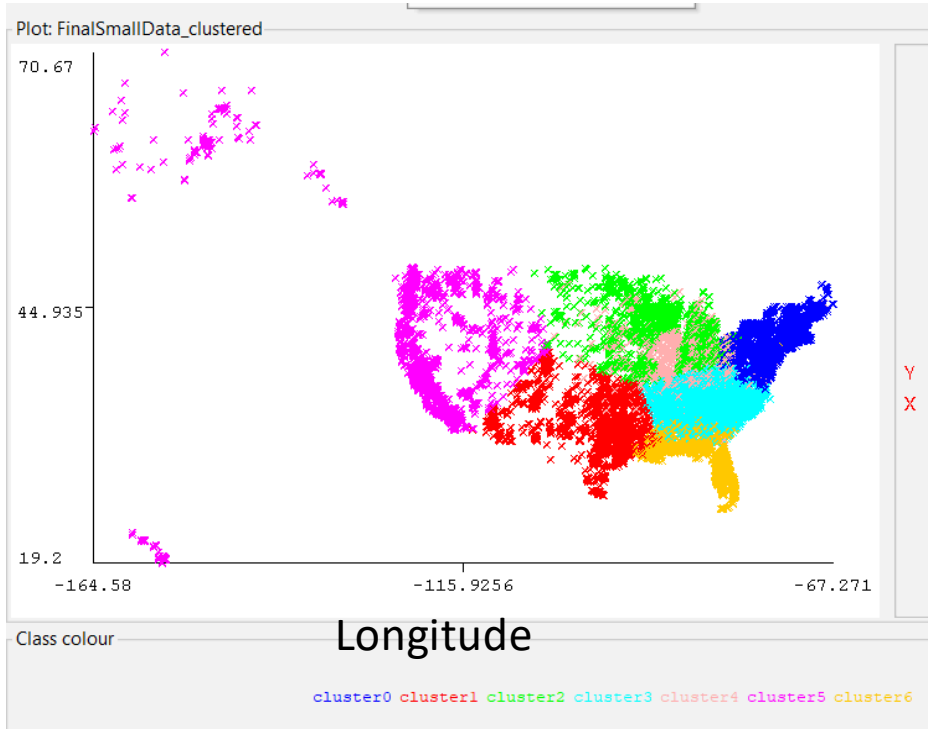
7 Clusters Including State Districts



Notable Clusters: 0,4,2

```
Final cluster centroids:
                                        Cluster#
Attribute                  Full Data            0            1            2            3            4            5            6
                          (158171.0)   (32000.0)    (7790.0)   (23616.0)   (21969.0)   (26978.0)   (16868.0)   (28950.0)
=============================================================================================================================
state                        Illinois     Michigan     Colorado Pennsylvania      Florida     Illinois   California      Georgia
n_killed                       0.2519       0.2437       0.2457          0.2       0.3044        0.236       0.3268       0.2364
n_injured                      0.4912       0.5127       0.3316       0.4593       0.4753       0.5963       0.4305       0.4857
state_house_district          55.3897      44.4426      43.8471      54.6396      64.0629      49.5089      28.5344      85.7539
state_senate_district         20.4782       7.6622      25.8661      21.4777      19.1481      20.1442       14.881      36.9612
latitude                      37.5557       38.185      40.5612      40.2359      30.0686      39.9075      38.4504      36.8334
longitude                    -89.3494      -84.8187     -109.8814     -75.8969     -87.4838     -88.6239     -120.1385     -83.9587



Time taken to build model (percentage split) : 2.46 seconds

Clustered Instances

0      16803 ( 21%)
1       4090 (  5%)
2      11907 ( 15%)
```

7 Clusters Without Districts



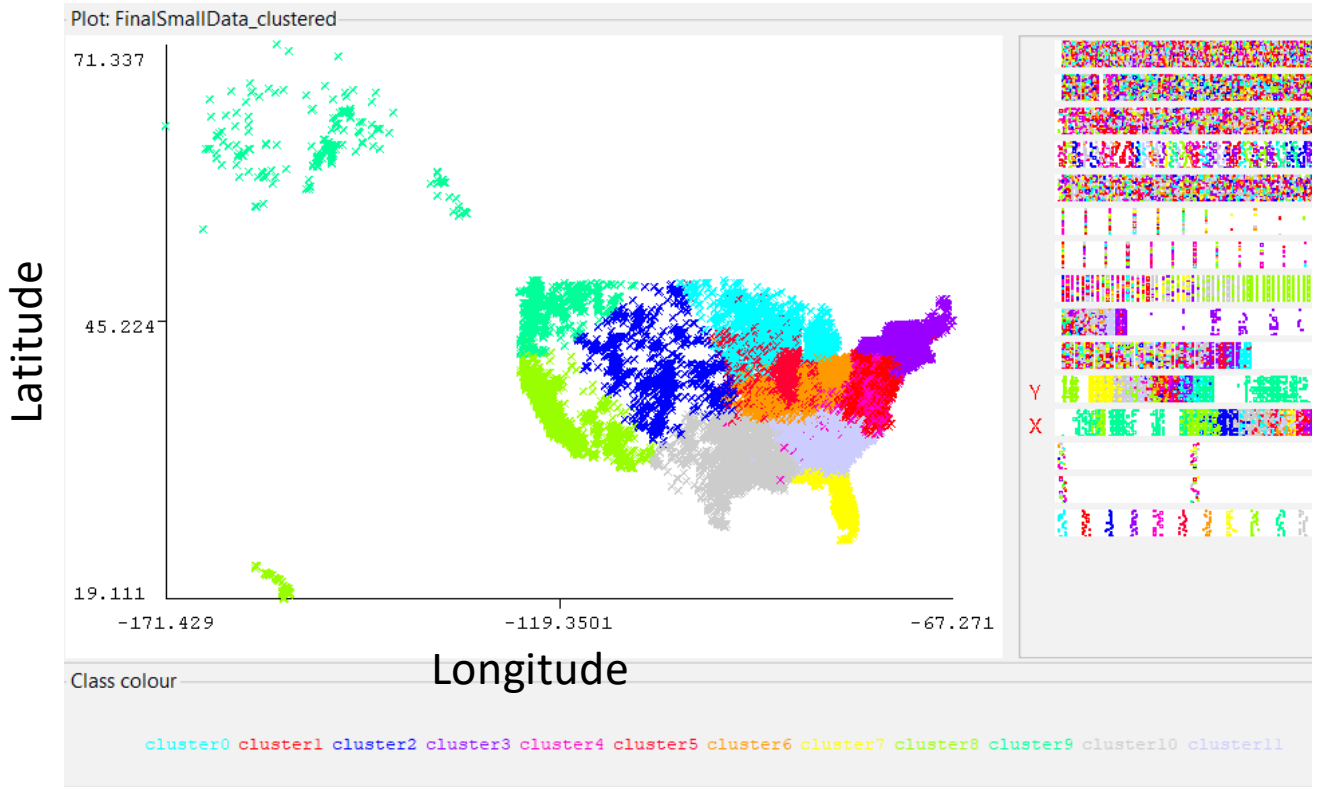
Notable Clusters: 2,4,6

Final cluster centroids:

Attribute	Full Data (158171.0)	Cluster# 0 (33032.0)	1 (18339.0)	2 (16388.0)	3 (27713.0)	4 (25397.0)	5 (18920.0)	6 (18382.0)
state	Illinois	New York	Texas	Wisconsin	North Carolina	Illinois	California	Florida
n_killed	0.2519	0.1666	0.3441	0.3335	0.3345	0.1394	0.2993	0.2228
n_injured	0.4912	0.4069	0.4094	0.0962	0.3221	1.1054	0.4081	0.568
latitude	37.5557	40.6436	33.3178	41.6563	35.2195	40.4662	39.9158	29.6504
longitude	-89.3494	-75.2236	-98.7048	-88.8765	-83.9584	-87.3146	-120.4095	-84.7907

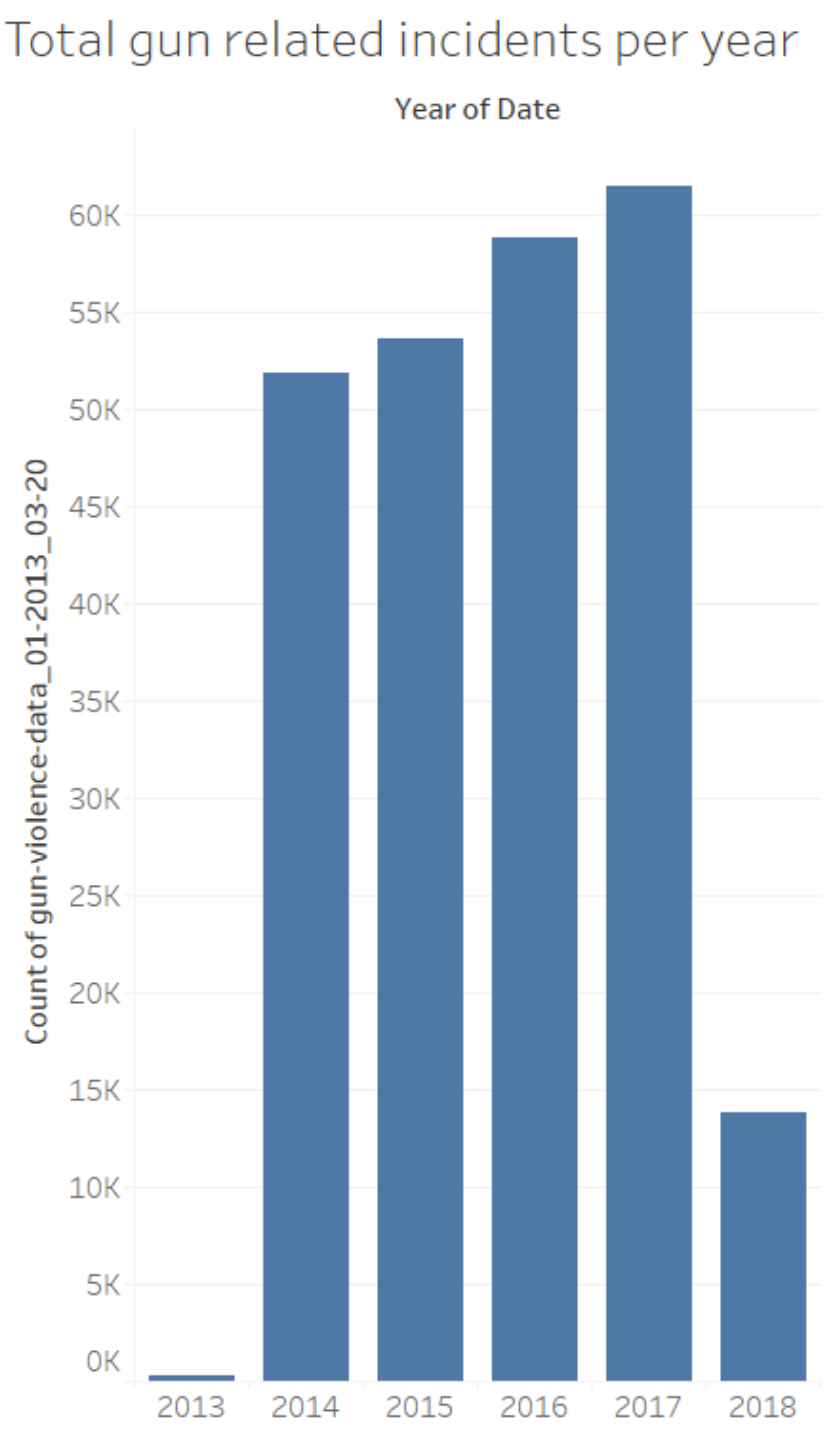
12 Clusters Without Districts

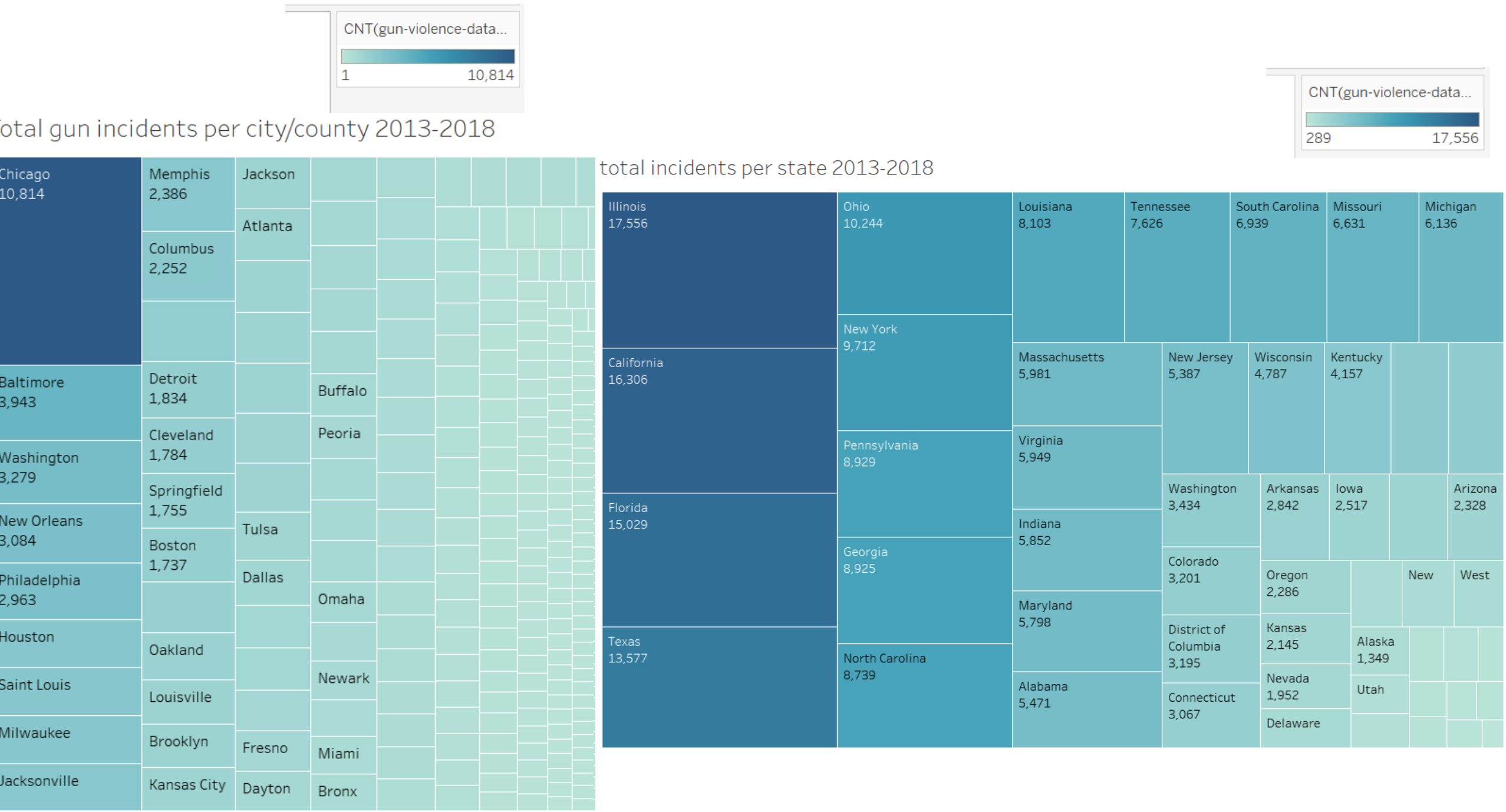
Notable Clusters: 1,4,5,6

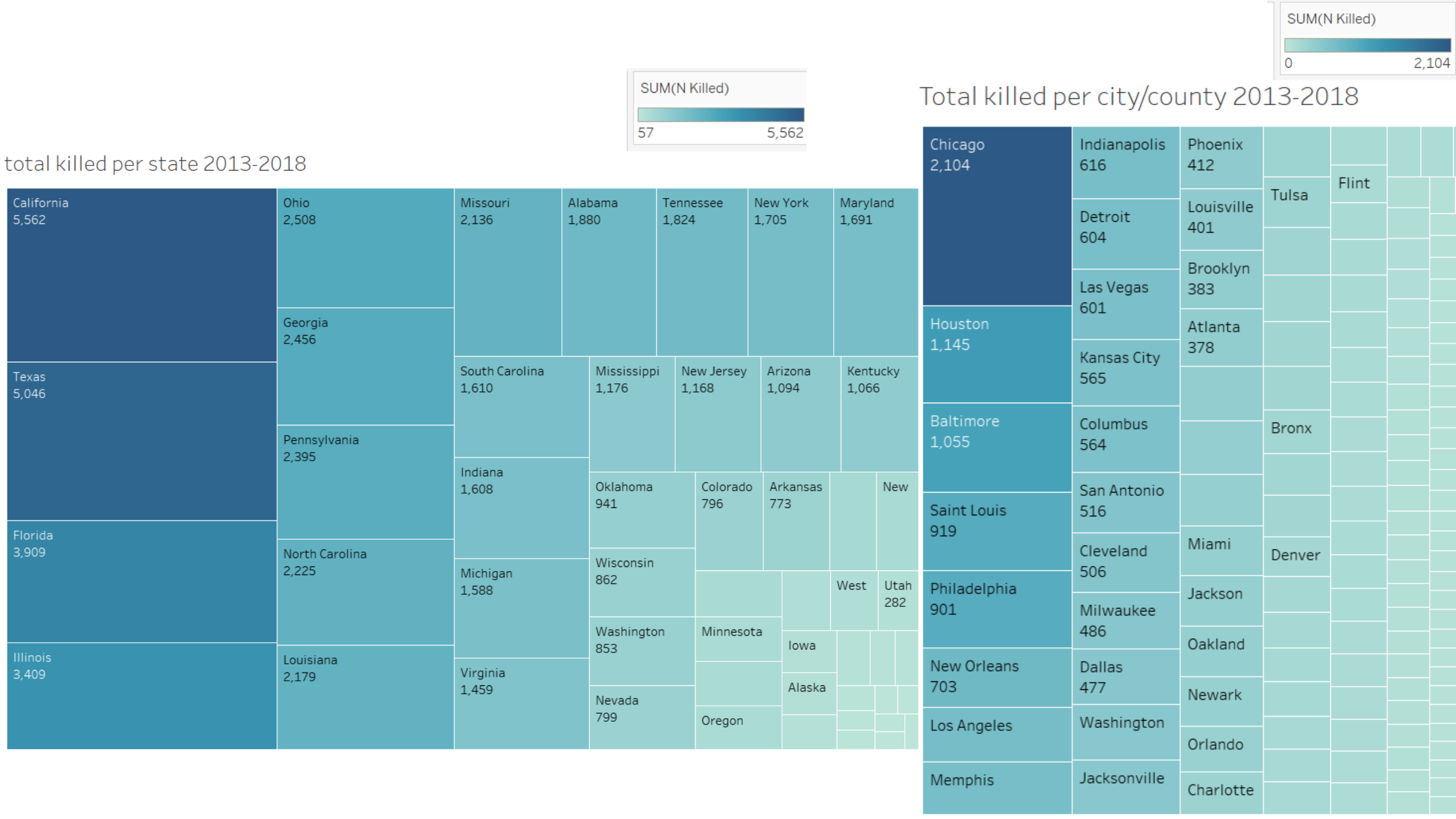


Cluster#											
0	1	2	3	4	5	6	7	8	9	10	11
(15109.0)	(17692.0)	(7711.0)	(23575.0)	(17315.0)	(29218.0)	(23595.0)	(15082.0)	(20837.0)	(7482.0)	(30245.0)	(31793.0)
Michigan	Pennsylvania	Colorado	New York	Pennsylvania	Illinois	Ohio	Florida	California	Washington	Texas	Georgia
0.1993	0.4604	0.2423	0.103	0.0788	0.1411	0.4019	0.2553	0.3596	0.2169	0.3235	0.225
0.2419	0.0051	0.2812	0.3588	1.38	0.9371	0.2592	0.4639	0.468	0.314	0.4653	0.4132
42.9825	38.9739	39.5799	41.8909	38.4485	40.6918	39.2336	28.4946	35.7059	48.845	31.767	34.1635
-88.8966	-77.1701	-105.7778	-73.6074	-77.785	-88.8656	-86.3574	-81.9167	-118.713	-125.3603	-94.3907	-83.9496

Data visualization



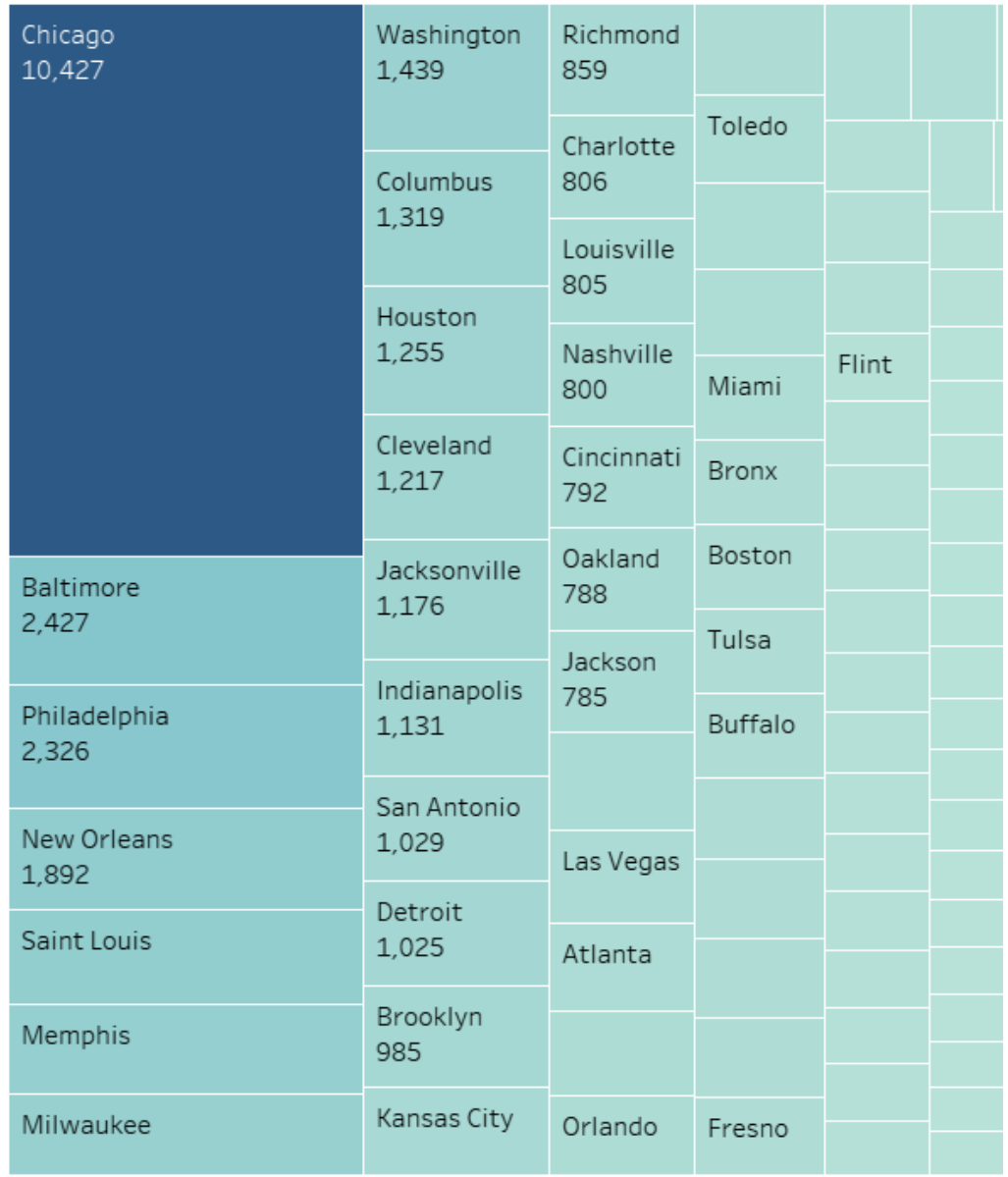

CNT(gun-violence-data...
1
10,814
Total gun incidents per city/county 2013-2018
Chicago
10,814
Baltimore
3,943
Washington
3,279
New Orleans
3,084
Philadelphia
2,963
Houston
Saint Louis
Milwaukee
Jacksonville
Memphis
2,386
Columbus
2,252
Detroit
1,834
Cleveland
1,784
Springfield
1,755
Boston
1,737
Oakland
Louisville
Brooklyn
Kansas City
Jackson
Atlanta
Tulsa
Dallas
Fresno
Dayton
Buffalo
Peoria
Omaha
Newark
Miami
Bronx
CNT(gun-violence-data...
289
17,556
total incidents per state 2013-2018
Illinois
17,556
California
16,306
Florida
15,029
Texas
13,577
Ohio
10,244
New York
9,712
Pennsylvania
8,929
Georgia
8,925
North Carolina
8,739
Louisiana
8,103
Tennessee
7,626
South Carolina
6,939
Missouri
6,631
Michigan
6,136
Massachusetts
5,981
Virginia
5,949
Indiana
5,852
Maryland
5,798
Alabama
5,471
New Jersey
5,387
Wisconsin
4,787
Kentucky
4,157
Washington
3,434
Colorado
3,201
District of Columbia
3,195
Connecticut
3,067
Arkansas
2,842
Iowa
2,517
Arizona
2,328
Oregon
2,286
Kansas
2,145
Nevada
1,952
Delaware
Alaska
1,349
Utah
New
West


SUM(N Killed)
57
5,562
total killed per state 2013-2018
California 5,562
Texas 5,046
Florida 3,909
Illinois 3,409
Ohio 2,508
Georgia 2,456
Pennsylvania 2,395
North Carolina 2,225
Louisiana 2,179
Missouri 2,136
Alabama 1,880
Tennessee 1,824
New York 1,705
Maryland 1,691
South Carolina 1,610
Mississippi 1,176
New Jersey 1,168
Arizona 1,094
Kentucky 1,066
Indiana 1,608
Oklahoma 941
Colorado 796
Arkansas 773
New
Michigan 1,588
Wisconsin 862
West
Utah 282
Washington 853
Minnesota
Iowa
Virginia 1,459
Nevada 799
Alaska
Oregon
SUM(N Killed)
0
2,104
Total killed per city/county 2013-2018
Chicago 2,104
Houston 1,145
Baltimore 1,055
Saint Louis 919
Philadelphia 901
New Orleans 703
Los Angeles
Memphis
Indianapolis 616
Detroit 604
Las Vegas 601
Kansas City 565
Columbus 564
San Antonio 516
Cleveland 506
Milwaukee 486
Dallas 477
Washington
Jacksonville
Phoenix 412
Louisville 401
Brooklyn 383
Atlanta 378
Miami
Jackson
Oakland
Newark
Orlando
Charlotte
Tulsa
Bronx
Denver
Flint

SUM(n_injured)

0 10,427

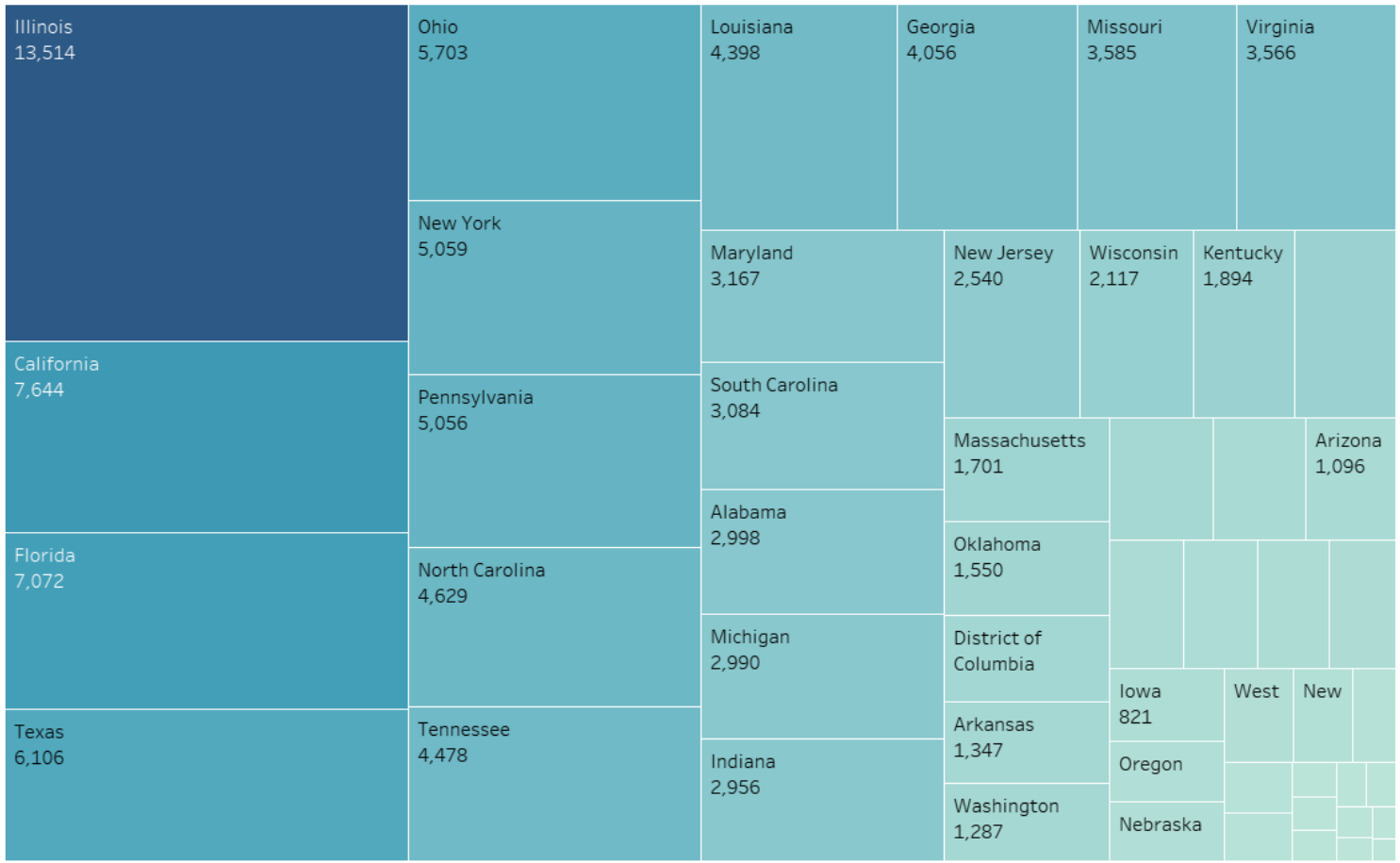
total injured by city/county 2013-2018

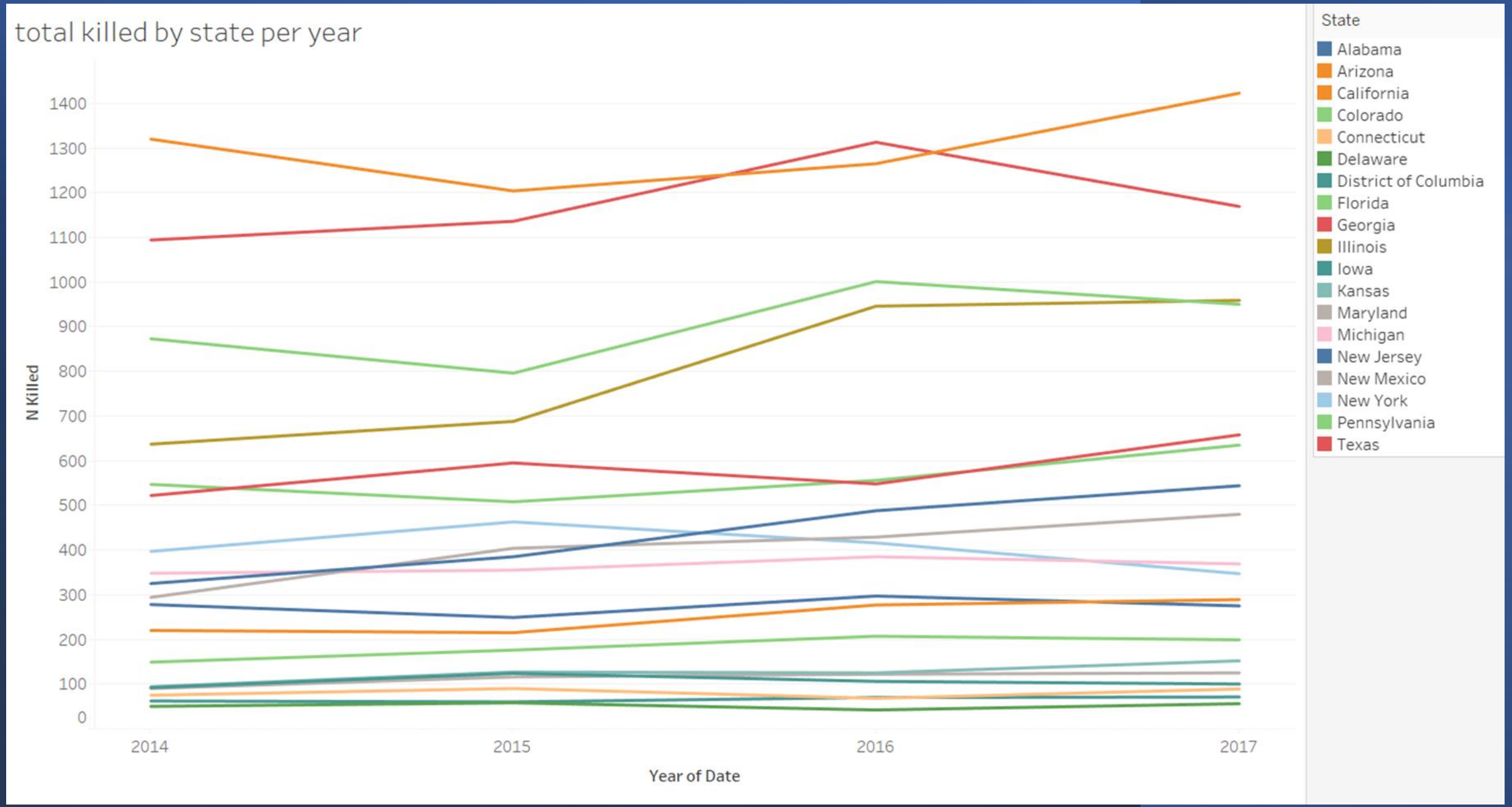


SUM(n_injured)

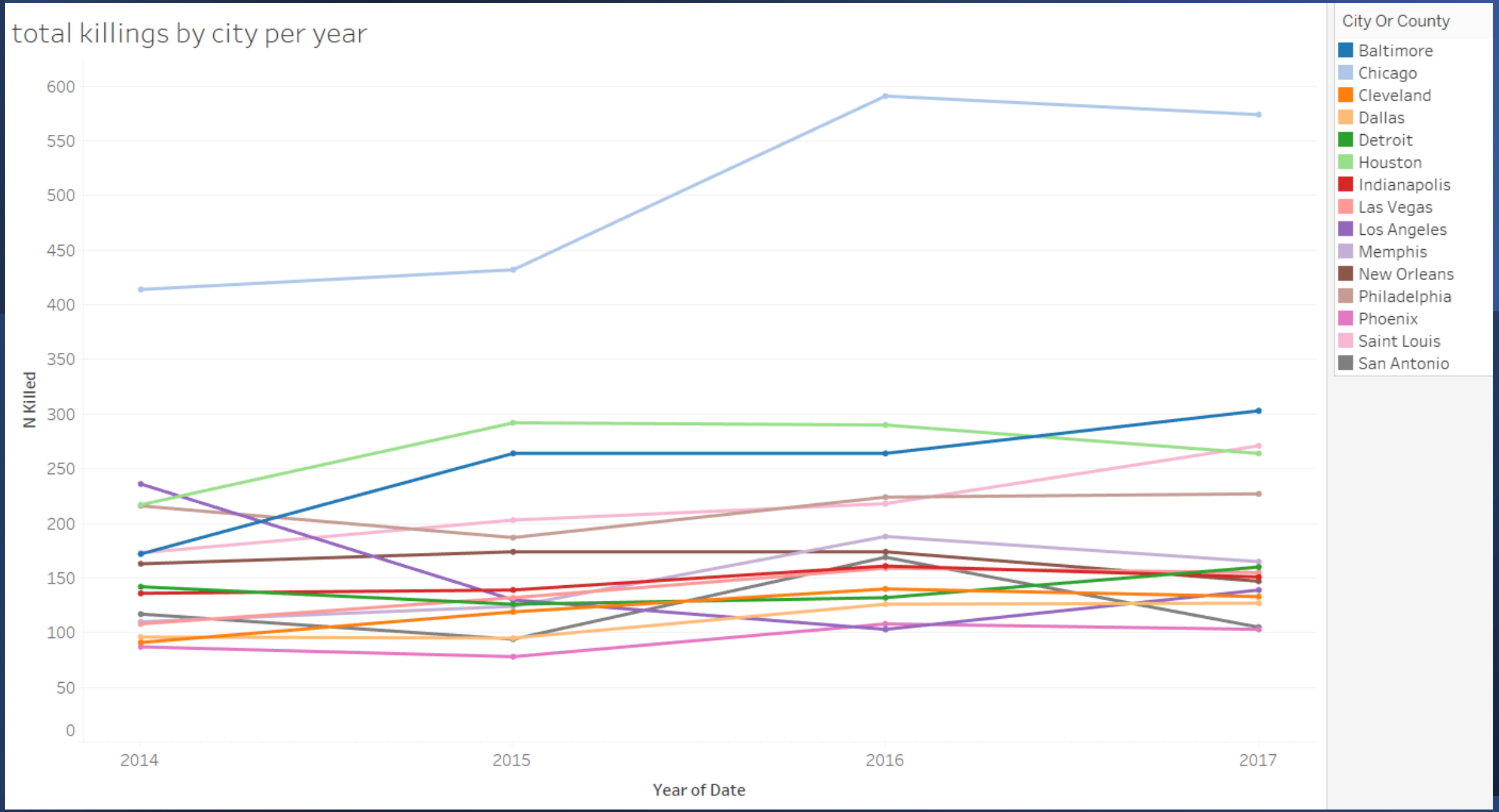
52 13,514

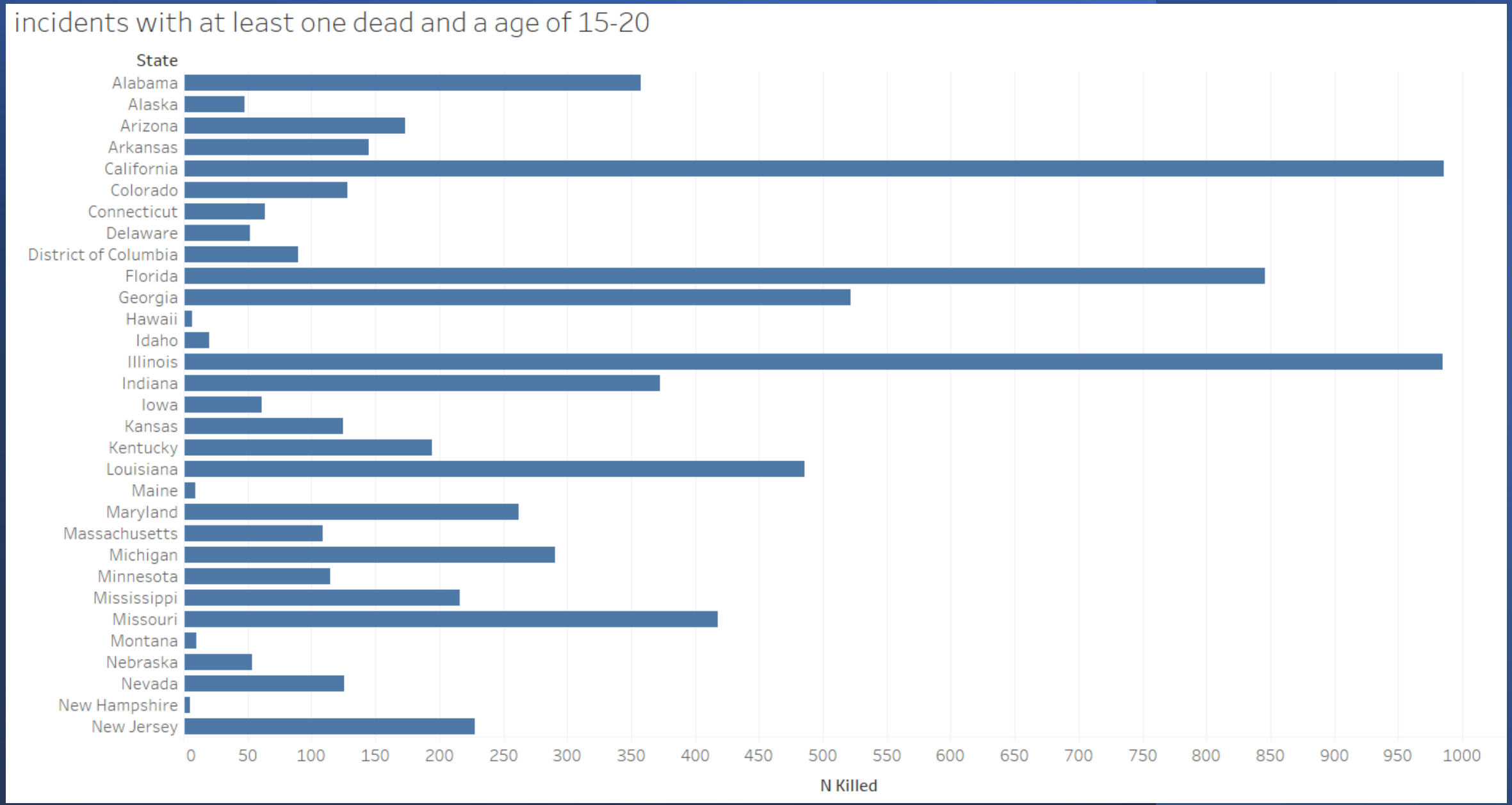
Total injured by state 2013-2018

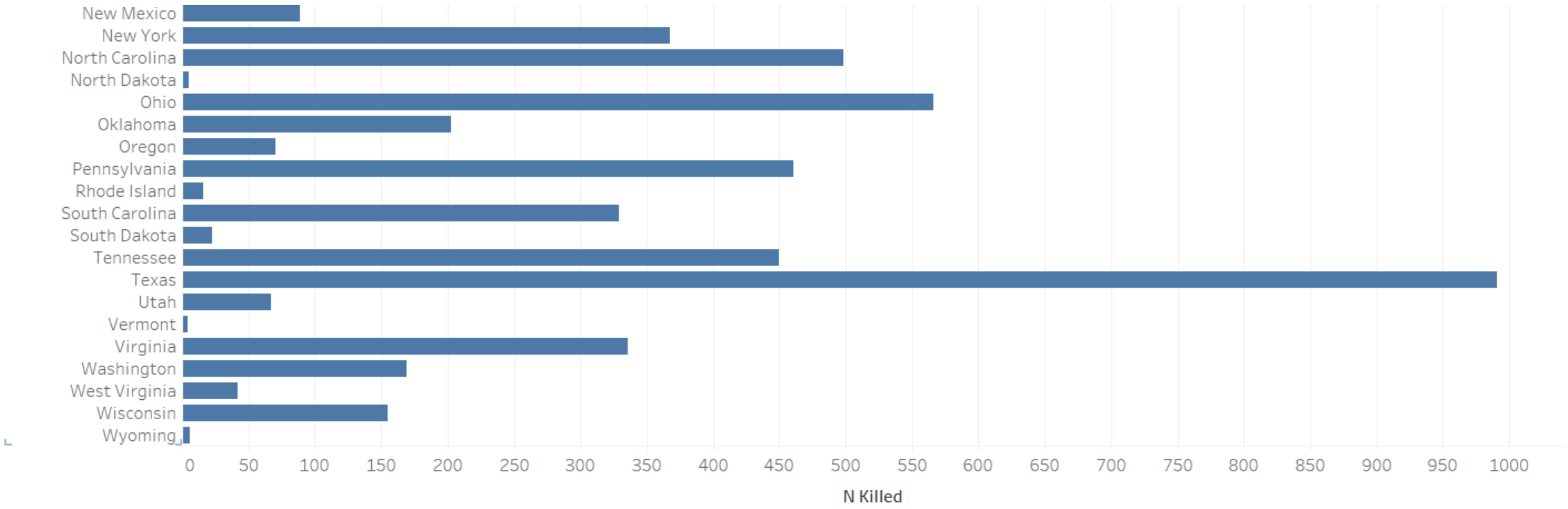



total killed by state per year
N Killed
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
2014
2015
2016
2017
Year of Date
State
Alabama
Arizona
California
Colorado
Connecticut
Delaware
District of Columbia
Florida
Georgia
Illinois
Iowa
Kansas
Maryland
Michigan
New Jersey
New Mexico
New York
Pennsylvania
Texas

total killings by city per year




incidents with at least one dead and a age of 15-20
State
Alabama
Alaska
Arizona
Arkansas
California
Colorado
Connecticut
Delaware
District of Columbia
Florida
Georgia
Hawaii
Idaho
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Massachusetts
Michigan
Minnesota
Mississippi
Missouri
Montana
Nebraska
Nevada
New Hampshire
New Jersey
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
950
1000
N Killed


New Mexico
New York
North Carolina
North Dakota
Ohio
Oklahoma
Oregon
Pennsylvania
Rhode Island
South Carolina
South Dakota
Tennessee
Texas
Utah
Vermont
Virginia
Washington
West Virginia
Wisconsin
Wyoming
0
50
100
150
200
250
300
350
400
450
500
550
600
650
700
750
800
850
900
950
1000
N Killed

Total killed per capita in select states:

- These numbers were tabulated by hand using population data from the Census Bureau and gun data from our Kaggle data set.
    1. California:
        - 2014: 3.419 killed
        - 2015: 3.093 killed
        - 2016: 3.229 killed
        - 2017: 3.615 killed
    2. Texas:
        - 2014: 4.057 killed
        - 2015: 4.135 killed
        - 2016: 4.703 killed
        - 2017: 4.131

Total killed per capita in select states continued:

1. Florida:
   - 2014: 4.398 killed
   - 2015: 3.938 killed
   - 2016: 4.856 killed
   - 2017: 4.531 killed
2. Illinois:
   - 2014: 4.943 killed
   - 2015: 5.350 killed
   - 2016: 7.378 killed
   - 2017: 7.504 killed
3. Maryland:
   - 2014: 4.935 killed
   - 2015: 6.749 killed
   - 2016: 7.146 killed
   - 2017: 7.968 killed

Total killed per capita in select cities:

1. Chicago:
   - 2014: 15.181 killed
   - 2015: 15.857 killed
   - 2016: 21.754 killed
   - 2017: 21.172 killed
2. Houston:
   - 2014: 9.679 killed
   - 2015: 12.768 killed
   - 2016: 12.556 killed
   - 2017: 11.395 killed
3. Baltimore:
   - 2014: 27.582 killed
   - 2015: 42.408 killed
   - 2016: 42.841 killed
   - 2017: 49.632 killed

Total killed per capita in select cities continued:

1. Los Angeles:
   - 2014: 6.035 killed
   - 2015: 3.300 killed
   - 2016: 2.598 killed
   - 2017: 3.496 killed
2. Jacksonville:
   - 2014: 11.612 killed
   - 2015: 11.434 killed
   - 2016: 12.151 killed
   - 2017: 14.461 killed
3. San Antonio:
   - 2014: 8.150 killed
   - 2015: 6.420 killed
   - 2016: 11.358 killed
   - 2017: 6.948 killed
4. Oakland:
   - 2014: 20.586 killed
   - 2015: 21.520 killed
   - 2016: 18.292 killed
   - 2017: 14.138 killed

State laws

- We had a difficult time gathering data on how many laws each state really has on its books.
- States with the most laws such as California, Illinois, Maryland, etc. still saw high levels of gun violence, specifically in major cities.
- On the flip side there are states like Texas, Florida, Alabama, etc. that have significantly less regulation and gun laws on their books and still have high levels of gun crime, also in the major cities.

Under reporting, duplicates, suicide

- Most of our data was specifically geared to events other than suicide.
- We tried to weed through to show events that were actually violent.
- In this data set, suicide was specifically left out by the creators unless it was the result of an incident after a police officer had shown up on scene.
- Fairly certain we have no duplicates.
- A lot more states are under reporting numbers or categorizing them as something else.

Conclusion:

- There is a different problem here then just guns, and more gun laws have not made any significant impact on crime.
- There is an argument to be made than more gun laws have made the problem worse.
- Prescriptively we believe more gun laws will only hurt the American people. We have the second amendment for a reason and should err on the side of extreme caution when proposing legislation that infringes on our second amendment rights when there are other problems contributing to the gun crisis in America.